Analysis and modeling of the radio channel for secret key generation

Taghrid Mazloum

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2016-ENST-0012

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Communications et Électronique »

présentée et soutenue publiquement par

Taghrid Mazloum

le 12 Février 2016

Analyse et modélisation du canal radio pour la génération de clés secrètes

Directeur de thèse : **Professeur Alain Sibille**

Jury

Mme Martine Liénard, Professeur, Université de Lille	Rapporteur
M. Claude Oestges, Professeur, Université Catholique de Louvain	Rapporteur
M. Bernard Uguen, Professeur, Université de Rennes1	Examinateur
M. Benoit Denis, Ingénieur de recherche, CEA-Leti	Examinateur
M. Jean-Pierre Barbot, Maître de Conférence, ENS Cachan	Examinateur
M. Alain Sibille, Professeur, Télécom ParisTech	Directeur de thèse
M. François Delaveau, Ingénieur expert, Thalès	Invité

T H È S E

TELECOM ParisTech
Grande école de l'Institut Mines-Télécom - membre fondateur de ParisTech
46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2016-ENST-0012

EDITE - ED 130

ParisTech Doctorate

T H E S I S

To obtain the Doctor's degree from

TELECOM ParisTech

Speciality « Communications and Electronics »

presented and defended publicly by

Taghrid Mazloum

12 February 2016

Analysis and Modeling of the Radio Channel for Secret Key Generation

Supervisor : **Professor Alain Sibille**

THÈSE

Jury

Mrs Martine Liénard, Professor, Université de Lille	Reviewer
Mr Claude Oestges, Professor, Université Catholique de Louvain	Reviewer
Mr Bernard Uguen, Professor, Université de Rennes1	Examiner
Mr Benoit Denis, Research Engineer, CEA-Leti	Examiner
Mr Jean-Pierre Barbot, Associate Professor, ENS Cachan	Examiner
Mr Alain Sibille, Professor, Télécom ParisTech	Supervisor
Mr François Delaveau, Expert Engineer, Thalès	Invited member

TELECOM ParisTech
Grande école de l'Institut Mines-Télécom - membre fondateur de ParisTech
46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

Résumé

La sécurisation des communications sans fil prend une importance croissante aussi bien pour les besoins des personnes privées que les besoins professionnels et institutionnels. Bien que la cryptographie symétrique assure largement la confidentialité des données, elle est pénalisée par la génération et la distribution de clés secrètes. Des études récentes indiquent que les caractéristiques intrinsèques de la propagation peuvent être exploitées afin de renforcer la sécurité. En effet, le canal radio fournit une source d'aléa partagée par l'émetteur et le récepteur, à partir de laquelle des clés secrètes peuvent être générées. Cette méthode de génération de clés secrètes ("secret key generation", SKG) repose sur la réciprocité du canal de transmission, sur la richesse de la propagation multi-trajets et sur la décorrélation spatiale des caractéristiques du canal.

Dans ce manuscrit, nous nous intéressons à la SKG, avec comme objectif de relier les propriétés du canal radio à la qualité des clés générées. Le travail mené a permis d'analyser l'impact du canal sur la performance de la SKG en relation avec le nombre de degrés de liberté (DOF), dans différentes conditions en particulier dans des environnements statiques. Nous avons développé un modèle statistique simplifié du canal qui montre une mémoire spatiale résiduelle bien au-delà d'une distance de quelques longueurs d'onde (scénarios spatialement non-stationnaires). Puis, nous avons étudié des canaux plus réalistes en environnements extérieur et intérieur (respectivement grâce à des données déterministes simulées et à des mesures) et mettant en évidence l'effet de la dispersion du canal dans des domaines délai et angulaire. Les résultats montrent que, même pour des bandes modérées (compatibles avec les normes IEEE 802.11a, g/n/ac), le seul DoF fréquentiel ou son association avec le DoF spatial est souvent suffisant pour générer des clés longues, à condition d'utiliser une méthode efficace de quantification des coefficients complexes du canal. La qualité de la clé est en outre évaluée après les étapes postérieures à la SKG, i.e. la réconciliation et l'amplification de confidentialité, qui jouent un rôle très important pour aboutir à un schéma robuste et sécurisé.

Abstract

Nowadays, the security of ubiquitous wireless communications becomes more and more a crucial requirement for private, professional and institutional needs. Even though data is widely protected via symmetric ciphering keys, a well-known difficulty is the management of such keys, both regarding their generation and their distribution. In the recent years, a set of works have addressed the exploitation of inherent characteristics of the fading propagation channel towards security. In particular, secret keys could be generated from the wireless channel, considered as a shared source of randomness, available merely to a pair of communicating entities. This approach of secret key generation (SKG) relies on the reciprocity property of the transmission channel, on the richness of the multipath propagation and on the spatial decorrelation of channel characteristics.

In the present dissertation, we are interested in relating the radio channel properties to the quality of the generated keys obtained from SKG. Accordingly, we analyzed the channel degrees of freedom (DoF) and their impact on the SKG performance in different channel conditions, especially in static environments. We first developed a simple stochastic channel model, focusing on the security with respect to the eavesdropper side, which appears to be impacted by a residual channel memory well beyond a few wavelengths distance (in spatially non-stationary scenarios). Then, we investigated more realistic channels in both outdoor and indoor environments (respectively from simulated ray tracing data and through measurements) and highlighted the effect of channel dispersion in the delay and angular domains. The results show that, even for moderately wide band (such as standardized in IEEE 802.11a, g/n/ac), the sole frequency DOF or its association with the spatial DOF is often enough for generating long keys, provided an efficient quantization method of the complex channel coefficients is used. The key quality is further assessed after the subsequent steps of the SKG scheme, i.e. information reconciliation and privacy amplification, which turn out to play an essential role in achieving a robust and secured scheme.

Acknowledgments

If I have seen further it is by
standing on the shoulders of giants.

Isaac Newton

This work would not have been possible without the help and support of many people to whom I am really grateful.

First and foremost, I am deeply grateful to my advisor, Prof. Alain Sibille, for his time, support, guidance and encouragement. Throughout my PhD journey, I have always been inspired by his invaluable advice, immense knowledge, rigor and motivation. He taught me to never give up and to attack the problem from different sides. It has been, simply, a true privilege to be his student.

I would also like to sincerely thank the Thales engineering team, Renaud Molière, Christiane Kameni Ngassa, Claude Lemenager, and in particular François Delaveau, for their collaboration and meaningful discussions.

I extend my gratitude to honorable Prof. Martine Liénard and Prof. Claude Oestges for their review of my thesis manuscript and valuable comments. Moreover, my thanks go to Prof. Bernard Uguen, Dr. Benoit Denis and Dr. Jean-Pierre Barbot, for agreeing to be part of my thesis evaluation committee.

I would like to thank all members of the Comelec department at Telecom Paris-Tech for the excellent scientific environment they provide. Special thanks go to Dr. Christophe Roblin for his advice, especially when preparing to my PhD defense.

To my colleagues and friends, thank you for listening and supporting me especially during the stressful moments. Special thanks go to Amal Abdul-Razzak for her friendship as well as for every single moment we have shared in Paris and elsewhere.

Last, but not the least, I owe my tremendous gratitude to my parents for their love, moral support and constant encouragement. All my loving thanks to my mother

and my father who believed in me. I am eternally grateful to you, my parents, to my sisters and my brothers. Special thanks to my brother Youssef for sharing all the sad and happy moments in Paris.

Contents

List of Figures

List of Tables

List of abbreviations

AoA	Angle of arrival
AoD	Angle of departure
ApEnt	Approximate entropy
APS	Angular power spectrum
BER	Bit error rate
BS	Base station
BST	Base station transmitter
BW	Bandwidth
CDF	Cumulative distribution function
CIR	Channel impulse response
CQA	Channel quantization alternating
CSI	Channel state information
DoF	Degree of freedom
DS	Diffuse scattering
EM	Electromagnetic
FDD	Frequency division duplex
FFT	Fast Fourier transform
GSCM	Geometry-based stochastic channel model
i.i.d.	independently and identically distributed
KDC	Key distribution center
K-S	Kolmogorov-Smirnov
LOS	Line of sight
MIMO	Multiple-input multiple output
MKSM	Multiple key single map
MPC	Multipath component
MS	Mobile station
NB	Narrow band
NIST	National Institute of Standards and Technology
NLOS	Non line of sight
OFDM	Orthogonal frequency division multiplexing

PDP	Power delay profile
PHYLAWS	Physical layer wireless security
PhySec	Physical security
PSD	Power spectral density
QM	Quantization map
QR	Quantization region
QuaDRiGa	Quasi deterministic radio channel generator
RL	Ray launching
RMS	Root mean square
RMS-DS	Root mean square delay spread
RSS	Received signal strength
RT	Ray tracing
SIMO	Single-input multiple-output
SKG	Secret key generation
SKSM	Single key single map
SNR	Signal to noise ratio
SSF	Small scale fading
TDD	Time division duplex
TPT	Telecom ParisTech
UWB	Ultra wide band
VNA	Vector network analyzer
WB	Wide band
WDS	Without diffuse scattering
WSSUS	Wide sense stationary uncorrelated scattering

Chapter 1

Introduction

1.1 General motivations

Given the growing prevalence of wireless communications, their security becomes a major concern for various applications such as broadband internet, e-commerce, bank services, health monitoring, terrorism and military operations. Indeed, the broadcast nature of the wireless medium makes it vulnerable to various attacks, e.g. eavesdropping, man-in-the-middle attack, etc. Although classical security mechanisms are widely used to protect data transmission, they present many challenges showing a lack of confidentiality. This lack is proved by e.g. the monitoring of Angela Merkel's smartphone during years [2]. Therefore, trends to strengthen security are going towards new security paradigm acting at the physical layer [3, 4].

Traditionally, a set of protocols provides security (e.g. confidentiality, authentication, integrity) by encrypting data using cryptographic keys. In symmetric encryption methods, the main drawback is the key management, which includes key generation and distribution, since the same secret key is used for both data encryption and data decryption [5]. More clearly, challenges stem from, on one hand, the need to share a prior key between each legitimate user and a key distribution center (KDC) which is responsible for delivering secret keys. On the other hand, the KDC may not be easily accessible in scenarios such as dynamic mobile networks and wireless sensor networks.

However, these symmetric key management issues are alleviated by asymmetric techniques where a pair of public and private keys, generated by each user, is used to secure communications [5]. Nevertheless, the usage of these public-key cryptosystems is restricted for covert key exchange through the Diffie-Hellman algorithm [6] since

they suffer from high computational cost. Moreover, the robustness of conventional cryptography relies on computational constraint on the attacker. However, with the continuous progress of high power computing, unconditionally secured systems are more and more required [4].

Recently, several researches investigate information-theoretic security techniques where the illegitimate user (let us call here Eve) of a wireless communication is assumed to be enabled with unlimited computing power and only her information about the propagation scenario may help her to break the data privacy [4, 7]. In this respect, a special approach to physical layer security (PhySec) field [4] intends to achieve wireless communications and data protection by exploiting the inherent properties of the wireless propagation channel such as reciprocity, multipath fading and noise. On one hand, security over the radio channel may be achieved through secrecy coding schemes (e.g. LDPC, Lattice and polar codes [8, 9, 10]) that require advantages at the legitimate part over the eavesdropper part, such as the signal to noise ratio (SNR). On the other hand, for cryptographic purposes, PhySec enables legitimate parties to generate independently identical secure key bits from the reciprocal radio channel. Both of these approaches, often combined with beamforming techniques and artificial noise [11, 12, 13, 14], are targeted in the European project PHYLAWS (i.e. physical layer wireless security) [15], which supports the present PhD thesis work.

1.2 Brief history of physical layer security

Information-theoretic security was firstly introduced by Shannon in 1949 [16], where a secrecy system had been defined in relation with Shannon's previous work in information theory. Shannon proposed to use the one-time pad, where messages are encrypted using the binary addition (XOR), in order to achieve perfect secrecy, where the codeword conveys strictly no information about the initial message. Although no restriction on the computational power of Eve was adopted, she was assumed to have no useful information to break the data confidentiality. We note that this perfect secrecy is impractical since it relies on unrealistic assumption where Eve can obtain the codeword (i.e. the encrypted message) without any error [4], while in contrast, the radio channel is corrupted by intrinsic noise.

In 1975, Wyner [3] revisited Shannon's information theoretic secrecy by considering a realistic assumption, i.e. Eve obtains noisy observations owing to the intrinsic noise of the wireless propagation. This implied the definition of the secrecy coding

concept through a proposed channel model called the "wiretap channel". Accordingly, the protection of data entails that the channel seen by Eve should be degraded with respect to legitimate entities channel, in other words, the Alice/Bob channel capacity should be greater than that of Eve. Otherwise, wireless communication cannot be secure.

Alternatively, in 1993, Ahlswede, Csiszar and Maurer [17, 18] proposed to distil a shared secret key from a shared source of randomness, even if Alice ad Bob do not have an advantage over Eve. This may be achieved through public discussions over an error-free authenticated channel. The channel advantage may be provided by exploiting the reciprocity law and the spatial decorrelation (intrinsic properties of the electromagnetic channel, other than the SNR) in order to distil a shared secret key from the channel randomness instead of using secrecy codes. The resulted secret key may provide data confidentiality through e.g. conventional symmetric encryption methods.

1.3 Thesis Context

In this work, we are mainly interested in wireless channel-based secret key generation (SKG), which seems to be an efficient alternative to conventional key distribution. Authorized users jointly generate a common secret key from a shared source of randomness, which is not directly accessible to an eavesdropper. This theoretical foundation goes back to Ahlswede, Csiszar and Maurer in 1993 [17, 18]. While the first implementations of SKG concern quantum physics [19, 20], their application is limited owing to many challenges, e.g. high cost. Hence, nowadays much attention goes towards the ubiquitous wireless channel as a shared source of random secret keys. Practical implementations have early been introduced by Hassan et al. [21, 22] by establishing secret keys from the information phase of a frequency-selective fading channel. Fundamentally, when channel reciprocity applies, typically when legitimate parties use the same frequency at the same time instant, they share the same propagation channel. Randomness is ensured through multipath fading, which results in decorrelation properties in the spatial, temporal and frequency domains. Consequently, an eavesdropper is probably not able to efficiently exploit her own measured channel in order to crack the key.

While channel reciprocity is a crucial requirement for key reliability, i.e. ensuring the same key for both legitimate users (typically referred to as Alice and Bob), it

does not hold perfectly in practical scenarios owing to electronic hardware differences between transceivers. Moreover, even in time-division duplex (TDD) systems where the same frequency is used, the forward and reverse channels may differ from each other when the environment changes between the two channel estimation instances, in other words when the channel estimation is not occurring within the same coherence time. All these issues, including noise, limit in practice the number of shared bits that may be reliably extracted from a single channel observation. Therefore, a secret key of a sufficient length generally results from the concatenation of several sub-keys, obtained from multiple channel observations. Subsequently, especially in time-varying channels, an issue related to the latency in the key generation process appears, which requires further investigations in order to harvest more randomness from a single channel estimation.

Furthermore, the random character of the key is essential in making eavesdropping extremely difficult or, on the other hand, which requires a small correlation between the channel samples seen by Alice/Bob and by the eavesdropper. This also entails statistical independent channel samples. Indeed, key randomness may be achieved through random channel variations, which occur inherently in small scale fading (SSF) owing to constructive/destructive combination of dispersive multipaths. This may be heavily accomplished by user movement, e.g. when using a smartphone or a connected watch emerging from the internet of things. Obviously, if we consider the case of a fixed laptop in an office, some channel variability may be provided if people are moving around, but experiment shows that the degree of channel variation incurred by such changing environment is quite weak, unless of dense crowds. So, a general questions is: is the randomness of the key sufficiently guaranteed? If not, how may the security performance be improved in static environments? Moreover, how much correlated information can Eve obtain about the legitimate channel and subsequently about the key?

1.4 Thesis objectives

The work reported in this PhD thesis takes place in the context of the European project PHYLAWS [15], which aims to improve the protection and confidentiality of wireless communications through approaches operating at the physical layer level, mainly. PHYLAWS intends to identify and test security approaches in both theoretical and experimental manners in order to elaborate efficient techniques that are simple to implement and consume few resources. The project outputs will thus benefit

a variety of existing and future standards for a large set of needs.

PHYLAWS intends to investigate several techniques able to enhance the level of security. Among them, the secrecy coding scheme assumes a message to be protected by exploiting the channel noise instead of using encryption keys. Hence, the "secrecy capacity" can be defined by the number of message bits that could be transmitted securely and reliably, per channel use, between legitimate users while Eve is unable to decode it. This approach requires an advantage on the legitimate link over the eavesdropper link, otherwise, the secrecy capacity is null. Along this line, artificial noise and jamming techniques are targeted within PHYLAWS project, in order to confuse the attacker and sustain legitimate users.

Alternatively, PHYLAWS considers another aspect of PhySec, which is SKG from common randomness provided by reciprocal radio channels, as introduced above. Since the implementation of PhySec techniques, especially SKG, will be deeply impacted by the radio channel characteristics, it is crucial to assess the main limitations of this technology in realistic scenarios. Therefore, the main objective of this PhD thesis is to study the fundamental role of multipath fading channels in security and, subsequently, to relate the characteristics of the radio channel to the security of the SKG process. Moreover, we intend within PHYLAWS to devise adequate channel models, which imply specified measurements or simulations, where SKG behavior is assessed either using information-theory or through the practical implementation of a key agreement scheme.

1.5 New contributions

To summarize, the general aforementioned limitations in SKG concern the quality of the key (i.e. reliability, randomness, size and secrecy), especially when SKG is distilled from time-variant channels and in static environments. Accordingly, this dissertation targets the strategy of generating shared keys from a reciprocal radio channel that is seen as a shared source of randomness available at legitimate parties. The main objective is to assess SKG performance with respect to realistic propagation channel features. This is achieved through the usage of several channel models, e.g. stochastic or deterministic, accounting for Alice/Bob/Eve locations and features. In this context, the contributions of the present dissertation are:

1. Investigate the channel degrees of freedom (DoF) existing in either the space or the frequency/delay domain, even in the joint space-frequency domain, in

the generation of shared secure keys [23, 24]. Intuitively, such an investigation intends to harvest more randomness from a single channel observation, and thereby, to alleviate the issue of the lack of time-variability in static environments.

2. Analyze the quality of the generated keys in relation with the channel properties (delay spread, angular spread, shadow fading, LOS/NLOS, etc.) in different propagation environments including both indoor and outdoor. The key robustness is assessed either in terms of size of the generated keys or of difficulty/impossibility for Eve to reconstruct Alice/Bob's key.

3. Define and "play" with a parameter based stochastic channel model, in order to gain deep insight on the relation between SKG performance and physical characteristics of the propagation and of the environments. Specifically, we develop a multi-link channel model to account for the spatial correlation between channels seen by Bob and Eve, covering both stationary and non-stationary regions.

4. Use deterministic channel models "close to reality" for sets of outdoor environments of main interest, and see how much the channel characteristics investigated impact SKG performance.

5. Carry out indoor channel measurement campaigns, which will allow to implement SKG in different environments and settings, considering varying separation distances between users on one hand and LOS/NLOS propagation conditions on the other.

6. Analyze and compare the robustness of the generated key after each phase of SKG (i.e. quantization, information reconciliation and privacy amplification), while considering different approaches to compute the quantization maps.

1.6 Organization of the dissertation

Following the brief above introduction designed to motivate the main topic of the dissertation, **Chapter 2** provides in a first part an overview of the radio propagation channel since it heavily impacts the security level bring by PhySec approaches. In the second part, we describe the principle of wireless channel-based secret key generation and review the prior works in the relevant area of SKG.

We explain in **Chapter 3** both the metrics and the methods used in the the present work, to assess the robustness of the generated keys. In particular, we describe the quantization algorithm, which maps the complex channel coefficients into a stream of key bits. Moreover, we explain how to investigate the channel variability in either the space or the frequency domain in order to increase the key rate per single channel observation, with a specific emphasis on the delay dispersion through an exponential decaying power delay profile (PDP).

A simple channel model is then developed in **Chapter 4** and focuses on the amount of information disclosed to Eve, which may cover several scenarios, including stationary and non-stationary ones with respect to Bob-Eve distance. To adequately model the spatial correlation between Bob and Eve, the proposed geometry-based stochastic channel model employs a per path shadow fading correlation [25].

SKG, in relation with the exploitation of different types of channel degrees of freedom, is investigated for deterministic outdoor channel models in **Chapter 5** and for measured indoor propagation in **Chapter 6**. The security is addressed from Alice/Bob point of view, as well as from Eve point of view.

While we emphasis, in the aforementioned chapters, on the SKG performance right after the channel quantization phase, we subsequently dedicate **Chapter 7** to the full implementation of the SKG scheme, including reconciliation and privacy amplification, and analyze the performance gain with respect to the sole quantization. In addition, we investigate the impact of the way the quantization map is computed.

Finally, in **Chapter 8**, we summarize the work of the thesis and we draw some conclusions after the comparison of the SKG behavior in different environments.

Chapter 2

State of the art on secret key generation from random channels

In the context of wireless communications, the propagation channel plays an important role towards the improvement of security services at the physical layer level. A deep understanding of this communication medium is thus crucial. Hence, the first section of this chapter is dedicated to a brief overview of the main notions regarding the propagation channel, especially those relevant to PhySec. The second section subsequently provides a state of the art about the propagation-based secret key generation, including a description of the different steps of the SKG strategy.

2.1 Fundamentals of wireless propagation channels

A radio communication system, as depicted in Fig. 2-1, aims to convey a message between two remote users through a propagation channel. At the transmitter side, the message is encoded into an electrical signal which is suited to efficient transmission and radiated into space by the antenna, in the form of an electromagnetic (EM) wave. On the other side, at the receiver, the original information is reproduced from the received EM waves, assuming it can be decoded without errors. The medium through which these signals propagate between the transmit and the receive antenna is called the wireless propagation channel. The wireless transmission channel is defined as the combination of the propagation channel and of the antennas. Fig. 2-2 depicts the difference between the transmission and the propagation channels.

The time-variant received signal $y(t)$ is the convolution of the time-variant trans-

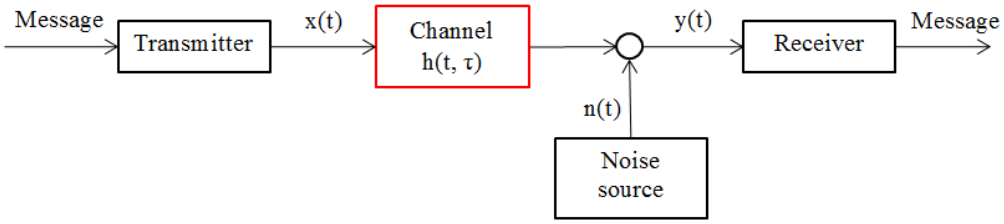


Figure 2-1: Communication system.

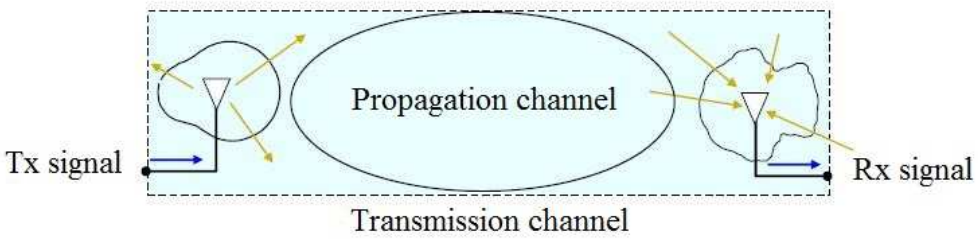


Figure 2-2: Transmission vs. propagation channel.

mission channel response $h(t, \tau)$ and the emitted signal $x(t)$, expressed as follows:

$$y(t) = (h * x)(t) + n(t) \tag{2.1}$$

where $*$ denotes for convolution. τ is the time delay. The received signal is affected by an additive noise $n(t)$, which may be generated inside the receiver itself or caused by some radio interferences. $n(t)$ is often approximated as a white Gaussian random variable.

Before reaching the receiver antenna, the emitted signal undergoes several physical interactions with the environment such as specular reflection, edge diffraction and diffuse scattering, as shown in Fig. 2-3. Hence, the received signal is the superposition of several replicas of the transmitted signal where each partial wave is attenuated, delayed and phase shifted according to the traveled distance as well as to the type of the physical interaction. This wave propagation phenomenon is referred to as *multipath propagation*. Accordingly, the propagation channel impulse response may be approximately described as:

$$h(t, \tau) = \sum_{n=1}^{N_p} \beta_n \delta(\tau - \tau_n) \tag{2.2}$$

where β_n and τ_n are respectively the complex gain and the delay of the nth partial

wave. N_p is the total number of paths occurring in the channel. This discrete multipath channel model neglects the frequency dependence of the interactions mentioned above and is specially valid for narrow or moderate bands.

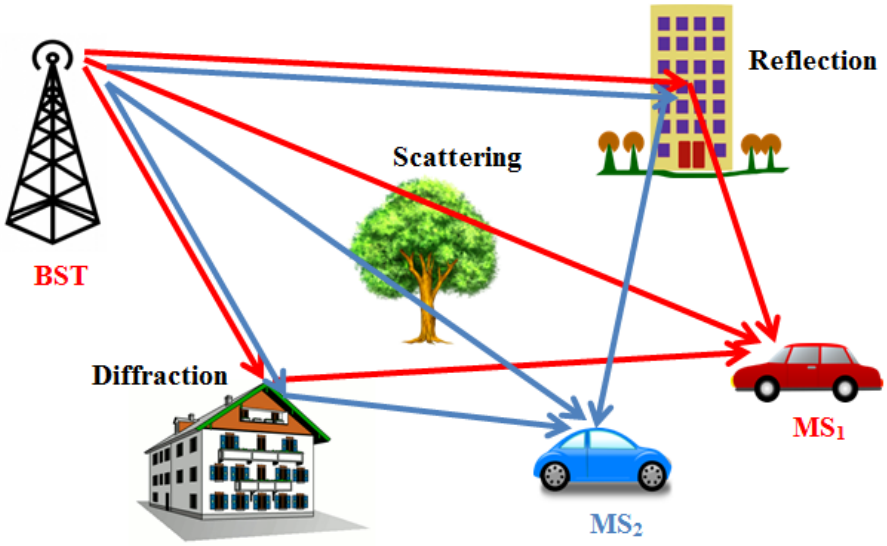


Figure 2-3: Location-specific multipath propagation.

2.1.1 Channel reciprocity

The channel impulse response measured between two transceivers is the same regardless of the direction of transmission, since EM waves undergo the same physical interactions in both directions and there are usually no magnetic (non reciprocal) interactions. This is known by the channel "reciprocity" property, which holds in TDD systems where the same frequency band is used for both uplink and downlink [26, 27]. One benefit of such a property in wireless communications is, e.g., intended to enhance transmission efficiency such as throughput by providing channel state information (CSI) at the transmitter side without additional feedback [28]. Recently, the reciprocity has been exploited in PhySec in such a manner that the shared CSI between two entities is exploited to protect exchanged data against an eavesdropper [22], in particular by extracting identical encryption keys.

However, establishing radio reciprocity may be limited by some practical issues. In fact, in TDD systems, the channel is required to remain invariable during the channel estimation phase in both directions. In other words, the channel should be estimated during the coherence time in order to reduce channel discrepancies.

Moreover, the asymmetric radio-frequency electronic hardware in the transmit and receive communication chains may break the reciprocity property [29, 30], owing to the unidirectional characteristics of certain components. This challenge is mostly addressed by performing calibration [31, 32, 33] and has been tested for SKG purposes in [34]. Nevertheless, although reciprocity is not valid in frequency-division duplex (FDD) systems, there have been attempts to mitigate its lack in such systems through alternative approaches, for example by using a frequency correction algorithm [35]. However, for instance, the reciprocity in FDD systems is not sufficiently resolved to sustain SKG algorithms.

All over the thesis, we concentrate on a communication scheme appropriate for perfect reciprocity or nearly so, such as TDD. Then, any two entities (who may be legitimate terminals) may privately share common information extracted from the reciprocal channel, from which secure key bits may be established in order to encrypt data and thereby protect wireless communications. More clearly, because of the absence of an explicit feedback from receiver to transmitter, a third party (who may be a malicious attacker) has no access to the shared information considered as a source of secret keys, unless she exploits different methods, e.g. her own measurements.

2.1.2 Path loss, shadowing and small scale fading

Fig. 2-4 visually depicts the spatial variation of a propagation channel, which is often expressed in terms of three physically identified phenomena according to the spatial scale, i.e. the long distance path loss, the shadow fading and the small scale fading [36].

Path loss is explained by the average attenuation of the power of the EM wave with the transmitter-receiver separation distance d as the wave propagates through space. If we consider propagation between two isotropic antennas in free space, the signal experiences the following path loss (commonly known as the Friis transmission equation) in dB:

$$PL_{fs} = 20 \log_{10} \frac{4\pi d}{\lambda} \tag{2.3}$$

where λ is the wavelength. Besides the distance, the path loss depends also on the environment, the propagation medium and on the antenna location. Indeed, it is largely impacted by interaction effects such as reflection, refraction and diffraction, in addition to the minimal free-space loss. Accordingly, all these elements contribute to define, based on measurements and theoretical analysis, an empirical model for the

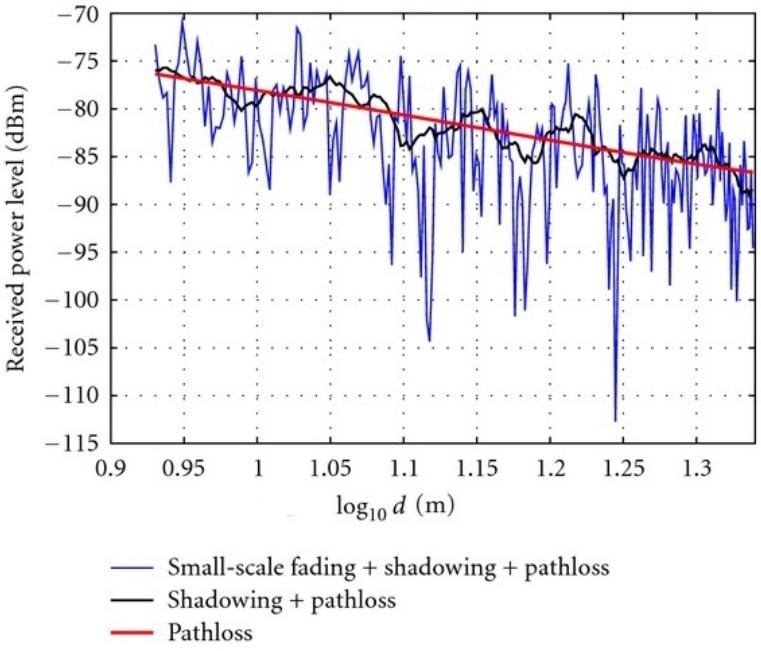


Figure 2-4: Path loss, shadowing and small scale fading.

average large scale path loss as follows:

$$PL(d) = PL(d_0) + 10\gamma \log_{10} \frac{d}{d_0} \tag{2.4}$$

where d_0 is a reference distance. γ is the path loss exponent that expresses how fast the path loss increases with distance. This parameter strongly depends on the propagation environment. According to measurements, the value of γ is in the range of 2 to 6. While $\gamma = 2$ characterizes free space, we can found $\gamma \leq 2$ for scenarios presenting waveguide effects such as corridors. Moreover γ can reach the value 6 in scenarios where there is a severe attenuation, e.g. terminals separated by multiple floors/walls in an indoor environment.

Although the path loss changes mainly with the distance, two terminals at a fixed distance from a base station do not receive the same power since they may not receive the same dominant paths. In fact, according to the local surrounding media, some paths may be blocked by obstacles such as building or hills for only one receiver while they still reach the other. This phenomenon is known as shadowing where random slow power variation occurs around the mean path loss. Shadowing is often modeled by a log-normal random variable with zero-mean and a local standard deviation, usually in the range of 3-10 dB [37]. Shadowing is also known as medium-scale fading

since the changes in the received power occur for distances in the range of tens of wavelengths or more.

Furthermore, the received power still changes randomly for smaller distances (on the order of the signal wavelength) owing to the multipath propagation where the combination of all partial waves may interfere constructively or destructively. The changes in the interfering combination over a short distance is explained by the phase shift experienced by the different paths when the receiver moves. This multipath propagation effect is called either small-scale fading or fast fading. It is noteworthy that the amplitude fluctuation is rapid and is often modeled either by a Rayleigh distribution or by a Rician distribution. The first distribution describes environments where interference is caused by a large number of paths uniformly distributed in space and having almost the same average power. However, in the case where a path (generally, but not necessarily, the line of sight (LOS)) dominates the propagation channel, the amplitude variation is better described by the Rician distribution.

While the path loss variation is deterministic owing to the dependence with the distance, the multipath fast fading is a random process. Therefore, the fading channel and the related differing physical mechanisms will make it very difficult for an eavesdropper to reconstruct Alice-Bob's channel accurately, especially for rich scattering environments, if she is located at some distance from the legitimate terminals (Fig. 2-3).

2.1.3 Channel degrees of freedom

In conventional communication system, the small scale fading is detrimental because the receiver is unable to correctly reproduce the original transmitted message when the channel is in a deep fade resulting from destructive multipath interference. In order to mitigate this effect, the transmitter sends several time the same message over the time-varying channel, until the fading turns toward constructive interference where the message is correctly received. This means that the message should be retransmitted after the channel *coherence time*, which is a statistical measure of the time duration over which the channel response is considered invariant. In other words, the coherence time is typically the time interval within which two channels measured over two different time instances are highly correlated. The coherence time is inversely proportional to the Doppler frequency, which is related to the rapidity of channel variations.

Fortunately, SSF may be efficiently mitigated by exploiting the channel degrees of

freedom (DoF) existing either over the frequency domain or in the space domain. The idea consists of providing statistically independent multiple parallel fading channels, which depends mainly on the richness on multipaths and their dispersion. Equivalently, multiple DoF may bring improvements to the performance of a communication, such as the channel capacity through multiplexing gain in multiple input/multiple output - MIMO - techniques. In the frequency domain case, the channel correlations are related to the dispersion of the propagation paths in the time domain. This can be described by the root mean square (RMS) delay spread, computed as follows:

$$\sigma_\tau = \sqrt{\frac{\int_0^{\tau_{max}} (\tau - \bar{\tau})^2 P(\tau) d\tau}{\int_0^{\tau_{max}} P(\tau) d\tau}} \tag{2.5}$$

and the mean delay $\bar{\tau}$ is defined as follows:

$$\bar{\tau} = \frac{\int_0^{\tau_{max}} \tau P(\tau) d\tau}{\int_0^{\tau_{max}} P(\tau) d\tau} \tag{2.6}$$

where τ_{max} is the maximum excess delay and $P(\tau)$ is the average power delay profile (PDP) computed over a set of time-varying CIRs as follows:

$$P(\tau) = E\{|h(t,\tau)|^2\} \tag{2.7}$$

The equivalent of the RMS delay spread in the frequency domain is the channel *coherence bandwidth*, which provides a statistical measure of the range of frequencies over which the channel undergoes a similar fading. In other words, the coherence bandwidth is the frequency interval within which two sub-carriers are highly correlated. We note that the coherence bandwidth is inversely proportional to the RMS delay spread, although there is no unique relation between both.

Equivalently, the space diversity is investigated through the use of multiple antennas, where the parallel channel correlations rely on the angular dispersion of the multipaths. The higher the angular spread, the smaller the channel spatial correlations. The RMS angular spread and the mean angle $\bar{\phi}$ are given respectively in Eq. 2.8 and Eq. 2.9, where ϕ is the azimuth angle and $P(\phi)$ is the average power angular spectrum. We note that this definition for σ_ϕ is specifically valid for small or moderate angular spreads [38]. More general formulas could be found in [38, 39].

$$\sigma_\phi = \sqrt{\frac{\int_0^{2\pi} (\phi - \bar{\phi})^2 P(\phi) d\phi}{\int_0^{2\pi} P(\phi) d\phi}} \tag{2.8}$$

$$\bar{\phi} = \frac{\int_0^{2\pi} \phi P(\phi) d\phi}{\int_0^{2\pi} P(\phi) d\phi} \tag{2.9}$$

2.1.4 Propagation channel models

For a long time, a multitude of propagation channel models has been widely proposed and developed in the literature in order to express the EM propagation characteristics that are involved in the functioning of wireless communication systems, and which can support their evolution (e.g. from 2G to 4G and beyond). Channel models are generally developed in order to be suited to specific applications, and the model is required to be as simple as possible and with a limited number of parameters. Until now, PhySec has not been considered as a scheme of interest for wireless communications. The main peculiar aspect resides in being a three terminal scenario, where the presence of Alice, Bob and Eve has to be modeled in the same propagation environment. Although multi-link channel models have been investigated in the context of relaying [40], a model that is able to address scenarios for three terminals with two in proximity of each other is needed. Additionally, this model must have sufficient accuracy in modeling the spatial correlation between the neighboring terminals.

Many classifications are possible to define channel models. An effective one is presented in [41] and relies on the type of modeling approach. On one side, physical models rely on the EM wave propagation by describing the double-directional multipath structure of the channel. These models are typically antenna independent, i.e. the antenna can be embedded in post-processing. On the other side, non physical analytical models characterize the channel in a mathematical/analytical way without explicitly considering wave propagation.

Physical models

Ray-Tracing (RT) and related techniques such as ray-launching is a deterministic site-specific method that is able to reproduce any channel characteristic and is therefore suitable to be used in PhySec scenarios. Nevertheless, only very few contributions in this context are found in the literature [42, 43]. The main drawback of RT resides in its computational needs, especially when the environment becomes complex. It also requests a very precise database of the geometrical environment, as regards the EM aspects. The model is inherently usable in a multi-user scenario, although the amount of computing need may accordingly increase. A great advantage of RT is its capability to describe long distance channel variations, since it is able to compute the multipath

structure at any Rx-Tx positions. However, there are two major drawbacks: i) Long simulation times are required in the case of a massive computation, especially for covering large areas and for many frequencies. ii) Pure specular path propagation does not fully render the complexity of real channels, e.g. through "diffuse scattering", which stems from the roughness of reflecting surfaces [44]. Conversely, diffuse scattering may be included in RT tools, although this incurs even more computation requirements and very long simulation times.

As opposed to RT, in geometry-based stochastic channel models (GSCM) the locations of scatterers and their characteristics are defined in a statistical way, according to certain probability distributions. GSCMs are less computationally greedy than RT and they can generate a variety of scenarios instead of a specific one. They are able to simulate all channel characteristics, but only in the environment where they have been parametrized. Moreover, only a subset of GSCM is multi-user/multi-link oriented, e.g. Winner II [1] and COST 2100 [45].

The last and simplest, from the computational point of view, category of models that relies on the physical wave propagation are the stochastic models that originate from the Saleh-Valenzuela model [46]. This model relies on a cluster assumption where clusters are defined in a purely statistical way, via cluster delay and intra-cluster ray delay distributions. The amplitude of channel coefficients is Rayleigh distributed and the phase uniformly distributed over $[0,\ 2\pi]$. This has been extended including angular behavior for MIMO scenarios [47].

Analytical models

Among analytical models, we mention the correlation-based models that characterize the MIMO channel matrix statistically in terms of the correlations between the matrix entries. Popular models are the Kronecker model and the Weichselberger model. The Kronecker model describes the propagation channels by separate transmit and receive correlation matrices [48]. However, the Weichselberger model [49] aims at overcoming the restriction of the Kronecker approximation in separating transmit and receive sides and at modeling the correlation properties at the receiver and transmitter jointly.

Non stationary channel models

Non stationary channel models pose another challenge, since there is extra complexity with respect to small scale channel modeling. In addition to modeling the charac-

teristics of interest of the radio channel at a given Tx/Rx location (e.g. multipath structure and parameters), which should typically be done for the variety of environments concerned by the targeted applications, it is needed to model the macroscopic changes experienced by the channel when the device moves (e.g. along a route). This may be useful within a given type of environment (e.g. indoor) or from one type to another (e.g. outdoor to indoor). The channel changes thus need to express the channel variations of the medium scale fading (typically: shadowing), as well as the slower changes (typically: variation in the long distance average attenuation).

The literature is still scarce, although it is a recognized relevant issue, especially now that more and more nomadic wireless devices of all sorts will be used in the coming years (by walking people, in vehicles, etc.). One possible approach is to use a GSCM in order to recover the spatial variability from the placement of scatterers seen from the mobile over its route [50, 51, 52, 53]. A simple manner to model shadowing is to use the lognormal model for the attenuation (i.e. a normal distribution in dB), which allows to generate channel coefficients stochastically from correlation coefficients in a well-known manner for a Gaussian stochastic process [54].

2.2 Wireless channel-based secret key generation

In the recent years, a set of works have addressed inherent characteristics of the fading propagation channel in an information-theoretic framework oriented towards security. In fact, a well-known problem in classical cryptography is the management of ciphering keys, both regarding the generation and distribution of these keys. A way to alleviate such difficulties is to use a common source of randomness for the legitimate terminals, not accessible to an eavesdropper [17, 18]. This is the case of the fading propagation channel, when exact or approximate reciprocity applies [21]. Fig. 2-5 depicts a communication scenario where Alice and Bob wish to communicate securely in the presence of an eavesdropper (Eve). SKG seems to be efficient for several systems, such as local area networks [55], mobile networks, wireless sensor networks [29, 56] including both body area network and emerging internet of things in which the energy consumption is of major importance.

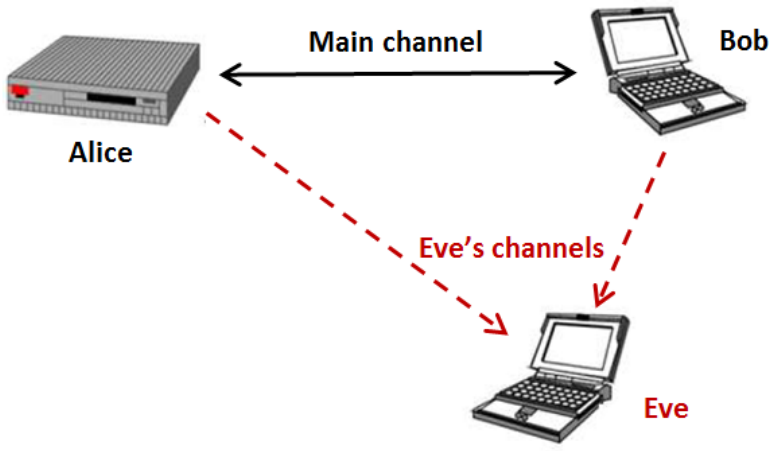


Figure 2-5: Wireless communication scenario.

2.2.1 SKG basics

Establishing secure keys from reciprocal radio channel may be achieved through four consecutive steps: i) channel estimation, ii) channel quantization, iii) key reconciliation, and iv) privacy amplification. An illustrative scheme is depicted in Fig. 2-6.

In a TDD system, such as IEEE 802.11, Alice and Bob estimate their CSI by successively sending each other a known training signal, using the same frequency band. According to Section 2.1.1 and owing to the EM reciprocity law, the CSIs at both Alice and Bob sides are very similar. Therefore, assuming they use a common quantization algorithm, they are able to jointly translate their measured channel information into a shared string of cryptographic key bits, which may be used by the upper-layer protocols in order to consolidate security. We note that the channel estimation phase and the key quantization phase, together provide shared randomness to Alice and Bob. This is equivalent to the "randomness sharing" step of the "sequential key distillation" strategy presented in [4].

Despite that radio channels may not be reciprocal especially when estimated with ordinary commercial devices, we assume in this thesis that the reverse and forward channels are reciprocal and that the slight channel estimation inaccuracies owing to TDD are lumped into channel noise. Therefore, the time-varying channel is estimated as follows:

$$\hat{h}(t) = h(t) + n(t) \tag{2.10}$$

where the time-varying additive noise $n(t)$ is supposed to be a complex Gaussian

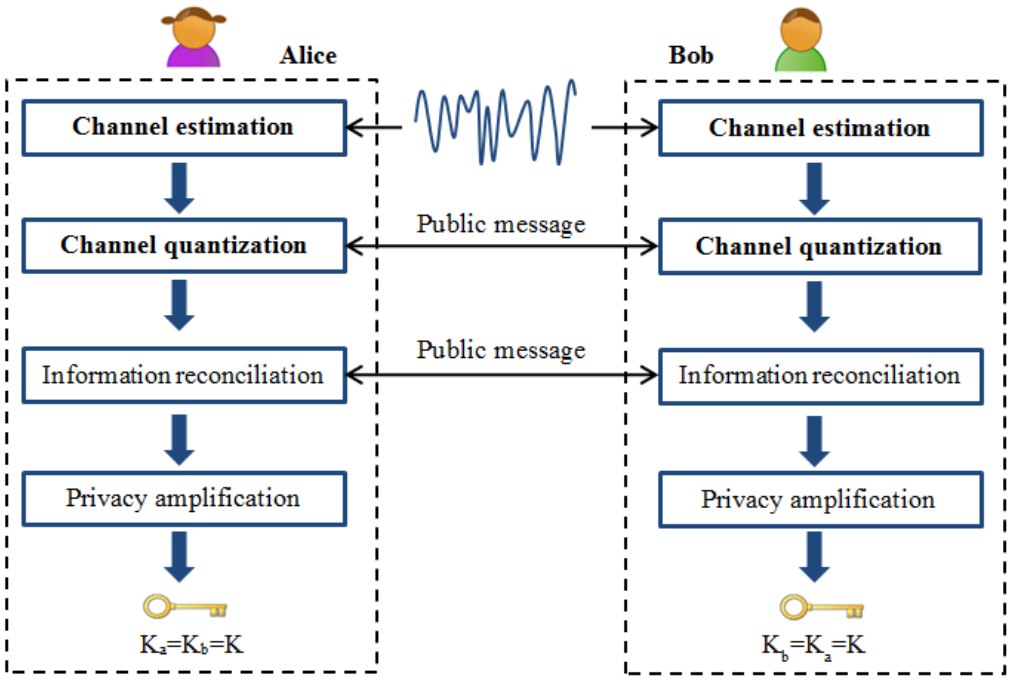


Figure 2-6: The four steps of SKG to agree on the same key.

random process. We also assume that the noise signals of differing terminals are independent.

Clearly, channel noise may lead to mismatches between Alice-Bob keys, subject to the SNR and the quantization scheme. Fortunately, such a key disagreement may be diminished, as a simple first step, through an efficient quantization technique (see Section 3.2.1) employing suitable censoring schemes [57, 58, 59, 60]. These algorithms may need to exchange appropriate public information, without revealing any useful information about the bit value to Eve. Since key disagreement may still occur, a reconciliation step [7] is required where Alice and Bob publicly exchange messages, termed syndromes, in order to agree on the same shared key bits with vanishing error probability. The correction of mismatched bits may involve linear error correcting codes [7, 60, 61, 62, 63, 64, 65], e.g. Hamming code [64], LDPC codes [7, 61, 62, 65] and polar codes [66], as examples. Moreover, key agreement is targeted through the use of cyclic error correcting codes such as BCH code [55].

During the reconciliation phase, Eve attempts to exploit the shared public syndrome in order to guess as many key bits as she can. However, this information may not be relevant, especially when the amount of the public information is small, which in turn means that a small bit disagreement probability results after the quantization phase. Nevertheless, a last step, called privacy amplification [67], seems essential in

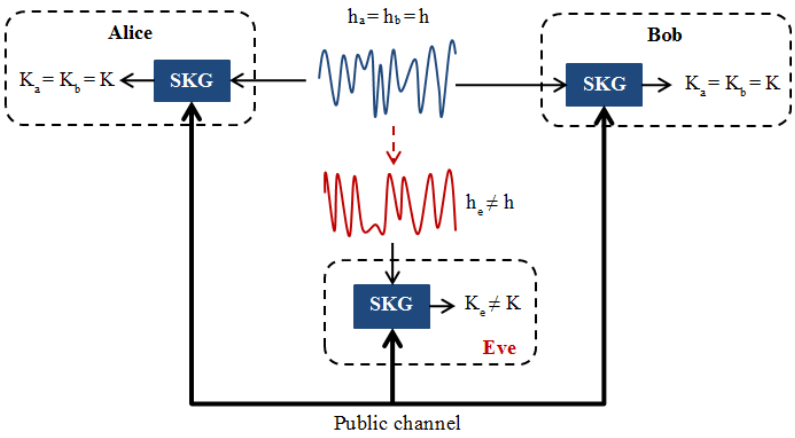


Figure 2-7: SKG in the presence of an eavesdropper.

rendering the key more random through e.g. hash functions, at the cost of a reduced key length. However, indirectly, the key length reduction by privacy amplification means that one of the first success criteria in SKG is the key length, so to have a sufficient length margin. This part of the whole SKG mechanism, i.e. information reconciliation and privacy amplification, is addressed in Chapter 7.

2.2.2 Security versus an eavesdropper

In the present work, we focus mainly on passive attacks where the eavesdropper is able to measure the channel when either Alice or Bob transmits training signals. She is also assumed to know the quantization algorithm used to establish a sequence of key bits. Moreover, Eve is supposed to intercept all the public communications. Therefore, Eve's goal is mainly to exploit all this knowledge in order to collapse the key space and make its search of the exact key easier. The idea underneath SKG from random channels is that Eve's gain of information about the key will be negligible in a typical scenario, rich in dispersive multipaths. In such a scenario, the location-specific propagation channel is expected to rapidly decorrelate in the space or frequency domains, owing to rapid signal fluctuations (Fig. 2-3). Thus, a passive eavesdropper who intercepts the communication between Alice and Bob, experiences a channel significantly decorrelated from that experienced by legitimate users. As a consequence, when Eve attempts to derive a key from the radio channel, the key has few common features with that extracted by both Alice and Bob, making hacking in practice unachievable. Fig. 2-7 depicts such a typical scenario.

However, channel decorrelation depends to a large extent on the nature of the

propagation environment (e.g. multipath richness) as well as on the relative distance between Eve and Alice/Bob. Furthermore, Eve might resort to some techniques in order to gain insight about the wireless transmission between Alice and Bob and therefore degrade the level of security. In [68], the degree of security is discussed according to the statistical knowledge acquired by Eve, for example how Eve may use her knowledge about the multipath structure in order to reconstruct the channel coefficients and subsequently attempt to guess a portion of the secret key. In [42], Eve uses ray tracing methods in order to predict the channel response, which depends on side information such as the environment map and EM characteristics and on her localization accuracy of the legitimate terminals.

Nevertheless, when considering active attacks, Eve may impersonate a legitimate user since the channel over which Alice and Bob communicate is not authenticated. This problem, also known as "man in the middle", is extremely serious since difficult to counter. In this respect, the authors in [60] proposed an authentication scheme based on a level crossing algorithm. This kind of attacks is taken into account in the PHYLAWS project where a technique based on furtive and adaptive radio signals aims to protect the earliest stage of PhySec, in particular SKG [69]. Discussions about different attacks and countermeasure may be found in [70]. Accordingly, we do not consider active attacks in the present work.

2.2.3 Review of SKG

Prior works address issues encountered in the SKG process from either information-theoretic perspective or from a practical view. In the former, bounds on secret key rate are computed where simple analytical expressions are just found for perfect fading Gaussian channels, for either NB system [25, 58, 61, 62, 71, 72] or for WB/UWB technology [63, 65, 73]. However, real measured channels may deviate from Gaussian statistics, for example when a path dominates the propagation channel or when limited interfering paths exist. Although secret key capacity is computed numerically for both Nakagami and Suzuki fading channels in [74], this does not seem efficient since a large amount of statistical data is required and more complications appear for complex valued channel coefficients. Therefore, assessing security from the information-theoretic side is not always suitable.

On the other hand, practical assessments consist of quantizing the channel information (e.g. the received signal strength (RSS), the phase information, the CIR, etc.) into a sequence of secret bits (Section 3.2.1), where security may be assessed through

bit error rates and through achieved secret key rates. In this respect, most prior works proposed quantization protocols, often combined with a reconciliation scheme, in order to enhance the security [29, 59, 60, 75, 76, 77, 78]. In particular, they intend to deliver key rates approaching the secret key capacity, after applying the four key generation steps shown in Fig. 2-6. The authors in [71] target the key extraction from the complex channel coefficients of a complex Gaussian channel model. They show that the secret key rate may be within 1 bit from the secret key capacity, by adapting the number of extracted bits according to the SNR, using gray codes and LDPC codes with soft decision for error corrections.

Therefore, it is clear that most prior works focus on the role played by the quantization algorithm in retaining security, while few papers address the relation between SKG and the propagation channel features. In [58], the authors emphasize the role of multipath richness in 1) increasing the amount of randomness of MIMO technologies and 2) keeping the key safe from Eve. SKG was also discussed with respect to the Doppler frequency [60], to the LOS/NLOS character [58, 79] and to the coherence time [63]. Premnath et al. [29] applied practical SKG algorithms to the RSS of real channels, measured in different environment settings, including indoor and outdoor conditions and stationary and mobile terminals/intermediate objects. The measurements using 802.11-based laptops exhibited the weakness of SKG behavior in nearly static environments, where the entropy of extracted key bits is very low, whereas it turns out to be higher in mobile environments.

Hence, channel variability is a crucial requirement to establish long random secret key bits. The quality of the key in part depends on the statistical independence between key bits, which to some extent can be reduced to the lack of correlation between channel samples. Such an independence stems from sufficiently separated samples, in whatever domain sampling might be, which involves the physical propagation mechanisms and characteristics of the radio environment. On the other hand, several authors indeed emphasize the benefit of decorrelating channel observations, especially highlighting MIMO systems [59, 62].

Security in static environments was targeted by creating channel fluctuations through beamforming techniques provided by steerable array antennas [55, 80, 81]. Moreover, fast SKG in such environments was achieved by using opportunistic beamforming (without the need of a specific antenna design) and the frequency diversity [82]. Furthermore, terminals/scatterers mobility can be compensated through the investigation of the channel degrees of freedom (DoF), covering both the spatial domain and the frequency/delay domain, which express the channel richness and its

expected impact on randomness [24, 58, 83, 84, 85]: the richer the channel, the better the chance to exploit this richness usefully towards security. The spatial diversity existing in MIMO systems is investigated theoretically and experimentally in indoor channels [58]. In [65], the frequency diversity in orthogonal frequency division multiplexing (OFDM) systems is investigated theoretically in SKG techniques by assuming independent and identically distributed (i.i.d.) channel transfer functions. However, no validation of this approach was performed in well specified measurement campaigns and no quantitative information about the achieved key length was provided.

From the eavesdropper point of view, several works assume that Eve is not able to access correlated channel information when she is located more than a half or at most a few wavelengths away from both legitimate terminals [55, 61]. This assumption is valid in the case of rich omnidirectional multipaths, where the channel decorrelates rapidly over the specified domain. Nevertheless, SKG performance has been investigated statistically and empirically for a simple scenario where Eve is very close to Bob[1] [58, 86]. Moreover, a simple correlation model, which relies on a low-pass filter, is investigated equivalently in [84, 87]. However, the authors in [64] proved by measurements that spatial correlation can be found even for larger separation distances. In particular, shadow fading seems to be critical in PhySec, while shadow fading correlations [37] between Bob and Eve are likely to affect the information accessible to the eavesdropper and thus to impact the confidentiality. In this context, we have presented in [72] a very simple version of a GSCM, relying on the presence of scatterers between Alice and Bob/Eve in order to capture the common characteristics between Alice/Bob and Eve's channels without a restriction to stationary region. An extension of this model will be presented in Chapter 4.

2.3 Conclusion

This chapter presented an overview on the state of the art of both the wireless propagation channel and the secret key generation. In the first section, a definition of the wireless propagation was given, with special emphasis on features relevant to SKG, such as the multipath propagation, the channel reciprocity and the three types of channel fading. Moreover, an overview on existing channel models was given. These models can not be directly exploited in the context of PhySec, in particular the SKG scheme which seems to be very exigent. Indeed, on one hand, the SKG performance analysis requires a multi-user channel model that accounts for the intrinsic spatial

[1]Both Bob and Eve are assumed to share the same multipath components

channel variability between two users in proximity of each others, i.e. Bob and Eve. Such a requirement is targeted in Chapter 4, by elaborating on a GSCM. On the other hand, the quality of the key extracted from the radio channel heavily relies on the multi-path channel richness, which should be modeled with a sufficient accuracy (e.g. taking into account the diffuse components of the radio channel in RT simulations, ref. to Chapter 5).

Furthermore, in the second section, the strategy of deriving keys from the propagation channel was described, by discussing some scenarios for the eavesdropper. Related works in the area of SKG were also cited, with particular emphasis on SKG limitations.

In order to evaluate the quality of an *on-the-fly* generated key with respect to different propagation settings, we describe, in the following chapter, appropriate metrics and performance methods. Specifically, we emphasize the relation between key quality and channel features when investigating delay-dispersive channels.

Chapter 3

Secret key generation performance evaluation methods

This chapter is dedicated to discuss different methods for evaluating the performance in the establishment of secret keys and in relation to the randomness provided by fading propagation channels. We first examine, in Section 3.1, both qualitative and quantitative metrics that reveal the robustness of the generated key, mainly regarding its length and its random character. These metrics may also be classified as theoretical or practical, where the latter is evaluated after quantizing the channel information into a stream of key bits. In contrast with the theoretical metrics where the SKG performance is merely impacted by the propagation channel features, the practical SKG behavior is also influenced by the employed quantization algorithm. Hence, most earlier works intended to develop quantization protocols attempting to offer performance approaching the theoretical one. Nonetheless, since our ultimate goal is to relate the generated key robustness to the propagation channel characteristics, we chose to implement an efficient quantization algorithm selected from the literature, sensitive to fine channel characteristics and whose parameters can be adjusted. We argue and describe this algorithm in Section 3.2.

Subsequently in Section 3.3, we target how to investigate the channel degrees of freedom (DoF) existing in either the space or the frequency domain, with a goal to harvest as much randomness as possible from a single channel observation, given that channel noise or channel estimation errors limit the amount of random information shared between Alice and Bob. One advantage of exploiting the DoF as much as possible is to mitigate the requirement of terminal movements, in order to provide independent channel samples. This is a critical constraint in static environments.

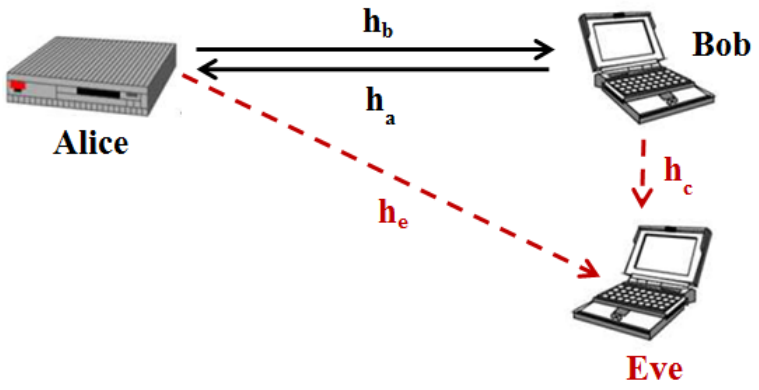


Figure 3-1: The terminals' propagation channels involved in the SKG scheme.

In other words, exploiting the channel DoFs may lead to address the SKG limitations explained deeply in Section 1.2 and discussed with respect to the literature in Section 2.2.3.

3.1 Metrics for SKG assessment

A cryptographic robust key is characterized by its length and its randomness. From the computational constraint point of view, the importance of the key length resides in avoiding a brute force attack, where the eavesdropper systematically checks all possible keys until the correct one is found. Nevertheless, novel schemes (such as investigated in the PHYLAWS project [15]) can bring an useful extra degree of security without requiring extremely long keys. This can be done by reinforcing the security protocol, e.g. through a limited validity time for valid keys [88] or through more sophisticated approaches [69]. On the other hand, from the perspective of information-theory, a high key randomness is essential to avoid significant information leakage to an eavesdropper, making life easier for her in reducing the key search space [68].

Therefore, we discuss in this section how to quantitatively assess the quality of the generated key, in terms of length and randomness, from both theoretical and practical perspectives. Moreover, we present the channel correlation coefficient as a qualitative metric, since, on one hand, the channel reciprocity is relevant to secret sharing between Alice and Bob, on the other hand, the spatial decorrelation impacts the amount of information exposed to Eve.

3.1.1 Information theoretic bounds on key length

The reciprocal fading channels measured by Alice ($\hat{\mathbf{h}}_a$) and Bob ($\hat{\mathbf{h}}_b$) are considered to contain a common randomness that should be quantized in order to obtain shared key bits (see Fig. 3-1 for an illustrative scheme of the channels between members of the trio). Thus, in an information-theoretic framework, the maximum amount of random information reliably shared between Alice and Bob is measured by the mutual information between their legitimate channels, or:

$$I_K = I(\hat{\mathbf{h}}_a, \hat{\mathbf{h}}_b) \tag{3.1}$$

We recall that the terminals measure noisy channel gains, as modeled in Eq. 2.10. The available key bits are, on one hand, entirely secure only in case Eve experiences channels ($\hat{\mathbf{h}}_e$ and $\hat{\mathbf{h}}_c$) that are statistically independent from those measured by the legitimate terminals, i.e. it is commonly considered that Eve should be sufficiently far from both Alice and Bob. Furthermore, it is commonly assumed that Eve has negligible information about the propagation scenario (e.g. terminals positions and surrounding environment) [68], so that she is unable to perform channel predictions through e.g. a ray tracing tool [42]. In general, not all of these bits are secure if the eavesdropper may have access to some insight about the legitimate channels. Therefore, the secret key rate I_{SK} can be obtained by evaluating the mutual information between the channels seen by Alice and Bob, given Eve's observations [7, 17, 89], or:

$$I_{SK} = I(\hat{\mathbf{h}}_a, \hat{\mathbf{h}}_b | \hat{\mathbf{h}}_e, \hat{\mathbf{h}}_c) \tag{3.2}$$

In the present work, we consider a simplified scenario where Eve is closer to Bob than to Alice, as seen in Fig. 3-1. As a result, we can assume that the channel Bob-Eve ($\hat{\mathbf{h}}_c$) is independent from the main channel Alice-Bob. Accordingly, the secret key rate is simplified to:

$$I_{SK} = I(\hat{\mathbf{h}}_a, \hat{\mathbf{h}}_b | \hat{\mathbf{h}}_e). \tag{3.3}$$

We also define the vulnerable key rate as follows:

$$I_{VK} = I_K - I_{SK} \tag{3.4}$$

Given that both available and secret/vulnerable key capacities are statistical quantities computed based on mutual information, and in order to numerically evaluate these theoretical bounds, we give in Eqs. 3.5 and 3.6 the basic definitions of the

mutual information between two random variables (X and Y) and also conditionally to a third one (Z) [90].

$$I(X,Y) = \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{3.5}$$

$$I(X,Y|Z) = \sum_x \sum_y \sum_z p(x,y,z) \log \frac{p(z)p(x,y,z)}{p(x,z)p(y,z)} \tag{3.6}$$

If the channels are jointly complex Gaussian random vectors with zero mean, the mutual information in Eqs. 3.5 and 3.6 can be easily computed starting from covariance matrices of channel observations [90]. In other words, the channels are required to follow either Rayleigh or Rician fading, where the latter case requires the removal of the dominant deterministic component [58, 71]. Nevertheless, I_K and I_{SK} are expressed by the following formulas:

$$I_K = I(\hat{\mathbf{h}}_a, \hat{\mathbf{h}}_b) = \log_2 \frac{|\hat{\mathbf{R}}_{aa}||\hat{\mathbf{R}}_{bb}|}{|\hat{\mathbf{R}}_{AB}|} \tag{3.7}$$

$$I_{SK} = I(\hat{\mathbf{h}}_a, \hat{\mathbf{h}}_b|\hat{\mathbf{h}}_e) = \log_2 \frac{|\hat{\mathbf{R}}_{AE}||\hat{\mathbf{R}}_{BE}|}{|\hat{\mathbf{R}}_E||\hat{\mathbf{R}}_{ABE}|} \tag{3.8}$$

where covariances with lowercase subscripts denote $\hat{\mathbf{R}}_{xy} = E\{\hat{\mathbf{h}}_x \hat{\mathbf{h}}_y^H\}$, while those with uppercase subscripts are covariances of stacked channel vectors, or $\hat{\mathbf{R}}_{XY\cdots Z} = E\{[\hat{\mathbf{h}}_x^H \hat{\mathbf{h}}_y^H \cdots \hat{\mathbf{h}}_z^H]^H [\hat{\mathbf{h}}_x^H \hat{\mathbf{h}}_y^H \cdots \hat{\mathbf{h}}_z^H]\}$. Explicit evaluation of these covariance matrices may be found in Appendix A.$|.|$ denotes for the determinant.

Unfortunately, the Gaussianity of the channel coefficients cannot be ensured in all scenarios. This will cause a significant difficulty in evaluating I_K and I_{SK}, since there is no closed-form expression for the mutual information in the general case. Hence, we just compute this mutual information for channels undergoing Rayleigh fading. Nonetheless, we can assess the security performance after translating the channel information into discrete key bits, as well explained in the next section, although there is less guarantee than provided by the direct application of information theory principles.

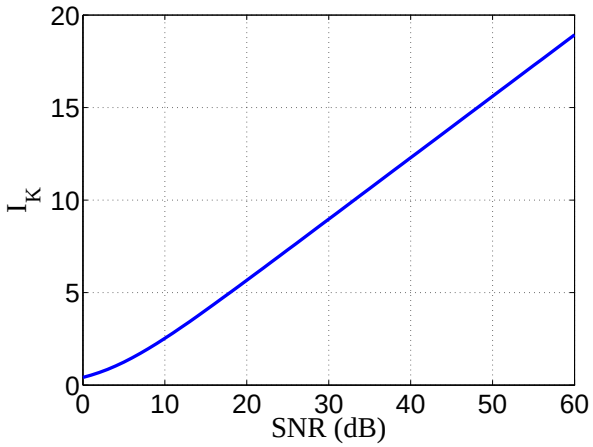


Figure 3-2: The impact of the SNR on I_K.

Impact of SNR

One main factor that limits the number of available key bits is the signal to noise ratio (SNR). We assume that the transmitted power increases proportionally to the number of transmission sub-channels. Thus, the average SNR, assumed for a single antenna single frequency link, is defined as:

$$\gamma = \frac{E\{|\mathbf{h}|_F^2\}}{N_h \sigma^2} \tag{3.9}$$

where $|.|_F$ and $E\{.\}$ denote respectively the Frobenius norm and the expectation over the channel realizations. σ^2 is the noise power while N_h is the length of the channel vector. Indeed, the components of the channel vectors refer to parallel multiple channels, which occur in either an OFDM system, a SIMO/MIMO system or the joint technologies.

We can simply examine the SNR impact by considering a narrow band complex Gaussian channel between Alice and Bob, with i.i.d. channel realizations. It is hence relatively easy to show that I_K is determined by the signal to noise ratio γ (or $SNR = 10\log_{10}(\gamma)$), assumed identical at Alice and Bob sides. I_K is thus expressed by the following formula [71]:

$$I_K = -\log_2\left[1 - \left(\frac{\gamma}{1+\gamma}\right)^2\right] \tag{3.10}$$

Obviously, the available key bits just reveal the accuracy of the analog to digital conversion, which is limited by the SNR. Consistently, we show in Fig. 3-2 that I_K

increases proportionally to the SNR in dB, at least beyond ∼10 dB. In most of the results presented in the following chapters, we consider the SNR to be equal to 15 dB, which results in $I_K(1) = 4.05$ bits per channel observation for $N_h = 1$. Indeed, while a higher SNR can be considered interesting, since providing more key bits and more randomness, the noise in Eq. 2.10 is not only an additive thermal noise, but also comprises channel estimation errors, imperfect calibration of the measuring chain etc. Since SKG might be used predominantly in low cost commercial devices, it can be anticipated that the effective lack of reciprocity between Alice and Bob channels (represented through the SNR) will be significant and assuming very high SNR would not be realistic.

Impact of the Rician K factor

Besides the impact of the noise, the robustness of the SKG heavily depends on the amount of randomness available in the reciprocal radio channel. Indeed, unconditionally security schemes rely on the fact that the eavesdropper should have insufficient information about the source from which keys are extracted. Accordingly, merely the random portion of the radio channel is beneficial for secret key generation [58, 79]. While Rayleigh fading channels are fully random, Rician fading channels contain a predictable part (i.e. typically the LOS path) that limits the amount of randomness. Thus, in order to appropriately apply the SKG, it is suggested to remove any predictable portion of the channel [58, 71, 91]. However, this cannot be achieved by just removing the channel mean, as suggested in [71], since the dominant path may have a changing phase owing to terminals motion. For that reason, the authors in [58] perform principal component removal by investigating the eigenvectors of the MIMO covariance matrix, the drawbacks of which stem from the need of sufficient statistics as well as sufficient DoFs. Moreover, since the deterministic part may not restrict to a single dominant path, other protocols attempt to predict and remove such a part through the use of, for example, a Markov chain [92] or a linear prediction approach [91]. Nonetheless, the task of removing the predictable portion of the channel is beyond of the scope of the present work, while the aim here is to simply quantify the impact of the limited amount of randomness on the SKG. For that reason, we often take in consideration the Rician K factor in the results presented below.

The K factor is defined as the power ratio of the dominant deterministic path to the remained scattered paths. We consider the case of a narrow band Rician channel resulting from the presence of a fixed dominant path h_{los} (with a unit mean power),

where the complex channel gain can be modeled as follows:

$$h = \sqrt{\frac{K}{1+K}} h_{los} + \sqrt{\frac{1}{1+K}} h_{scatt} \tag{3.11}$$

h_{scatt} results from the contribution of a sufficient number of i.i.d. scattered multipath components, assuming a rich scattered environment. Hence, according to the central limit theorem, h_{scatt} follows a zero-mean complex Gaussian distribution with a unit variance, i.e. $h_{scatt} \sim CN(0,1)$. We adequately evaluate the amount of shared randomness between Alice and Bob once the dominant path is removed. We assume that Alice and Bob see noisy versions of the reciprocal channel h with a fixed SNR of 15 dB (γ in linear scale). Since SKG exploits merely the scattered part of the channel, I_K depends on the SNR of such a part, i.e. $\gamma_{scatt} = \frac{\gamma}{1+K}$. Consequently, $I_K = -\log_2\{1 - (\frac{\gamma}{\gamma+(1+K)})^2\}$. We plot in Fig. 3-3 the variation of I_K with the Rician K factor for SNR=15 dB. It is straightforward to notice that I_K decreases as K increases. This results from the decrease in the received power as K increases after removing the dominant component, yielding a smaller SNR for the scatterered paths (γ_{scatt}). This shows that random keys may be extracted from Rician channels if the diffuse components has sufficient SNR, however the price to pay is the capability for the system to properly extract the dominant component.

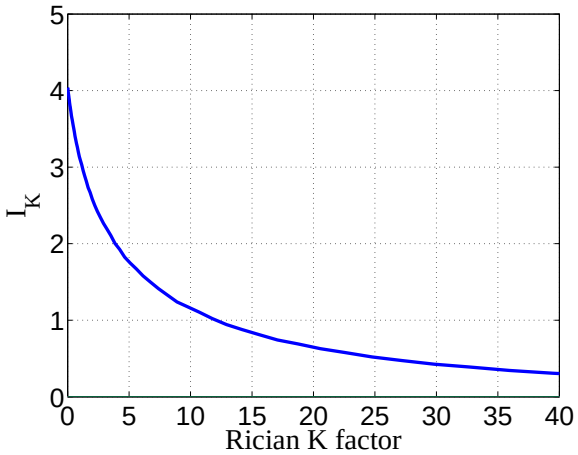


Figure 3-3: I_K as a function of the Rician K factor.

Vulnerability in an ideal scenario

It is usually assumed that Eve has negligible information about the main wireless channel provided that she is sufficiently far away from both Alice and Bob. This

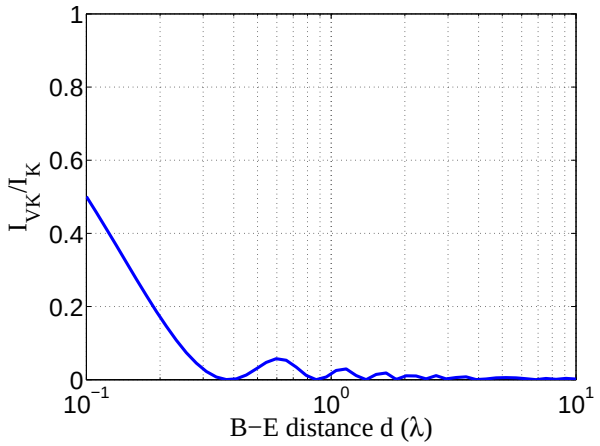


Figure 3-4: I_{VK}/I_K as a function of the separation distance d for Clarke's model.

means that the channel experienced by Eve is totally decorrelated from that measured by legitimate terminals. In fact, this represents an ideal scenario (i.e. the Clarke model) where the channel is assumed rich in paths uniformly dispersive in the angular domain. This brings rapid channel variations in either the time or the spatial domain, expressed by the Jakes correlation function [93]:

$$E\{\mathrm{h}_b, \mathrm{h}_e\} = J_0(2\pi d/\lambda) \tag{3.12}$$

where $J_0()$ is the zero-order Bessel function, λ is the wavelength, and d is the distance between Bob (B) and Eve (E). Accordingly, the normalized vulnerable key rate, which is plotted versus the distance d in Fig. 3-4, reveals that no information is disclosed to Eve after a few wavelengths. However, the reality is more complex and may significantly deviate from Clarke's scenario. Indeed, in practice, the channel is often dominated by a set of discrete multipaths, for which the angular spread may cover a moderate angle, which will require a larger separation distance to warrant sufficient decorrelation and an adequate degree of security. In the following chapters, we will consider the vulnerability issue in more realistic scenarios.

3.1.2 Channel correlations

Given that the information security metrics for Gaussian channels recalled above are based on second order quantities, the correlation between channel coefficients is responsible for the imperfect security performance. Clearly speaking, the larger the correlation between the legitimate channels, the more shared bits between Alice and

Bob can be expected to be successfully extracted. Moreover, the confidentiality is assured owing to the spatially decorrelated channel measured by the eavesdropper. Therefore, the correlation can be used to assess qualitatively SKG performance in terms of reliability and confidentiality.

We use the conventional channel correlation coefficient between two random complex variables X and Y of mean $\mu_X = E\{X\}$ and $\mu_Y = E\{Y\}$, defined as follows:

$$corr(X,Y) = \frac{E\{(X-\mu_X)(Y-\mu_Y)^*\}}{\sqrt{(E\{|X-\mu_X|^2\})(E\{|Y-\mu_Y|^2\})}} \tag{3.13}$$

where $(.)^*$ stands for complex conjugate. We consider two types of channel correlations: 1) the complex channel correlation coefficient (referred to as ρ_{xy}) where X and Y represent respectively the complex channel coefficients h_x and h_y and 2) the power envelope correlation coefficient, where X and Y are replaced by the channel powers $|h_x|^2$ and $|h_y|^2$, respectively.

3.1.3 Channel quantization performance

As discussed in Section 2.2.1, Alice and Bob jointly quantize their estimated channel information into a stream of key bits. If the EM reciprocity holds perfectly even without channel estimation noise (i.e. $\hat{\mathbf{h}}_a = \hat{\mathbf{h}}_b$), legitimate users are then able to adequately share the same string of key bits. Otherwise, they should reconcile their keys with the least amount of information leaked to Eve. For that reason, the bit error rate (BER) should be small. Hence, it is crucial to study the BER between the key bits extracted by both Alice and Bob, which is the ratio of the number of erroneous bits to the total number of bits[1]. Accordingly, we define the efficiency η (in bits/channel observation) as the average identical number of bits that can be extracted from a single channel observation even over N_h channel degrees of freedom (i.e. the case of stacked channel vectors), or:

$$\eta = N_h(1 - BER)\log_2(M) \tag{3.14}$$

This formula assumes i.i.d. channel coefficients even over the stacked channels, which refers to the channel DoF. As a consequence, this expression is not able to quantify the length of "random" bits. Furthermore, if Alice and Bob do not agree on a given symbol, it is relevant to discard the erroneous bit rather than the whole symbol since

[1]We choose to evaluate the BER instead of the symbol error rate (SER) in order to account for the reconciliation phase which operates at the bit level.

they are able to reconcile their bits in the steps that follow the bit extraction phase. For that reason, we use the BER in the expression of η instead of the symbol error rate as admitted in [58].

3.1.4 NIST tests

The fact that key bits are not statistically independent reduces the key quality since in an information-theoretic framework Eve may exploit any useful information to collapse the key space. In this context, the source of randomness, in addition to the admitted quantization algorithm, is the most critical aspect that mainly affects the key robustness behavior. Hence, we aim to assess the security performance in terms of randomness, which can be achieved using the NIST test suite [94]. We notice that these tests are not able to prove the perfect randomness of a key. However, each test shows if the key bits follow a certain expected behavior owing to key generation process [75]. There are 16 statistical tests in total, but we are not able to apply all these tests owing to limitations on the requirement of each tested key length. Table 3.1 shows length limitations for some tests applied in this dissertation, where m is the length in bits of the bit strings used in each test and N is the key length in bits. The remaining tests require very long keys and do not apply to the physical layer based wireless security scheme we here target.

Table 3.1: NIST tests limitations.

Mono-bit frequency	$N \geq 100$
Block frequency	$N \geq 100$
Runs	$N \geq 100$
Serial	$m < \lfloor \log_2 N \rfloor - 2$
Approximate entropy	$m < \lfloor \log_2 N \rfloor - 5$

Some tests try to show whether the sequence of bits has the statistical properties of a random sequence. Consistently, the "mono-bit frequency" and the "block frequency" tests investigate this randomness criteria on respectively the entire key bits and in sub-blocks. For example, if we consider the following sequence of N bits:

$$\underbrace{00...00}_{N/4}\underbrace{10...10}_{N/4}\underbrace{11...11}_{N/4}\underbrace{01...01}_{N/4},$$

we notice that it passes the mono-bit frequency test since 0 and 1 are equiprobable bits in the whole sequence whereas it is not the case in subblocks where too many

bits equal either to 0 or to 1 may be present, leading to failure of the block-frequency test. Another test, which is the "runs" test, checks whether the frequency of runs (uninterrupted strings of identical bits either 0 or 1) is that expected for a random sequence. In other words, it determines whether the transition between bits 0 and 1 is too fast or too slow. Accordingly, the sequence in the above example is considered random since the number of runs is very close to that expected for a random sequence (i.e. $N/2$ runs). However, the following sequence:

$$\underbrace{00...00}_{N/3}\underbrace{11...11}_{N/3}\underbrace{00...00}_{N/3}$$

is not random since only 3 runs are computed.

Both the "approximate entropy" (ApEnt) test and the "serial" test focus on the frequency of occurrences of all possible overlapping 2^m strings of m-bits length each, across the entire key bits. Their purpose is to compare the frequency of overlapping strings of several consecutive lengths against the expected result for a random sequence. To that aim, the ApEnt test uses two consecutive bit lengths (m and $m+1$) while the serial test uses three consecutive lengths (m, $m-1$ and $m-2$). Moreover the serial test differs from the ApEnt test by the fact that longer bit strings can be used in the former for the same key length, as shown in Table 3.1. According to both the ApEnt test with $m = 1$ and the serial test with $m = 2$, the first sequence example is supposed random since strings of 2 bits are almost equiprobable whereas the second key example fails these two tests. Furthermore if we consider strings of higher lengths, the first sequence may fail the tests. More information about these statistical tests can be found in [94].

For a single key, each randomness test indicates whether the key is accordingly random or not. Furthermore, in order to relate the quality of the randomness to the features of the radio channel, a set of generated keys is tested by each randomness test, which returns a percentage of sequences passing the test. Then we computed the "mean pass rate" by averaging the percentages of sequences passing each NIST test and thus over all the applied statistical tests, which provides a global assessment of the randomness for each specific scenario.

3.2 Channel quantization alternating (CQA) algorithm for SKG

Apart from the crucial impact of random wireless channel, security through SKG is, to some extent, related to the effectiveness of the quantization mechanism that is required to extract more random key bits with vanishing error rate. Thus, we firstly review various algorithms attempting to distill keys from various channel information, such as the RSS. Then, we describe the channel quantization alternating scheme, proposed in [58], which is used throughout this work to practically establish symmetric keys.

3.2.1 State of the art

The key rate may be increased by extracting several bits from a single channel sample. However, this may increase the number of mismatched bits extracted by both Alice and Bob, since samples become closer to quantization boundaries. In order to address this issue in either mono- or multi-bit extraction schemes, quantization algorithms seek to employ a suitable censoring scheme. While some algorithms increase Alice-Bob's key agreement by dropping down samples falling into a predefined guardband region [57, 58], the time required to construct a long secret key increases, which reduces the effectiveness of such algorithms. Alternatively, a more efficient protocol adapts the quantization map to each channel observation, by using at least two alternative maps [58, 59]. The main goal of these censoring schemes is to reduce the BER in order to render the reconciliation phase more efficient, with a minimum amount of information disclosed to Eve.

Furthermore, some algorithms intend to provide security by better ensuring the random character of the generated bits. In this respect, decorrelation transformations are proposed to be employed before the quantization step. In [59], a linear transform (discrete Karhunen-Love transform), which exploits the eigenvectors of the covariance matrix, is used to convert an estimated channel vector into decorrelated components. A similar approach is used in [62], where the eigenvectors are transmitted over the public channel between Alice and Bob. Other protocols attempt to exploit just the unpredictable portion of the channel information, by predicting and removing the deterministic components [58, 91, 92], as discussed above in Section 3.1.1. Despite the pros and cons of these approaches, since we aim to assess the amount of intrinsic random information in relation to the channel features, we are interested in a

simple approach that ensures preliminary randomness by defining statistically equal quantization intervals [58] without any further data processing.

Moreover, the channel information quantized into key bits impacts the security. Although the most common channel quantity used in practical SKG is the RSSI [29, 55, 56, 57] because this parameter is widely accessible in most radio receivers, the entropy of the generated keys is not very high. Alternatively, the channel phase information has been investigated and found to generate more random and secure stream of bits, such as in [65, 76, 95, 96]. However, either RSSI or the phase only partially exploit the richness of the channel information [65]. Another candidate for SKG is the channel impulse response (CIR) of either an OFDM, WB or even UWB channels, whose ability to support SKG techniques has been targeted through the literature [61, 64, 73, 75, 97]. Nevertheless, we can efficiently establish a sufficient long and random key by exploiting as much as possible channel information at once, which is achieved by making use of the joint real and imaginary parts of channel coefficients (complex CSI) [58, 62].

3.2.2 CQA description

Therefore, we specifically consider in the present work the channel quantization alternating (CQA) protocol proposed in [58], which offers the following advantages: 1) the exploitation of the full information contained in the complex CSI, 2) the definition of statistically equally probable quantization regions (QR), and 3) the decrease of the BER without samples rejection. A preliminary set of channel coefficients is required in an initial learning phase, in order to statistically define the quantization regions of a quantization map. Subsequently, it will be necessary to specify the nature of the channel observations that serve in the construction of the I/Q maps (learning phase), from which subsequent observations can be quantized according to CQA (SKG phase). Obviously, the more numerous observations can be used in the learning phase, the more precise will the maps be and the more similar Alice's and Bob's maps will be. Along this line, the more channel observations we have for the SKG phase, the longer and the more secure the keys will be, providing margin since key length is reduced after the privacy amplification step.

To generate secure key bits from the shared source of randomness, Alice and Bob separately employ CQA. Alice defines her map QRs by quantifying the cumulative distribution function (CDF) of each of the aggregated real and imaginary parts of the channel coefficients into $\sqrt{M}$ statistically equal quantization intervals, resulting

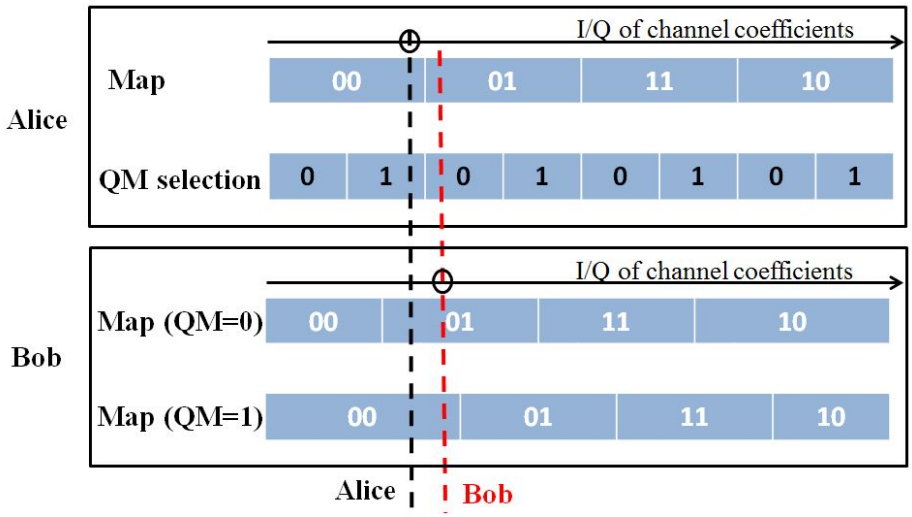
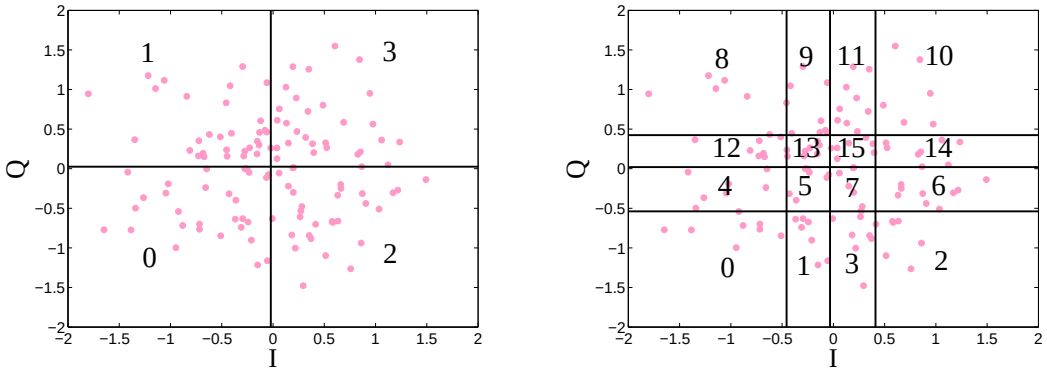


Figure 3-5: Illustration of the quantization intervals of the CQA scheme on one dimension I or Q with $M = 16$.

in M quantization regions. In seeking key mismatches reduction, Alice divides each quantization interval into two sub-intervals of equal probability. She then sets a quantization mapping (QM) index to 0 or 1 if the current sample falls within the first or the second sub-interval, respectively, and then publicly sends the QM to Bob. Fig. 3-5 illustrates the quantization intervals along a single dimension (I or Q) for both Alice and Bob and for $M = 16$ QRs (i.e. 4 quantization intervals in each single dimension). On the other hand, Bob first performs the same earlier step done by Alice and computes the preliminary boundaries from his own set of channel coefficients observations. Bob then defines two alternative maps by shifting the quantization boundaries, each one in a different direction, with the same probability (i.e. $1/4\sqrt{M}$), which corresponds to the half of each QM probability (please refer to [58] for further insights). Fig. 3-5 shows a simple example demonstrating the efficiency of CQA where Alice and Bob are able to obtain the same string of bits. Alice obtains the string "00" by using her own unique map. If Bob uses the same map, he obtains a different string, i.e. "01". Otherwise, if Bob uses the map labeled with QM=1 as indicated by Alice, he is able to reproduce the same string as Alice, i.e. "00". Hence, Alice publicly communicates the adequate QM index to Bob.

We note that we use the Gray coding in mapping each QR into a string of key bits, since symbols affected to two alternative QRs differ from only 1 bit, which in turn aids in diminishing the BER [71]. Fig. 3-6 depicts a particular channel realization in the complex I-Q plane and shows Alice's map by presenting the correspondence between symbols and QR, both for $M = 4$ and $M = 16$. We recall that the interest of



(a) M=4 (b) M=16

Figure 3-6: Correspondence between symbols and QR.

increasing M is to establish a certain key length with fewer channel samples. However, we will see in the subsequent chapters the conditions suitable to increase M.

3.3 Exploiting channel variability in SKG

As shown in Section 3.1.1, the available key rate I_K is limited by both the SNR and the deterministic components of the radio environment. This is pragmatically addressed through the accumulation of several sub-keys, which are required however to be independent. Indeed, the quality of the key in part depends on the statistical independence between key bits, which to some extent can be reduced to the lack of correlation between channel samples. Such an independence stems from sufficiently separated samples, in whatever domain sampling might be, which involves the physical propagation mechanisms and characteristics of the radio environment. Hence, channel variability is essential to achieve SKG.

Given the limitations of SKG when exploiting the time variability, achieved through movements of either the terminals or scatterers in the surrounding environment, we propose to investigate either the space, the frequency or joint space-frequency degrees of freedom (DoF). The benefit stems from the increase in the number of shared random bits per channel observation and the reduction in the time required to achieve a long secret key, for time variant channels. In the case of N_h independent degrees of freedom, which would result from independent parallel sub-channels with reciprocity between Alice and Bob limited by an SNR given by γ, the number of available key

bits is given as follows:

$$I_K = -N_h \log_2\{1 - (\frac{\gamma}{1+\gamma})^2\}. \tag{3.15}$$

However, this equality does not hold in the case of non independent parallel sub-channels. Hence, the security strongly relies on the sub-channels correlation degree. Intuitively, the less correlation between them, the more key bits per single channel observation.

SKG performance goes through maximizing key length, under the condition that key bits are as much independent as possible, which means capturing as many as possible parallel sub-channels and ensure as much as possible independence between these channels. Such a consideration will be at the heart of the work performed and described throughout this dissertation. In the present section, we mainly focus on the exploitation of parallel channels coming from the dispersion in the delay domain and try to get insights from a simple model of dispersive channels.

3.3.1 Space variability

Space variability stems from the use of several antennas at either the receiver or the transmitter side (e.g. massive MIMO [98]) although coupling and other effects may disturb this simple picture. Space variability also stems from varying positions for a single antenna, which is equivalent to time variability if the channel can be assumed static over the CIR duration.

The correlation between spatially variant channels relies in general on the angular characteristics of the radio propagation. If the multipaths are dispersive in the angular domain, the channels exhibit faster spatial decorrelation, especially in rich scattered environments as may be expressed by the Bessel function of Eq. 3.12. On the opposite, more correlated channels result from a highly directional multipaths scenario, for a given distance between points where the channel is measured. This may be expressed through the channel coherence distance. In dense and sufficiently omnidirectional scenarios, it is commonplace to consider that antennas separated by at least $\lambda/2$ provide sufficiently decorrelated signals (i.e. channel coefficients). However, this is less true in directional ones, such as rural scenarios or in street canyons with visibility from the BST [99]. We notice that the vulnerability of SKG in either static or LOS conditions may be reduced through the use of specific configurable antennas enabling to create artificial fluctuations by randomly changing the beam-patterns, thus giving

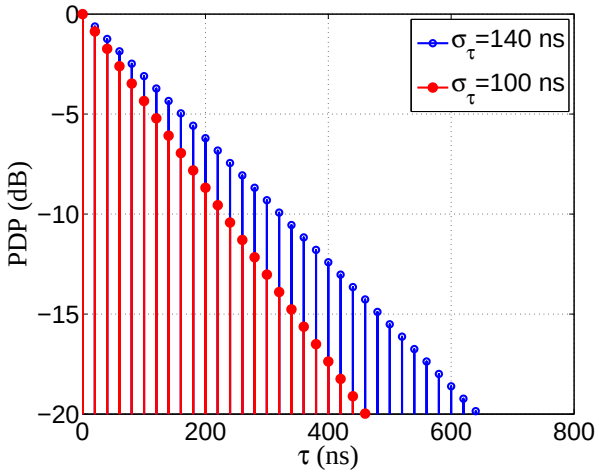


Figure 3-7: Some examples of PDP with different RMS delay spreads.

Alice/Bob an advantage over Eve [55, 81].

3.3.2 Frequency variability

In real world applications, a spatial degree of freedom may not always be available (e.g. in single antennas links with very stable channels). In such a case, it is recommended to find another source of channel variability, hence the need to exploit the frequency variability existing in WB/UWB channels.

We here describe a simple dispersive channel model, which will help identify the phenomena under SKG performance from correlated sub-channels and address the optimization of SKG schemes. It is based on periodic multipaths in the delay domain, with an exponentially decreasing mean power. Each path is Rayleigh distributed, i.e. independently Gaussian distributed with the same variance on the I and Q components. The two sole channel parameters are then the delay gap between each multipath and the delay spread (σ_τ), plus the SNR (Fig. 3-7). The instantaneous channel response writes:

$$h(t,\tau) = \sum_{n=1}^{N_{path}} \beta_n e^{j\phi_n} \delta(t - (n-1)\Delta\tau) \tag{3.16}$$

$\Delta\tau$ and N_{path} are respectively the delay resolution and the number of paths that are assumed statistically independent. ϕ_n is the random phase shift uniformly distributed within $[0, 2\pi]$ while β_n is the positive real-valued amplitude of the nth path, assumed

Rayleigh distributed with a mean power [100, 101, 102]:

$$E\{|\beta_n|^2\} = \exp\{\frac{-(n-1)\Delta\tau}{\sigma_\tau}\} \tag{3.17}$$

Notably, however simplified this model may seem, it is commonly used for representing channels in standards as advanced as IEEE 802.11ac [103, 104], complemented by several clusters rather than one, and with extra parameters specifying the spatial dependence of this channel (readily needed for multiple antenna systems).

Some examples of such PDP are displayed in Fig. 3-7 with both $\sigma_\tau = 140$ ns and $\sigma_\tau = 100$ ns. The maximum excess delay is set to 1980 ns, corresponding to $Npath = 100$ and $\Delta_\tau = 20$ ns. Starting from the CIRs, we perform a discrete Fourier transform in order to obtain 100 channel transfer functions within a BW of 50 MHz. Then, N_f (≥ 1) channel transfer functions are stacked within a vector used in the computation of the number of available key bits per single channel observation, i.e. $I_K(N_f)$. In other words, I_K is computed in the frequency domain. Nonetheless, it is also shown in [65, 63, 73, 71] that I_K may be alternatively computed in the time domain. In particular, the upper bound on the key rate may be expressed as follows[2]:

$$I_K \leq -\sum_{n=1}^{N_f} \log_2 \left[1 - \left(\frac{\gamma_t(n)}{1+\gamma_t(n)}\right)^2\right] \tag{3.18}$$

where $\gamma_t(n)$ is the SNR of the nth resolved path. The equality holds when the resolved paths are independent.

In order to assess the security behavior with the channel DoF brought by the frequency variability, we consider two different schemes with an increasing number of sub-carriers, as illustrated in Fig. 3-8. The first consists in studying the impact on the security of an increasing bandwidth with a fixed frequency separation ($\Delta f = 0.5$ MHz). More clearly, the BW increases proportionally to the number of investigated sub-carriers N_f (i.e. $BW = N_f \Delta f$) that are added around a central frequency. The second method investigates the increased number of sub-carriers within a fixed BW ($BW = 50$ MHz) which is achieved by decreasing Δf. Whatever the adopted method is, we assume a fixed and equal SNR for both Alice and Bob, i.e. $SNR = 15$ dB for which $I_K(1) = 4.05$ bits per single channel observation. Moreover, we note that the power spectral density (PSD) of both the sub-carrier and the noise is constant, which means that increasing N_f yields a proportional increase in the transmitted/received

[2]The total number of resolved paths is equal to the number of investigated frequencies N_f.

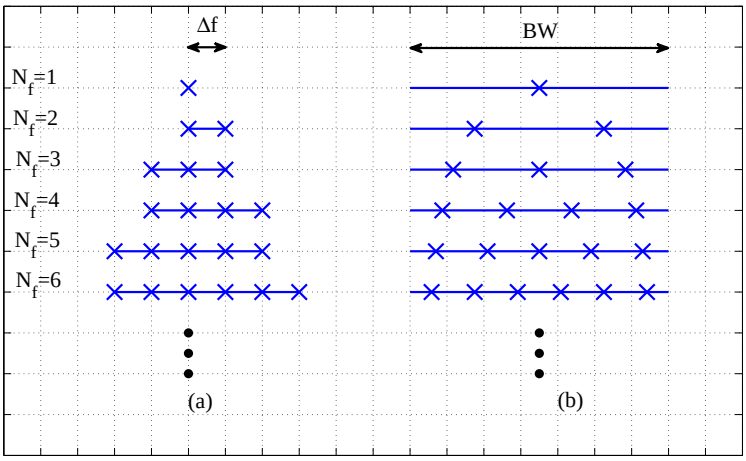


Figure 3-8: Frequency variability schemes with an increasing number of sub-carriers N_f: (a) an increasing BW and a fixed Δf, (b) a fixed BW and a decreasing Δf.

power.

An increasing BW with a fixed frequency separation Δf

Fig. 3-9 shows the evolution of the available key rate with respect to the number of sub-carriers, directly proportional to the BW, and thus for several time delay resolutions and for $\sigma_\tau = 100$ ns. It is shown that I_K linearly increases with N_f until a certain value beyond which the slope of the curve decreases showing a nearly saturation behavior. This means that until this critical value, the set of selected frequencies is not able to accurately resolve all the multipath components of the PDP [105]. However, when using N_f frequencies corresponding to $BW = 1/\Delta\tau$, the paths are perfectly resolved indicating a full exploitation of the channel DoFs. In fact, when N_f sub-carriers are used, the delays are quantized into N_f delay bins, each one with a gain obtained by using a sinus cardinal filter [65]. As the inverse of the BW is larger than $\Delta\tau$, several paths are overlapped to give correlated gains in several delay bins. However, when $1/(N_f\Delta f)$ is proportional to $\Delta\tau$, the paths are perfectly resolved as clearly illustrated in Fig. 3-10.

Once the CIRs are perfectly resolved, I_K still increases with N_f (or BW), even slowly. This is explained by the improvement on the SNR per resolved path owing to the increase in the total transmitted power as the number of investigated sub-carriers increases, according to Parseval theorem. Similar results are seen in [97]. However, we notice that if the energy is fixed in the frequency domain instead of the

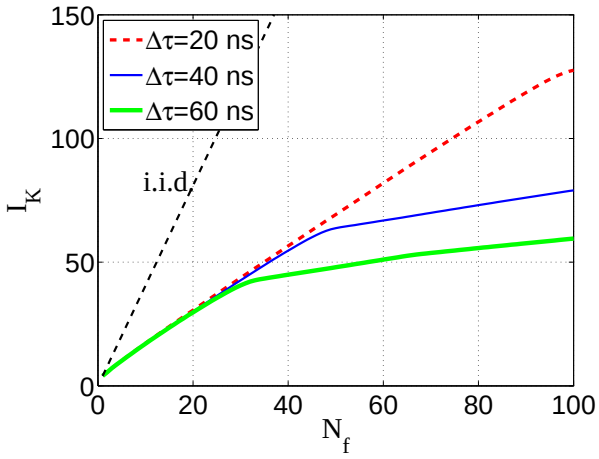


Figure 3-9: Evolution of I_K with respect to N_f and to the time resolution Δ_τ, for $\sigma_\tau = 100$ ns.

PSD, the authors in [63] show that there is an optimal bandwidth that maximizes the mutual information I_K since increasing the BW yields a decrease in the power of each sub-carrier.

Fig. 3-11 plots I_K as a function of N_f, for several RMS delay spreads and for $\Delta\tau = 20$ ns, while Fig. 3-12 presents the variation of I_K with σ_τ for both $N_f = 2$ and $N_f = 100$. Obviously, the available key rate increases as σ_τ increases. Indeed, the higher the RMS delay spread, the more randomness available in the radio channel through the multipath components. This is explicitly shown in Fig. 3-7 where the diffuse paths become with higher relative power for higher σ_τ. As seen in Fig. 3-12, the increase in σ_τ yields that I_K may approach the theoretical value corresponding to i.i.d. observations even for small bandwidths, e.g. when $N_f = 2$, where the paths are not perfectly resolved.

A fixed BW with a decreasing frequency separation Δf

Consider now the case where N_f frequencies are uniformly distributed within a fixed BW of 50 MHz (Fig. 3-8 (b)). The variation of I_K with respect to N_f is depicted in Fig. 3-13, where $\Delta\tau = 20$ ns and $\sigma_\tau = 100$ ns. For relatively small values of N_f, I_K increases dramatically with N_f in a manner such that it approaches the available key bits given by i.i.d. sub-carriers, in contrast to results shown in Fig. 3-11. Indeed, the fixed BW corresponds to the time resolution of the signal, leading to an observation of the fine structure of MPCs while the decreasing value of Δf determines the duration of the resolved PDP. As a consequence, as N_f increases, more independent paths are

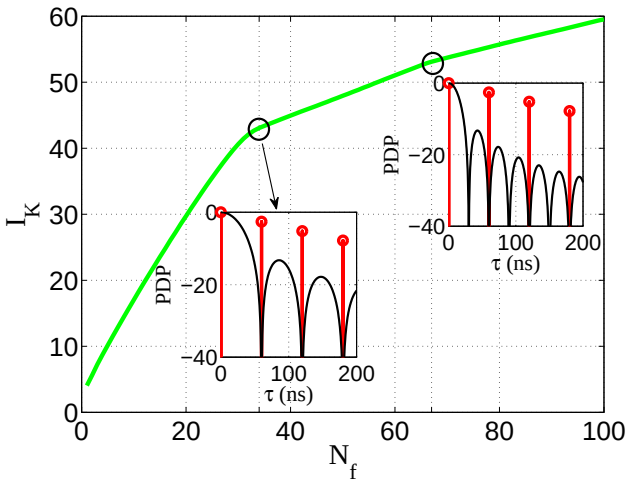


Figure 3-10: Equivalent time domain when perfectly resolved paths, for $\Delta_\tau = 60$ ns and $\sigma_\tau = 100$ ns.

resolved in the time domain, revealing more information to be exploited in the key generation process. In these conditions, it is clearly shown that I_K can be computed as in Eq. 3.18. When the number of sub-carriers becomes relatively dense, the SNR per additional resolved path decreases as N_f increases, revealing little information available to SKG, especially that the path power becomes near the noise power. This results in a diminishing slope in the curve of I_K. We note that a similar scheme is targeted in [62] regarding the space variability, which results in almost similar results.

While this analysis only considers the security from Alice/Bob side, studying from Eve's side needs further investigations, which necessitates modeling the correlation between Eve's observations and those experienced by Alice/Bob. Therefore, this more complicated scenario is addressed in the following chapters where more realistic channel models, e.g. deterministic ones, are considered.

3.3.3 Joint space-frequency variability

Intuitively, the smaller the coherence bandwidth, the more efficient will the SKG be able to exploit frequency variability. Unfortunately, the coherence bandwidth changes from an environment to another and is out of control. SKG performance should be achieved also in environments where the coherence bandwidth is small, which is a difficulty when no sufficient spatial variability is provided. As a way of mitigation, we propose to exploit jointly the space and frequency degrees of freedom, so to relax the requirements on each of both individually. A potential use case is that

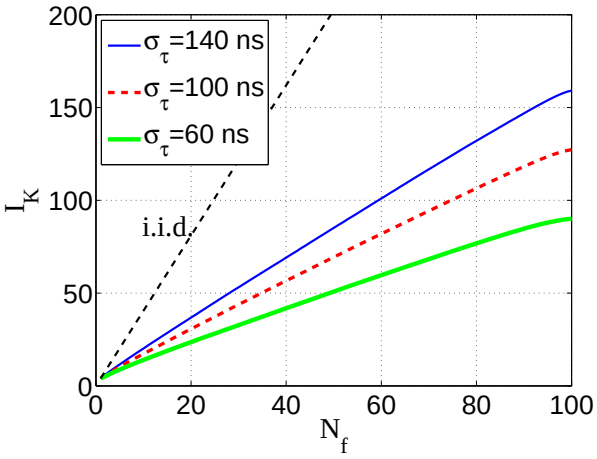
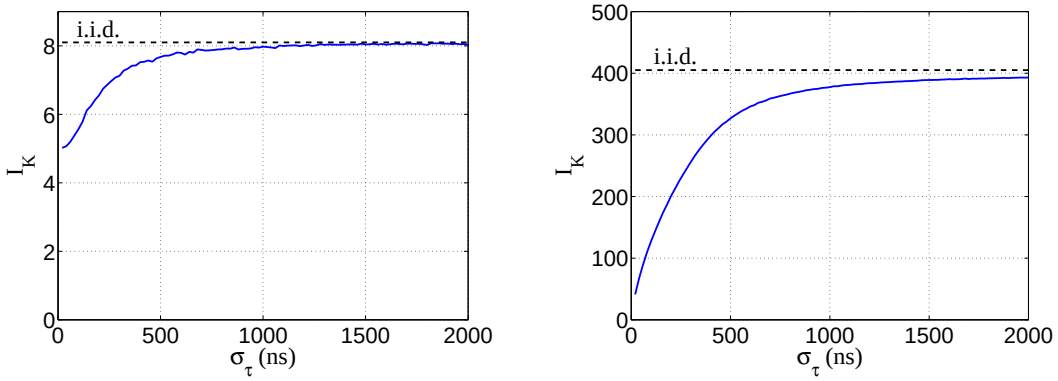


Figure 3-11: Evolution of I_K with respect to N_f and to RMS delay spread σ_τ, for $\Delta_\tau = 20$ ns.



(a) $N_f = 2$ (b) $N_f = 100$

Figure 3-12: Variation of I_K with the RMS delay spread for $\Delta\tau = 20$ ns.

of MIMO systems (such as IEEE 802.11n/ac), providing spatial variability, together with OFDM technology providing frequency variability.

3.4 Conclusion

We present in this chapter the relevant metrics, including theoretical bounds and that related to the quantization algorithm, which are used in the present dissertation to assess the SKG performance. Furthermore, we have discussed the impact of using various channel degrees of freedom, including spatial and frequency domain, on the security behavior.

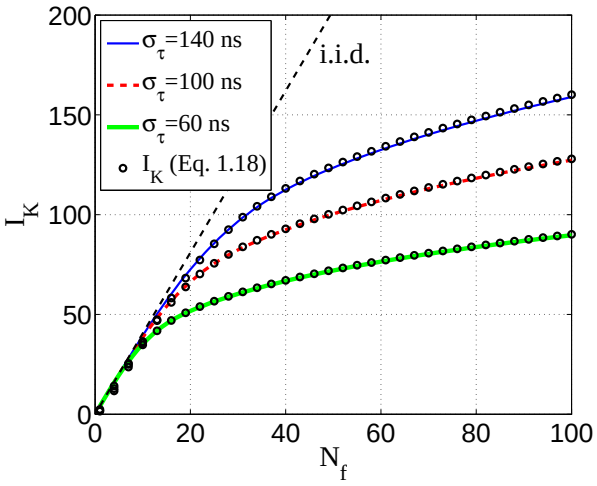


Figure 3-13: I_K as a function of N_f for $BW = 50$ MHz.

In particular, we address the potential SKG performance from delay dispersive channels. This has been done through a simple model of CIR, with equidistributed path delays, an exponentially decaying PDP and uncorrelated, Rayleigh faded path amplitudes. Although simple, this model has allowed to identify major features in terms of secret key rate (or available number of key bits) from a CIR, which can be summarized as follows:

- For a fixed sub-carrier interval in the frequency domain, the secret key rate increases roughly proportionally to the number of frequency sub-channels, although more slowly than for i.i.d. sub-channels, until a first breakpoint for which the number of available key bits equals the number of DoF computed from the noise level for independent paths. Successive breakpoints occur at each time the bandwidth "resonates" with the inverse of the inter-path delay. This means that at very high BW, the increase in the secret key rate gets lower and lower.

- When the bandwidth is fixed, increasing the number of sub-carriers nearly proportionally increases the key rate, at a "speed" equal of that for i.i.d. sub-channels. This comes from the fact that for few sub-carriers, their frequency separation is large and they can be assume uncorrelated. A more slow progression occurs below a certain number of sub-carriers, because of the onset of frequency correlation. This onset comes later, i.e. the available number of key bits can be higher, if the delay spread is higher, because of a smaller coherence bandwidth.

We will further see, in the following chapters (5 and 6), to which extent these observations would be seen in more sophisticated channel models as well as in nearly realistic channel models.

Chapter 4

Performance of secret key generation in non-stationary channels

In a PhySec perspective, the confidentiality of any wireless communication heavily relies on the scenario of the Alice/Bob/Eve trio [68, 106]. On one side, the propagation scenario of Alice/Bob (e.g. SNR, channel degrees of freedom, richness in dispersive multipaths, etc.) determines the maximum amount of random information that would be reliably shared between legitimate users. On the other side, the relative propagation scenario of Eve (e.g. SNR, distance, channel fading, etc.) affects the amount of shared bits that would be useful for the generation of a "secret" key. Therefore, given that the SKG is simply targeted from the Alice/Bob sides in the previous chapter (Section 3.3.2), we intend in the present chapter to deeply investigate the relation between the Eve propagation scenario and the secret key quality. In fact, the SKG analysis with respect to Eve is more critical and requires a multi-user channel model that accounts for the spatial correlation between two users in proximity of each other, i.e. Bob and Eve.

In [58], a simple channel model is proposed where Bob and Eve are assumed very close to each other, so they share the same multipath components (MPC). This scenario may be valid in wireless sensor networks in which sensors may be located in the same stationary region[1]. More general scenarios should account for larger separation distances between Bob and Eve, which mean not only several wavelengths but

[1]By stationary, we imply an area/volume of space where the statistics of the radio channel are constant

ultimately up to distances such that all radio channel characteristics are completely different (independent in mathematical terms). When this is achieved, the environments seen by Alice/Bob and Eve are necessarily different, which means that they are not in the same stationary spatial domain. Such a scenario is targeted in our previous work [72] through a very simple version of a "Geometry based Stochastic Channel Model" (GSCM), relying on the presence of scatterers between Alice and Bob/Eve in order to represent the propagation events at the origin of multipaths. However, the proposed model did not account for the shadow fading and its spatial dependence, involed in the spatial correlation.

Nevertheless, multi-user channel simulations are allowed through some GSCMs (refer to Section 2.1.4 for propagation channel models state of the art). In particular, the Winner II model [1] enables the simulation of correlated channels through the filtering of large scale parameters such as shadow fading. However, the main drawbacks reside in that the channel parameters are valid in a low temporal duration and that radio links involving different propagation scenarios cannot be simulated. While these issues are resolved to some extent in the QuaDRiGa model [107], physical layer security protocols are investigated through both Winner II and QuaDRiGa models in the PHYLAWS project [15], for either WIFI or LTE simulations.

Both Winner II and QuaDRiGa models are parameterized for some given scenarios. However, we aim to understand how the security is impacted by simply changing the channel features. Hence, in this chapter, we elaborate on the non-stationary channel model previously presented in [72]. We hence present further developments in order to better account for shadowing, which is a well-known feature of space variant channels beyond a few wavelengths. In order to account for channels varying over macroscopic distances and impacted by shadow fading, such an effect must be specifically involved in the model.

4.1 Shadowing

In narrow band systems, the total received power slowly fluctuates around the distance dependent path loss, following a log-normal distribution (i.e. in dB, a zero-mean normal distribution with a shadow fading standard deviation σ). In fact, such a phenomenon results from the obstruction of some paths due to the variation in the surrounding environment (i.e. interacting objects) around either the mobile station or the base station. As a consequence, in wide band systems, shadowing occurs per

clusters that gather multipath components of similar delays and angular characteristics [1, 52, 108, 109]. Measurements in outdoor urban environments showed that each cluster undergoes independent log-normal shadow fading with different standard deviation values [110].

While the correlation properties of the shadow fading process are targeted through the literature for the design of macroscopic diversity and handover schemes [37], such properties seem essential in the modeling of spatial variant channels in the context of PhySec. Typically, two types of spatial correlations are defined: the auto-correlation and the cross-correlation, both illustrated in Fig. 4-1. The former considers the correlation between two links connecting the base station (BS) to two locations of the mobile station (MS) at two different instances, while the latter corresponds to two links connecting the same MS to different BS. Several shadow fading correlation models are proposed and addressed in the literature. A summary could be found in [111]. In particular, Gudmundson [37] proposed a distance-dependent exponential decaying function to account for the spatial auto-correlation function, while the authors in [54, 112] consider the angle difference between the links in the modeling of the shadowing cross-correlation. Nonetheless, without loss of generality, we use in the present work the term of shadow fading correlation without differentiating between the auto-correlation and the cross-correlation [111].

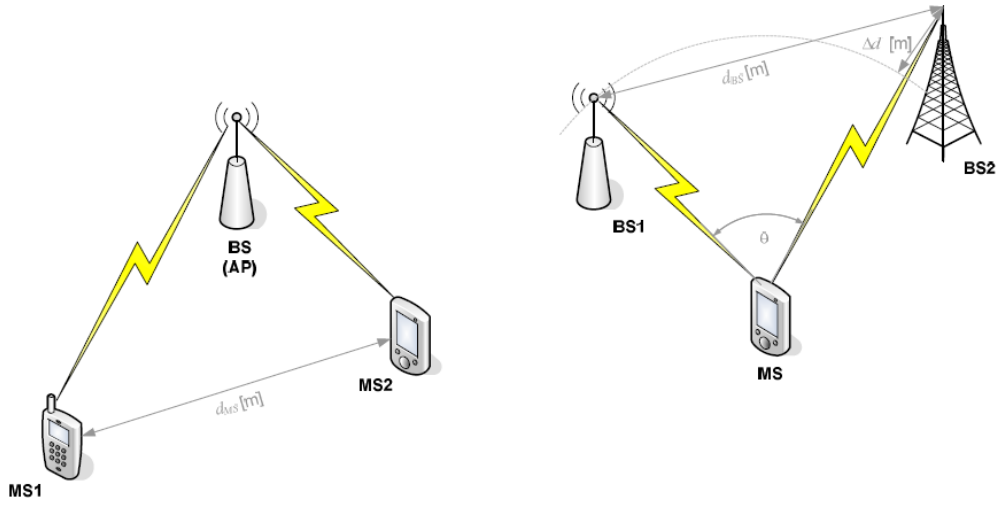


Figure 4-1: Shadow fading auto-correlation (right) and shadow fading cross-correlation (left) [1].

4.2 Disc of scatterers based channel model details

In order to accurately study the security provided by the SKG technique against any passive eavesdropper, we should address realistic cases of interest for the locations of Alice, Bob and Eve. Alice and Bob can commonly be a cellular base station (above or below roof tops of macro, micro or small cells), an indoor access point, or a terminal (either a fixed or a mobile one). More critical is the relative position of Eve in the Alice/Bob environment since this impacts the amount of correlated information between the channels seen by the trio. While it is unlikely that Eve can be very close to Alice/Bob when the latter is a BS, this is not fully impossible. Alternatively, Eve can potentially be very close, moderately far or very far from Alice/Bob. This may rely, respectively, on the wavelength, on the shadow fading distance or on the long distance attenuation. In other words, the location of Eve with respect to e.g. Bob may determine the distinct spatial characteristics between the channels seen by Bob and Eve from Alice. These differences can be ascribed to differences in some or all the multipath compenents (MPC) in terms of e.g. path amplitude and path direction, even polarization.

Briefly speaking, the channel MPCs seen by Bob and Eve can change according to the relative distance between them and also according to the environment. For example, the propagation channel components are likely to be more sensitive to the separation distance in a dense scattering urban environment than in a rural one. Therefore, in order to study the effect of the lack of spatial stationarity between Bob and Eve on the SKG, we consider a 2-D GSCM, where scatterers are uniformly distributed within a disc [113, 114] centered at Bob (Fig. 4-2). Eve is located within the disc at a distance d from Bob. The maximum separation distance is kept to a value low enough to avoid edge effects due to the finite size of the disc. Furthermore, the transmitter, Alice, is supposed far away from the disc so that we can consider rays arriving from a single direction $\vec{K}_A$ to the local scatterers. This situation occurs mostly in urban macro-cells, when the BS is located over rooftops and the angular scenario at the BS level is very directional. Each terminal is considered to be in non line-of-sight condition with respect to Alice. Hence, all the rays received by Bob/Eve originate from the scatterers, acting as secondary sources. We also assume that Bob and Eve are both equipped with an omnidirectional antenna.

In GSCMs, interacting objects over which the impinging wave is either reflected, diffracted or scattered are represented by discrete scatterers distributed in the surrounding medium. The scattering coefficient assigned to each scatterer may be mod-

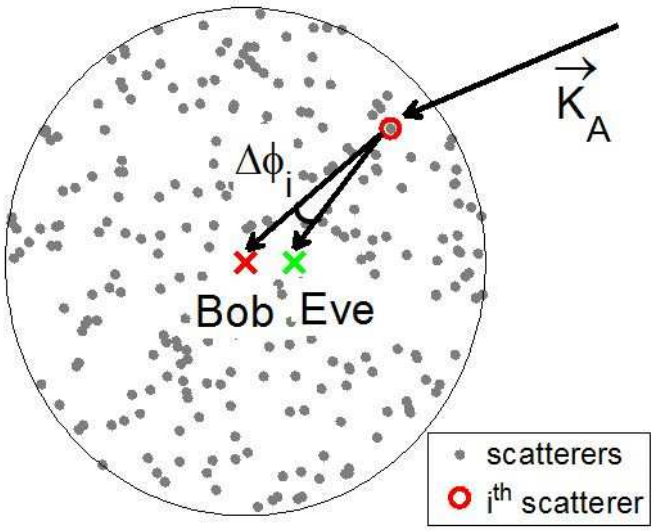


Figure 4-2: Geometrical representation of the communication scenario.

eled according to the physical interaction type, as shown in [52, 108] where the authors differentiate between the LOS path, the specular interactions (i.e. reflection and diffraction) and the diffuse scattering. However, for the sake of simplicity, we do not consider in the present work a specific physical mechanism but a more general one, where the impinging wave is re-radiated in different directions. Accordingly, each scatterer is assumed to act as a non-omnidirectional lossy re-transmitter, which is statistically independent from the others. Hence, Bob and Eve may not see the same power of rays. Moreover, in order to account for the shadow fading correlation between Bob and Eve, we assume that each scattered path emitted from the same scatterer is spatially correlated according to the following correlation coefficient [54, 112], which is based on an angle rather than a distance:

$$\rho_i = 0.5 + 0.5\cos\{\Delta\phi_i\} \tag{4.1}$$

where $\Delta\phi_i$ is the angle of departure difference at the ith scatterer, as shown in Fig. 4-2. More clearly, if S_{Bi} and S_{Ei} are the shadowing gains of the rays emitted by the ith scatterer towards respectively Bob and Eve,

$$E\{S_{Bi}S_{Ei}\} = \rho_i. \tag{4.2}$$

This may be achieved when:

$$S_{Bi} = a_{Bi} \tag{4.3}$$

and

$$S_{Ei} = \rho_i a_{Bi} + \sqrt{1-\rho_i^2} a_{Ei} \tag{4.4}$$

where the shadowing coefficients a_{Bi} and a_{Ei} are i.i.d. and follow a normal distribution with zero-mean and a standard deviation σ in dB. The correlation ρ_i accounts for the similarity/difference of the surrounding interacting objects around Bob and Eve. This is not revealed in the distance decay exponential model [37] which is rather preferred for especially closer receivers.

According to the scatterers distribution, physical path structures towards Bob/Eve, including directions, are determined from a simple geometrical relationship; hence, the multipath fading channel can be computed. Therefore, the narrow band single input single output (SISO) channel seen by Bob/Eve is defined as follows:

$$h_X = \sum_{i=1}^{N_S} \frac{10^{S_{Xi}/20}}{d_{Xi}} \exp[j(K d_{Xi} + \vec{K}_A.\vec{r}_i)] \tag{4.5}$$

where N_S, d_{Xi} and $\vec{r}_i$ are respectively the number of scatterers within the disc, the distance from an ith scatterer to X (Bob/Eve) side and the ith scatterer coordinate. Moreover, $K = 2\pi/\lambda$ and $\vec{K}_A$ are respectively the wave number and the wave vector of the plane wave emitted by Alice towards the disc. This equation also expresses that the power of the incoming wave from Alice is diffused by scatterers and attenuated by free space propagation according to the separation distance towards Bob/Eve [52, 108, 115, 116]. We note that, while we implement in our model the shadow fading correlation per diffused scattered paths, such correlation is applied merely for both LOS path and specular interactions in [52, 115].

In order to account for the channel estimation errors and noise, the channels are assumed corrupted by noise (see Eq. 2.10), as $\hat{h}_X = h_X + n_X$ where n_X is the noise estimation which can be modeled as zero-mean complex Gaussian random variable with variance σ_n^2. X denotes here A (Alice), B (Bob) and E (Eve). Moreover, the noise at the different terminals is assumed independent.

Usually, the shadow fading is accounted for once the small-scale fading (SSF) is removed by averaging [37]. However, in the context of PhySec, SSF is essential in providing randomness from which robust keys may be extracted. Therefore, we model both small scale fading and shadow fading (per path), where SKG metrics are assessed from SSF statistic while maintaining fixed the shadow fading parameters. Then a statistic on the SKG metrics may be derived owing to the shadow fading statistic. More clearly, we define how the random variables change according to these

two fading types, as follows:

- Shadow fading statistic: It is defined by different realizations of the environment, characterized by the macroscopic scatterers positions and the shadowing coefficients (a_{Bi} and a_{Ei}). This statistic leads to get a set of keys, or equivalently to get a set of each SKG metric, e.g. a set of I_K or I_{VK} values.

- SSF statistic: For a fixed shadow fading realization, where the macroscopic environment around Bob and Eve is fixed, scatterers are allowed to randomly move on a square grid of surface $5\lambda \times 5\lambda$ centered at the fixed macroscopic position, providing SSF channels through varying phases over 0 to 2π. In other words, SSF is provided by maintaining the same paths amplitudes and changing the phases according to the small scale spatial movement of scatterers. This SSF statistic is essential in order to generate a single key, or equivalently to get a single value of each SKG metric.

4.3 SKG performance evaluation

Table 4.1 presents the simulation parameters values. We assume that all terminals have the same SNR and that 250 scatterers are distributed uniformly within the disc, unless differently stated. The maximum separation distance is 1000λ providing equivalent statistics for both Bob and Eve. Each statistical quantity (correlation coefficient, information theoretic key bound and BER value) is computed from an SSF statistic with 250 realizations. Then, statistical distributions for these quantities are obtained from the combined macroscopic scatterers randomness and the shadow fading parameters randomness, with also 250 realizations. Thus, with 62,500 realizations in total, the results are considered quite accurate.

Table 4.1: Simulation parameters.

Frequency	2 GHz
SNR	15 dB
Disc radius	5000 λ
Scatterers number N_S	250

4.3.1 Channel correlations

As a qualitative assessment of SKG performance, we compute both the complex channel correlation (ρ_{BE}) and the envelope power correlation as in Eq. 3.13, involving respectively the complex channel coefficients seen by both Bob and Eve, and their powers. Fig. 4-3 shows the cumulative distribution functions (CDF) of complex channel correlation coefficients for both shadow fading standard deviation $\sigma = 3$ dB and $\sigma = 10$ dB, and for several separation distances d between Bob and Eve.

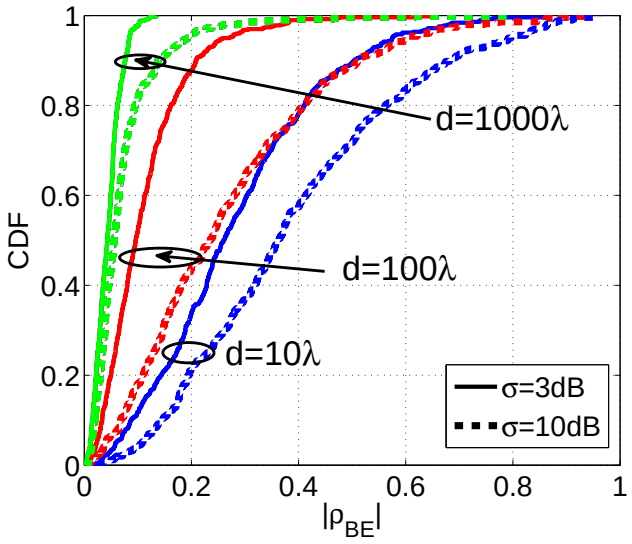


Figure 4-3: CDF of complex channel correlations.

Obviously, as Bob/Eve's separation distance d increases, the complex channel correlation $|\rho_{BE}|$ decreases. Indeed, this is ascribed to the fact that Bob and Eve do not see the same interferences as d increases since they experience different multipath structures, in particular the phase. Moreover, for $N_S = 250$, the average distance between scatterers is almost 600λ. For $d > 600\lambda$, the interferences seen by Bob and Eve become independent and $|\rho_{BE}|$ vanishes.

Furthermore, we notice that $|\rho_{BE}|$ increases when σ increases. Usually, increasing σ yields more rapidly channel decorrelations (e.g. σ is higher in NLOS channels than in LOS one, since the former may undergo more severe shadow fading). However, this is not the case here. This is due to the fact that the shadow fading is implemented per path and with the same standard deviation. In fact, the tail of the log-normal distribution, especially for $\sigma = 10$ dB (Fig. 4-4), indicates relatively rare instances in which there is a dominant scattering coefficient (i.e. $10^{S_{X_i}}$) with a relatively high power. Such a power may be either amplified if the scatterer is very close to the

terminal or attenuated in the opposite case, and thus owing to the free space attenuation. As a consequence, the model is able to reproduce both Rayleigh and Rician distributions. The proportion of these two distributions is impacted by the value of σ. When σ increases, the proportion of Rician channels increases. We here consider that channel amplitudes are Rician distributed if the Rician K factor is greater or equal to 1. The results show that almost 18% of the channels are Rician for $\sigma = 3$ dB whereas we have 30% for $\sigma = 10$ dB. Indeed, we find that the correlation increases for Rician channels, where a dominant predictable path exists. This is also the reason why we still have large correlation values for large separation distances.

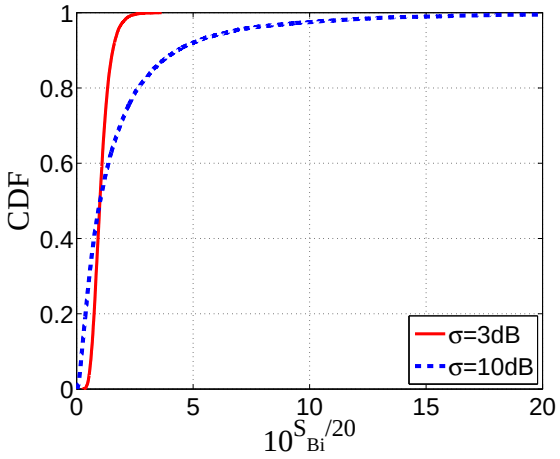
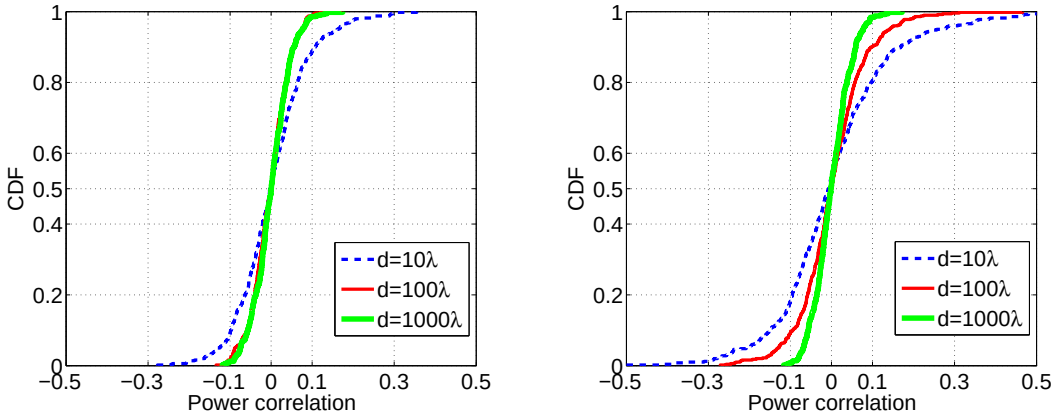


Figure 4-4: Log-normal shadowing per path for both $\sigma = 3$ dB and $\sigma = 10$ dB.

Fig. 4-5 depicts the CDFs of the power correlations for several Bob-Eve distances, for both $\sigma = 3$ dB and $\sigma = 10$ dB. As expected, the power correlation decreases when d increases. When Eve goes away from Bob, they see different multipath components, leading to a decrease in both complex and power envelope correlations. Moreover, due to the high proportion of Rician channels for $\sigma = 10$ dB, the variance of the power correlation is the largest in this case, which can be explained by more correlations resulting from less significant scatterers in the presence of a dominant path.

4.3.2 Vulnerable key rate

According to the chosen SNR, the maximum number of key bits I_K is nearly equal to 4.05 per channel observation. Fig. 4-6 shows the statistics of the relative vulnerable key rate I_{VK}/I_K as a function of several separation distances d between Bob and Eve, for both $\sigma = 3$ dB and $\sigma = 10$ dB. We recall that these statistics are computed over the



(a) $\sigma = 3$ dB (b) $\sigma = 10$ dB

Figure 4-5: CDFs of the power envelope correlations.

250 shadow fading realizations. The average and the variance values of the I_{VK}/I_K statistics are shown respectively in Fig. 4-7 and Fig. 4-8. While all these results only consider the case of $N_S = 250$ scatterers within the disc, Fig. 4-7 considers also the case where $N_S = 100$.

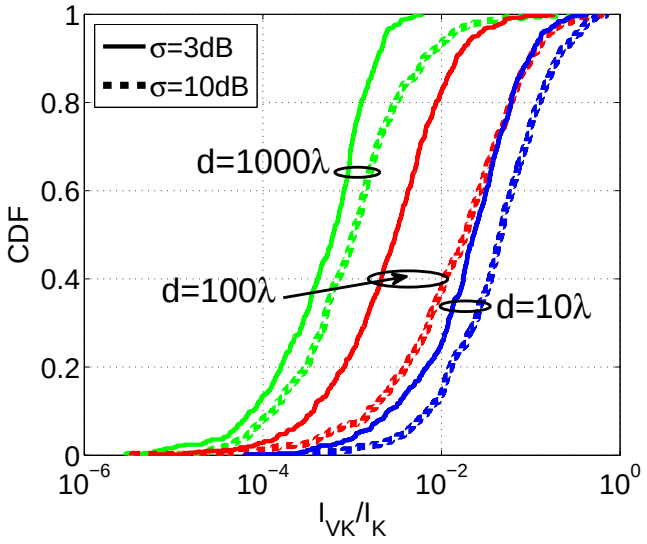


Figure 4-6: CDF of relative vulnerable key bits.

The results presented in both Fig. 4-6 and Fig. 4-7 show that I_{VK}/I_K decreases as either d increases or σ decreases. These results are consistent with those obtained for the channel correlations, revealing the strong relation between I_{VK} and the spatial channel correlations between Bob and Eve. For the smallest value of d (i.e. $d =$

0.1λ), Bob and Eve nearly see the same MPCs with high shadow fading correlation, which results in relatively high vulnerability. When Eve moves away from Bob while remaining very close (i.e. $d \leq \lambda/2$), the channel phase shifts decorrelate rapidly, yielding a high decrease in the vulnerability rate. Then, with larger values of d, it is apparent that the decrease on I_{VK}/I_K is quite progressive owing to the long distance spatial memory of the channel, even for relatively different structures of MPCs seen by Both Bob and Eve. Finally, as d gets larger values, the angle difference at each scatterer $\Delta\phi_i$ increases, yielding a less shadow fading correlation and subsequently I_{VK}/I_K vanishes.

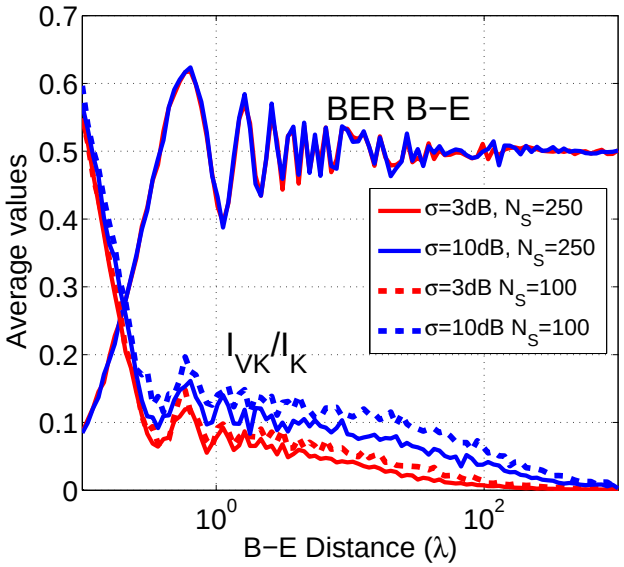


Figure 4-7: Averaged BER and vulnerable key bits as function of the separation distance d between Bob and Eve.

Furthermore, if the density of scatterers N_S decreases, the effective number of scatterers (i.e. the number of scatterers with relatively significant power) seen by Bob and Eve decreases, resulting in more vulnerability. Moreover, regarding the variance of I_{VK}/I_K shown implicitly in Fig. 4-6 and explicitly in Fig. 4-8, the security behavior changes according to each environment realization with more significant variation for moderate d values and for large σ values. This heavily relies on the effective number of scatterers that contribute to the most channel power and also to their relative position with respect to Bob/Eve which affect the shadow fading correlation. Briefly speaking, good security performance is provided for dense multipath propagation channels corresponding to the lower tail of the CDFs, whereas it is degraded for environments where a predictable dominant path exists corresponding to the higher tail of the CDFs.

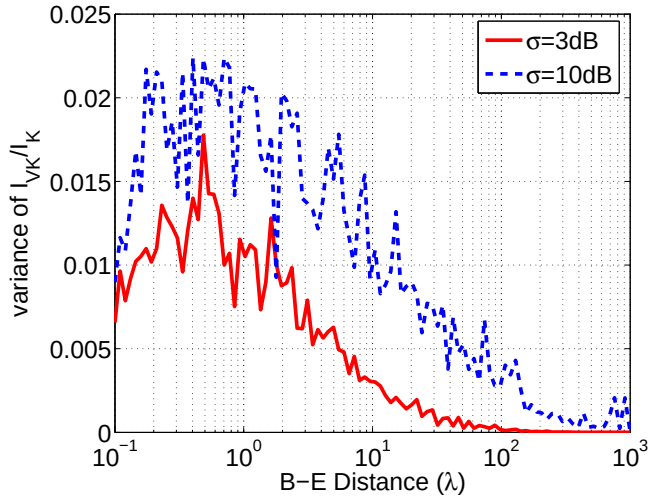


Figure 4-8: The statistical variance of I_{VK}/I_K as a function of d.

4.3.3 CQA performance

Alice, Bob and Eve quantify their complex channel coefficients into stream bits by extracting 1 bit from each I and Q parts (i.e. $M = 4$ quantization regions provide 2 bits), according to the channel quantization alternating (CQA) algorithm explained in Section. 3.2. Fig. 4-9 shows the CDFs of the bit error rate (BER) of Bob/Eve keys, for different values of σ and d, while Fig.4-7 shows the average values of the BER in addition to the I_{VK}/I_K. Although we have almost the same average value of the BER for different values of both d and σ, the behavior changes from one environment to another, as shown implicitly by the variance of each CDF, see also in Fig. 4-10. While σ does not impact the mean of the BER whatever d, it impacts the variance of the BER as shown in Fig. 4-10. It is interesting that the BER variance and the variance of I_{VK}/I_K have partly similar behaviors. Large values of σ favour few dominant scatterers, which results in more variability from one realization to another and consequently a higher variance on both related parameters.

When Eve goes away from Bob, the security is enhanced since the mean BER converges towards 0.5, which is consistent with the behavior of the average I_{VK}/I_K. Actually, the BER simply expresses the raw difference between the key bits directly extracted from the channel coefficients seen by Bob and Eve. The algorithm doesn't attempt to develop more powerful strategies in order to exploit the common characteristics between these channels. This is the reason why the remaining vulnerability expressed in I_{VK} beyond about one wavelength distance between Bob and Eve, is not reflected in the mean BER. However, we still see it in the BER variance.

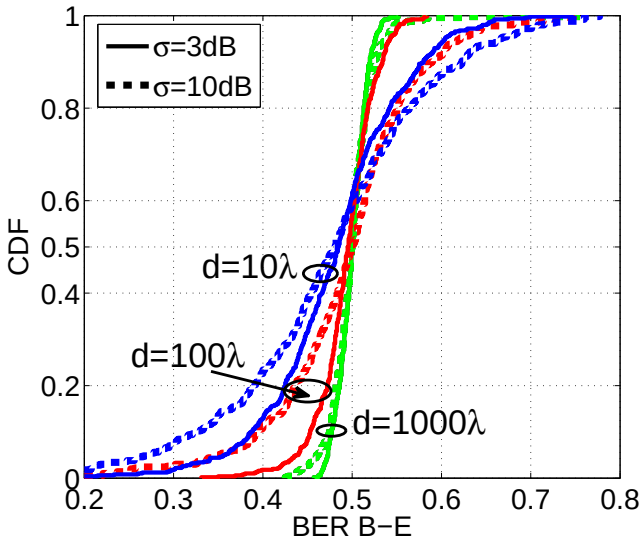


Figure 4-9: CDF of BER between Bob and Eve.

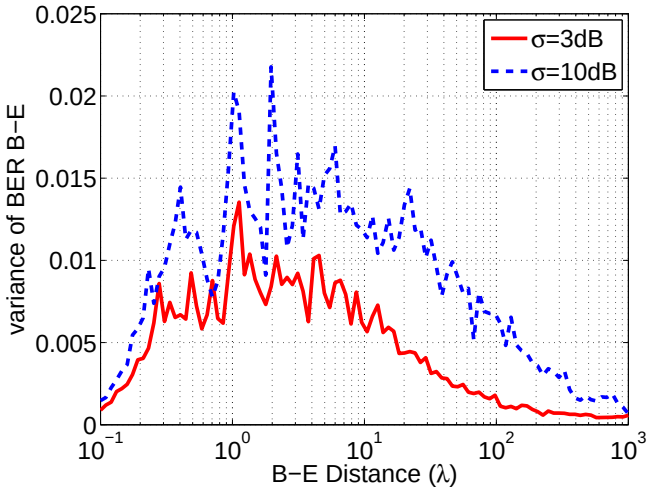


Figure 4-10: The statistical variance of BER as a function of d.

4.4 Conclusion

The GSCM model described above provides a realistic approach to channel richness (tunable through the number of scatterers) and shadowing, in addition to simple free space attenuation. This has been described by an extra log-normal attenuation and an angular dependent decorrelation between the shadow fading coefficients. The observation that, in spite of these features, the complex correlation coefficient between Bob and Eve's channels decreases relatively slowly with their distance (Fig. 4-3) demonstrates the role of long distance common characteristics. This expresses a

"spatial memory" of the channels, which is responsible for a tail in the number of vulnerable key bits (Fig. 4-7), as computed for Gaussian statistics generated from SSF channels for both Bob and Eve. This memory is not seen for the direct comparison between Bob and Eve's generated keys, making use of CQA. This means that such an algorithm, operating directly in the I-Q complex plane, is based on major changes of the channel as regards key confidentiality, but it cannot guarantee that a very clever Eve would not capture some of this long distance memory to acquire partial information on Bob's key.

Another unexpected feature is the enhanced correlation between channel coefficients for high shadowing coefficients, which may sound strange at first sight. This is basically an outcome of the model, which can be explained by the fact that under high shadowing, farther distant scatterers are less visible (the dominant ones, producing shadowing, mask them), which effectively reduces the channel richness. Although surprising, the observation is not meaningless. For instance, for an indoor scenario with thick absorbing walls, the channels are less rich and more spatially correlated than with partly transparent ones, where outside scatterers will contribute to multipaths.

Chapter 5

Security behavior through ray tracing channel models

The model developed and discussed in Chapter 4 is mainly intended to investigate the role of important physical phenomena, such as small scale fading and shadow fading, on the security performance provided by SKG. However, its role is not to be "realistic" with respect to the variety of possible environments concerned by the practical use of SKG techniques. In particular, as opposed to what is commonly done in radio channel modeling (see, e.g. Section 2.1.4), the parameters have not been, and are not intended to be, determined from channel measurement campaigns. A complementary approach, in this respect, is addressed in the PHYLAWS project [15], where the Winner II and its extension QuaDRiGa are considered for PhySec simulation purposes in LTE networks.

This being said, it is of interest to investigate SKG performance in a variety of environments and see what relation can be drawn between the characteristics of the channels and SKG security performance. Examples of site specific channel models are well known to be ray-tracing (RT) or ray-launching (RL) techniques, which compute the wave paths between the transmitter and the receiver deterministically, based on ray approximation (plane waves) and basic theory of the fundamental propagation events (reflexion, diffraction, transmission and the like). The main advantage of RT is that it takes into account the true scatterers, such as buildings, into the computation of rays between Tx and Rx. Its main drawback is the heaviness of computations, and above all the great difficulty in reproducing in the numerical model of the propagation environment the true and fine features of reality.

In this chapter, we present results on RT-RL simulations, which had been obtained

from a previous national (French) project and are re-used in the present SKG context. This may provide valuable added value and insight with respect to the model of Chapter 4, as well as of measurements results from the literature.

5.1 Environments and characteristics of the simulations

The radio wave propagation simulation software used here employs the Volcano tool commercialized by Siradel [117], which is a 2.5D ray-based method. Ray launching in the horizontal plane is used for determination of the multipaths angles of departure and arrival, while a modified Deygout method combined with a uniform theory of diffraction are used for each of these paths in the elevation plane [118]. An extension to this commercial tool has been privately provided by the comapny, in order to incorporate to some extent diffuse propagation. Indeed, it is well known that diffuse scattering (DS) plays an important role in the radio channel, up to a proportion that depends on the geometry, density and features of the scatterers. It is commonplace to consider that the energy carried through diffuse scattering can take between 20 % and 80 % of the whole energy in a channel impulse response (CIR). The DS is included using a model based on the Effective Roughness approach [119, 120]. Each wall surface is divided in a multitude of facets (tiles). Each of those radiates a scattering wave whose amplitude depends on a scattering coefficient S and on a directive scattering pattern, centered around the direction of specular reflection. An example of ray traces in the absence and in presence of DS can be seen in Fig. 5-1 for an indoor case.

The parameters involved in the reflexion/diffraction have been chosen according to proposed ones by the Volcano tool, based on a large set of previous measurement campaigns and parameters calibration achieved by SIRADEL. Indoor locations of mobile terminals (MT) are not taken into account, while the attenuation by vegetation is considered within the simulations. The simulation frequency used in the present analysis is 5.5 GHz.

The considered environments are locations in Paris and around, particularly:

- "Carrousel du Louvre (Paris)" (Fig. 5-2(a)) with two heights[1] for the base station antenna (Louvre4: 4 m and Louvre48: 48 m). The investigated zone had a size of 1200m x 360m and contained 4011 terminals positions (10m as spatial

[1]The heights are computed from the ground level.

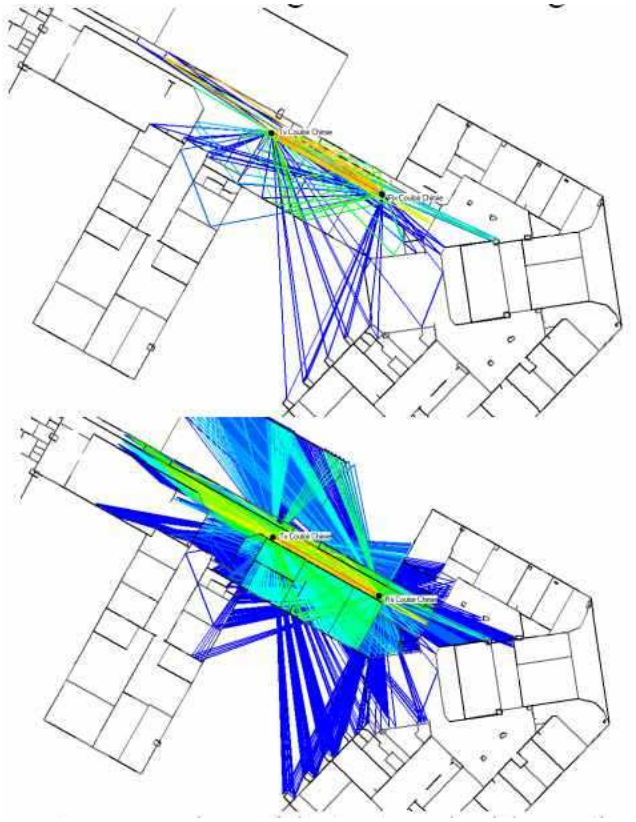
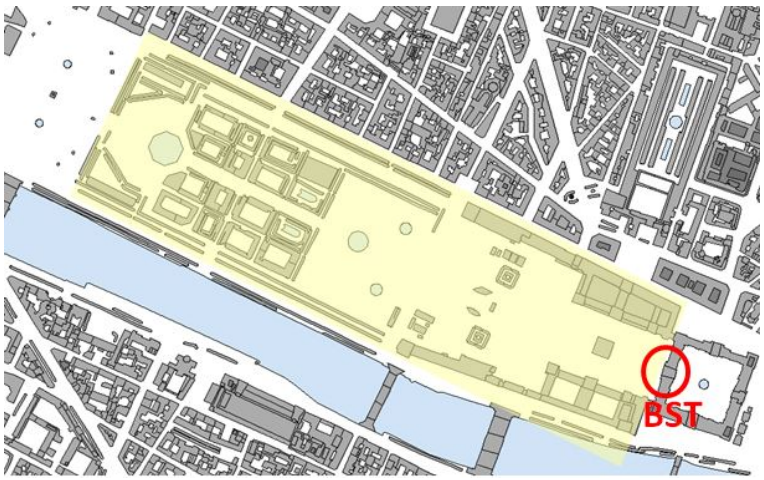


Figure 5-1: Example of an (indoor) deterministic simulation without (up) and with (down) diffuse scattering.

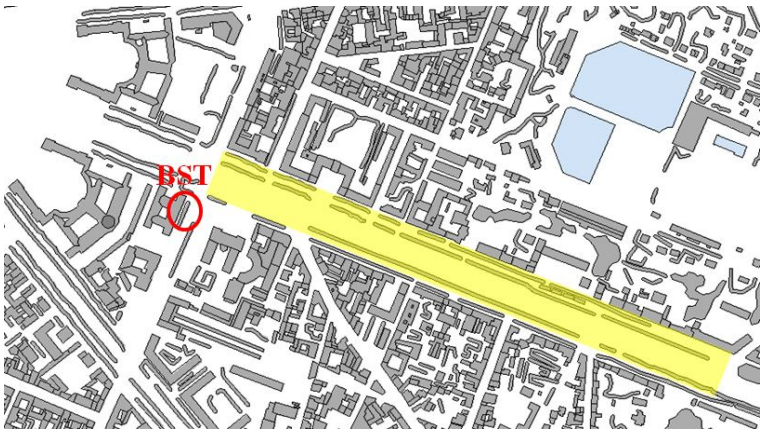
resolution).

- "Avenue de Paris (Versailles)" (Fig. 5-2(b)). The BST had a single antenna height (19 m), the investigated zone was of size of 850m x 60m and contained 2193 terminals positions (5m as spatial resolution).

Moreover, for the sake of comparison, we dedicate a scenario without diffuse scattering (WDS) implementation, which is the scenario Louvre48-WDS. It has the same description as Louvre48 except for the spatial resolution, which is 5m, and thereby of the number of terminals positions (15756).



(a) Louvre



(b) Versailles

Figure 5-2: The digital 2D maps highlighting the simulated area.

5.2 Propagation channel characteristics

5.2.1 Channel model

A ray tracing tool generally computes almost all electromagnetic waves propagating from the base station to a single or several receivers at different locations, taking into account the relevant propagation mechanisms such as reflection on walls and diffraction by building wedges and corners. Each received ray at position $\mathbf{r}$ is then characterized by its gain β_i, its time delay τ_i, its azimuth [2] angle of arrival (AoA) ϕ_i and its azimuth angle of departure (AoD) θ_i. The base station (BST) is at location $\mathbf{r}_t$. The combination of all computed N_{path} paths allows to calculate the parameter

[2] We do not consider elevation angles since the Volcano tool is a 2.5D tool.

of interest, i.e. the channel impulse response (CIR), as follows:

$$h(\mathbf{r},\tau) = \sum_{i=1}^{N_{path}} \beta_i \exp\{j(\mathbf{k}_{Ri}.\mathbf{r} - \mathbf{k}_{Ti}.\mathbf{r}_T)\}\delta(\tau - \tau_i) \tag{5.1}$$

where the wave vector at the receiver is defined as $\mathbf{k}_{Ri} = \frac{2\pi}{\lambda}[\cos\phi_i \ \sin\phi_i]$ while it is defined at the transmitter side as $\mathbf{k}_{Ti} = \frac{2\pi}{\lambda}[\cos\theta_i \ \sin\theta_i]$. λ is the wavelength.

The channel transfer function is then obtained by performing a Fourier transform on the CIR, as follows:

$$H(\mathbf{r},f) = \sum_{i=1}^{N_{path}} h(\mathbf{r},\tau_i)\exp\{-2\pi f\tau_i\} \tag{5.2}$$

We notice that the measured channels are space-varying and thereby the number of paths N_{path} changes from one location to another as well as the parameters of each multipath component (MPC).

Given that commercial receivers have limited performance, we just consider terminals' positions where the path loss verifies $PL < 125$ dB. Naturally, it is useful to restrict the analysis to the locations where the predicted received power does not fall below the receiver noise level. Moreover, we consider a receiver with a limited bandwidth of 200 MHz, which results in a time delay resolution of 5 ns. Accordingly, the paths falling into the same delay bin $\Delta\tau$ are vectorially added in the i-Q domain. Furthermore, the frequency separation is set to $\Delta f = 312.5$ KHz, corresponding to the interval between subcarriers in the IEEE 802.11.ac standard, which results in a maximum excess delay of 3200 ns, beyond which all computed paths are simply discarded.

5.2.2 Small scale fading statistics

As discussed throughout this dissertation, random channel fluctuations provided by SSF are essential for robust SKG. Since MPCs are predicted on positions separated by at least 5 m, which does not support SSF statistics, we need to reconstruct SSF channels. For that reason, we adopt the plane wave assumption and consequently, simply compute CIRs by deterministically shifting the phase of each path according to the distance. In this context, channel responses are estimated over a local area modeled by a 5λx5λ square grid, with a step of $\lambda/5$, centered at each terminal position. Subsequently, the Rician K factor is computed by fitting the amplitude statistics into

a Rician distribution, which includes pure Rayleigh fading for the extreme case where $K = 0$. Since an ideal Rayleigh distribution can never be achieved in practice, channels with $K \leq 1$ are here assumed to follow Rayleigh fading.

Table 5.1 presents the number of retained channels as well as the ratio of Rayleigh channels with respect to the former number. We notice that the ratio of Rayleigh channels is very small, whatever the considered scenario. This is due in part to the nature of the propagation environment, e.g. an open space for the Louvre environment. This may also in part be due to limitations on RT simulations, such as 1) the limited number of physical interactions, 2) the simplified model of diffuse scattering and 3) the simple description of the required maps (e.g. EM maps and geometrical maps).

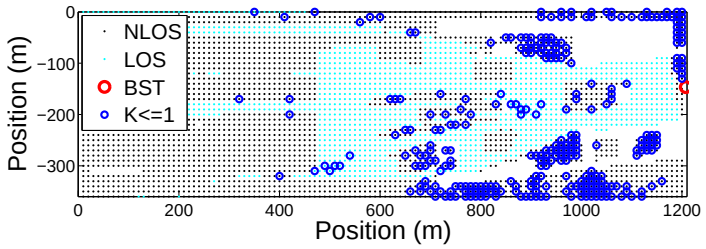
Table 5.1: Statistics of the computed channels.

Scenario	retained channels	Rayleigh (%)	Rayleigh passing K-S test (%)
Louvre48	2605	11.44	8.68
Louvre4	1922	12.75	5.2
Versailles19	1469	6.19	4.56

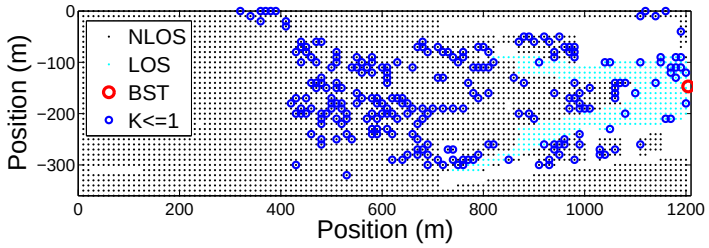
Fig. 5-3 displays, for each scenario, a 2D map localizing the BST as well as the total receiver positions, including LOS and NLOS conditions. Moreover, we marked receiver positions where $K \leq 1$. Regarding the Louvre environment, we find that, for higher antenna height (i.e. 48 m), receivers with $K \leq 1$ are located behind or on the corner of buildings, as seen in Fig. 5-3(a). However, for a smaller antenna height (i.e. 4 m) as shown in Fig. 5-3(b), the aforementioned positions are difficult to reach with several interfering paths, while channels with $K \leq 1$ are distributed in the open area as the receiver goes farther from the BST, yielding more attenuation in the dominant path. Regarding the scenario Versailles19, the positions with $K \leq 1$ are concentrated behind buildings.

5.2.3 Delay and angular spreads

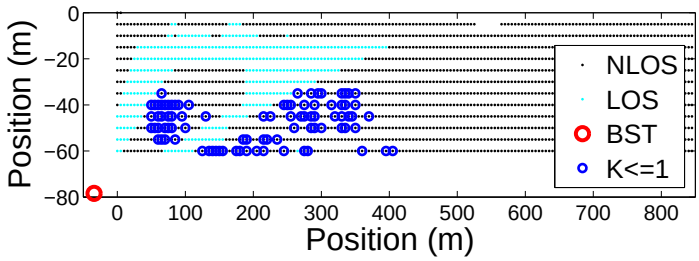
Both RMS delay spread (RMS-DS) σ_τ and angular spread σ_ϕ, given respectively in Eq. 2.5 and Eq. 2.8, are computed by just considering paths within 20 dB from the highest peak of the power delay profile (PDP) and the angular power spectrum (APS). We note that these spreads are evaluated at the MT side. Fig. 5-4 plots the statistics of the RMS delay spread for the 4 scenarios, differing between Rayleigh and Rician channels. Intuitively, Rayleigh channels provide RMS delay spreads higher than those



(a) The scenario Louvre48



(b) The scenario Louvre4

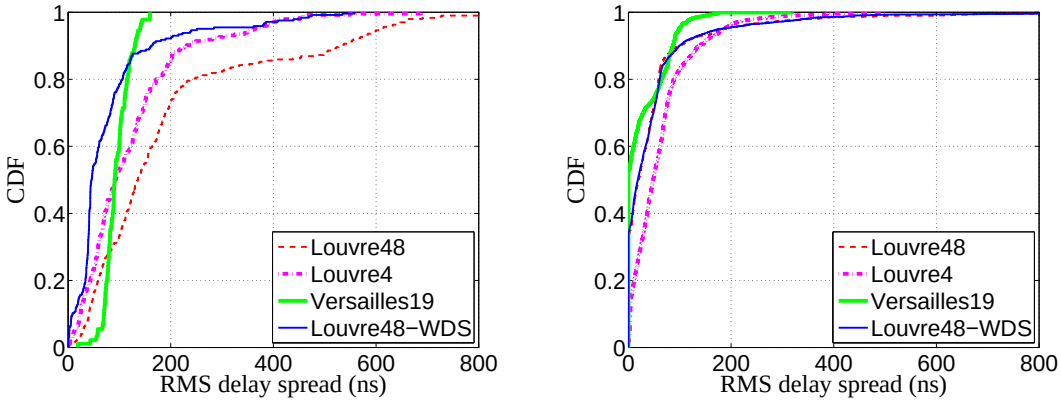


(c) The scenario Versailles19

Figure 5-3: The measured receiver locations in 2D maps.

observed in Rician fading, owing to the presence of a dominant path in the latter case. In particular, at 80% of the CDF, the maximum of the RMS-DS (90 ns) of Rician channels is less than the minimum (104 ns) of that of Rayleigh channels.

We focus our analysis on Rayleigh channels, which are of particular interest in generating randomness used to establish secret key bits. Accordingly, we consider now the RMS-DS of channels having $K \leq 1$, presented in Fig. 5-4(a). Given that time dispersion relies on geometric relationships between transmitter, receiver and the surrounding physical area, we remark that a higher antenna provides more time dispersive paths, when considering the Louvre scenarios. Indeed, as already discussed, the locations of Rayleigh channels differ between Louvre48 and Louvre4, where for the former more diffused paths result from the interaction with buildings. Furthermore, channels without diffuse scattering implementation obviously show smaller de-



(a) Rayleigh channels ($K \leq 1$). (b) Rician channels ($K > 1$).

Figure 5-4: RMS delay spread statistics.

lay spread values. Furthermore, the scenario Versailles19 reaches the smallest RMS-DS since the avenue is narrow relative to the large open area of the Louvre. The path lengths are smaller for the former scenario [121]. Nonetheless, small values of σ_τ may due to limitations in the RT simulations.

Some examples of PDPs corresponding to particular values of σ_τ are shown in Fig. 5-5. Considering the plots from right to left, the RMS delay spreads are respectively 15 ns, 360 ns and 728 ns. Particularly, we point out that, although the left PDP offers the highest delay spread, the channel presented by the middle PDP is the richer in MPCs within even the 10 dB threshold. Furthermore, we note that the discontinuous character of the PDPs stems from the discrete channel simulations but also from the limitations of the RT tool, which is not able to reproduce the full multipath density of a real channel.

Fig. 5-6 shows the statistics of the angular spread computed in the different scenarios for either Rayleigh or Rician channels. The same analysis as that done for the RMS delay spread can be made here. We recall that angular spread is statistically larger for Rayleigh than for Rician channels. Among the Rayleigh fading channels, Louvre48 offers higher angular spreads than Louvre4 owing to the relevant antenna height in the former scenario. We notice also the relevance of implementing diffuse scattering in enriching the radio propagation in more dispersive paths. Furthermore, some examples of APS are plotted in Fig. 5-7, focusing on particular values of the angular spread, corresponding to channels with $K \leq 1$.

In the following, we just retain Rayleigh channels that pass the Kolmogorov-Smirnov (K-S) test in order to adequately calculate the theoretical key bounds. The

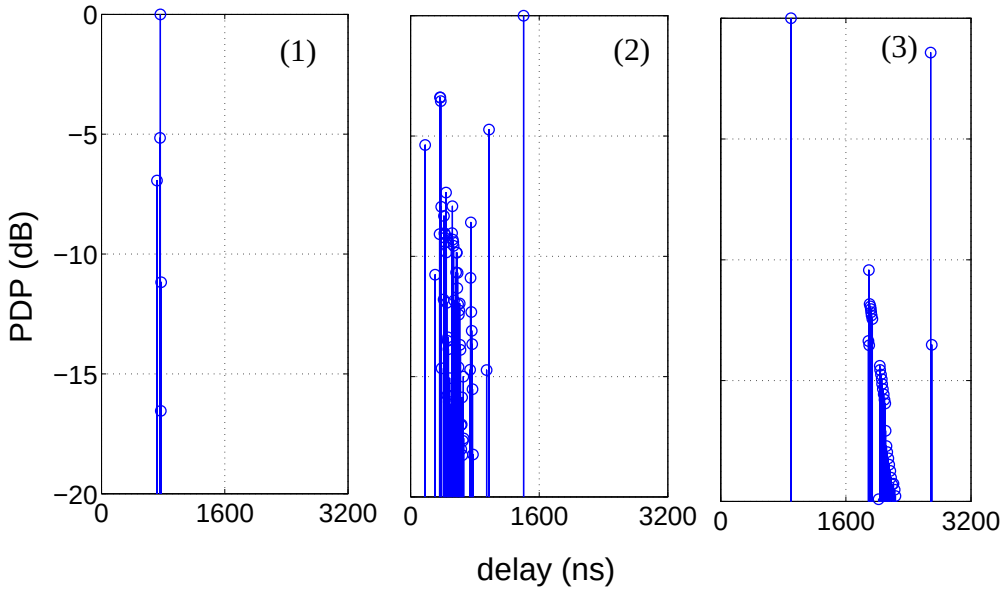
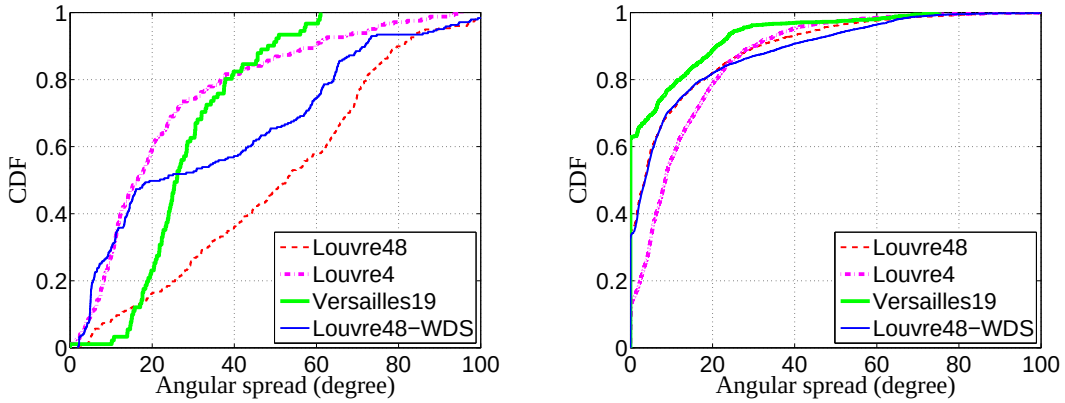


Figure 5-5: Some examples of PDPs corresponding to Louvre48 (right: $\sigma_\tau = 15$ ns, middle: $\sigma_\tau = 360$ ns, left: $\sigma_\tau = 728$ ns).



(a) Rayleigh channels ($K \leq 1$). (b) Rician channels ($K > 1$).

Figure 5-6: Angular spread statistics.

K-S test relies upon the computation of a statistical parameter, which is the maximum distance between the CDF of the empirical channel $F(h)$ and that of a reference distribution, here being the Rician distribution $F(h_{Rice})$. Explicitly, the K-S statistic is expressed as $D_{KS} = max(|F(h) - F(h_{Rice})|)$. The goodness of fit is revealed once D_{KS} is compared to a threshold according to a given significance level ($\alpha = 5\%$ being a commonly used value). Although we use the Rician statistic in the test, we just keep

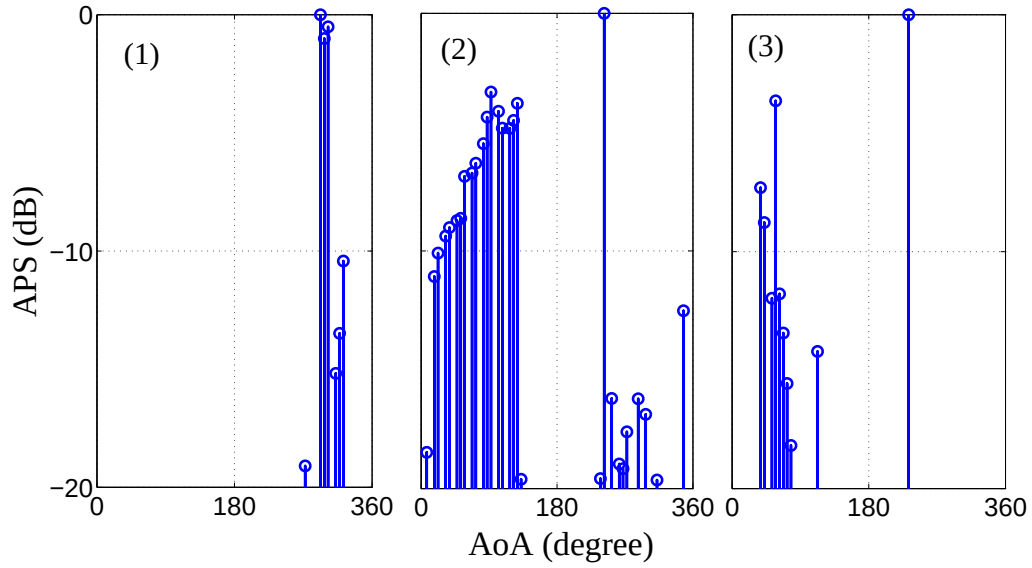


Figure 5-7: Some examples of APS corresponding to Louvre48 (right: $\sigma_\phi = 5.3$ degree, middle: $\sigma_\phi = 71.3$ degree, left: $\sigma_\phi = 115$ degree).

channels having $K \leq 1$. We point out that the number of channel passing the K-S test, presented in Table 5.1, is not so high and thereby reveals the limited number of significant scattered paths seen by most measured points, as already discussed above in Section 5.2.2.

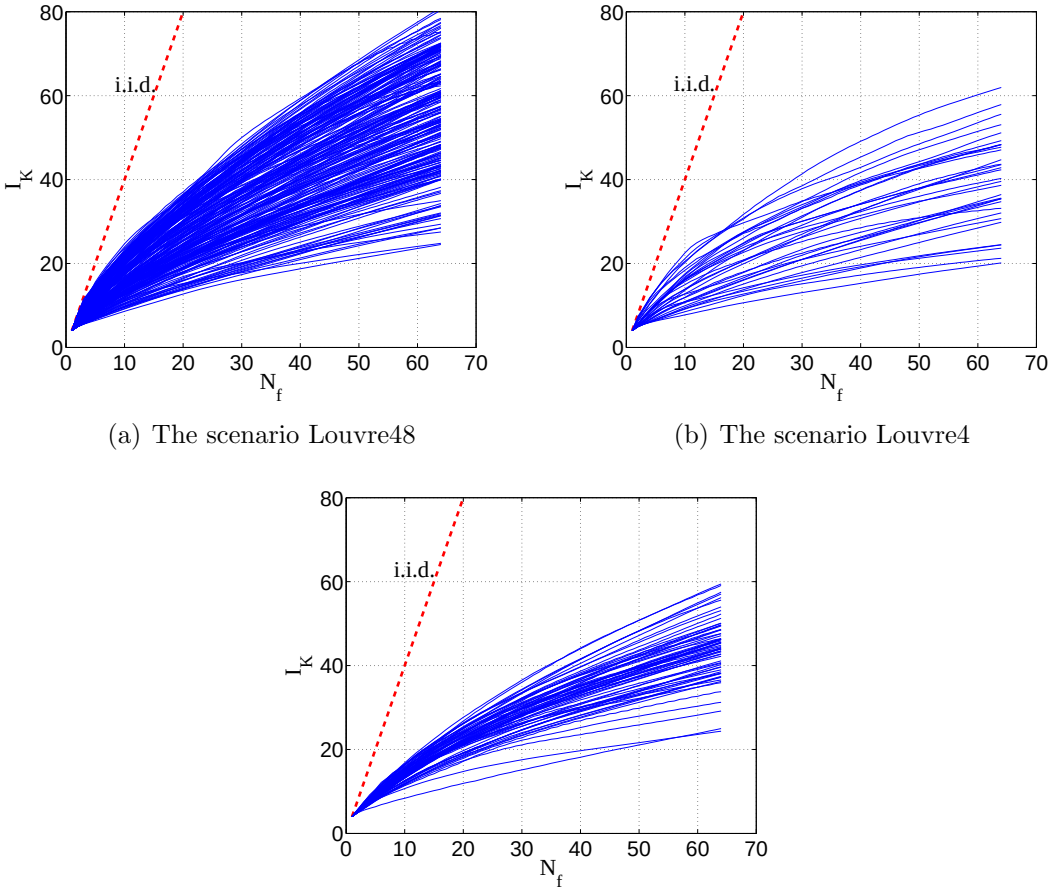
5.3 SKG from frequency variability

In order to assess the benefit brought by the frequency variability of an OFDM channel to wireless security, we consider in particular the effect of the increasing bandwidth with an increasing number of sub-carriers N_f and a fixed separation frequency Δf, as considered in Section 3.3.2 and illustrated in Fig. 3-8 (a). We recall that the BW is directly proportional to the number of investigated sub-carriers, i.e. $BW = N_f \Delta f$. We consider an OFDM channel with $BW = 20$ MHz, typical of the 802.11 standard, and a frequency separation of $\Delta f = 312.5$ KHz, which results in 64 total sub-carriers. We also recall that the frequencies are always equally spaced within the BW. We note that the security is addressed from both legitimate terminals side and the eavesdropper side.

5.3.1 Available key bits

5.3.1.1 BW impact by investigating more sub-carriers

Fig. 5-8 plots I_K as a function of N_f for each position showing a Rayleigh amplitude that passes the K-S test. In accordance with results presented in Section 3.3.2, it is shown that I_K increases with the BW (or N_f) in a sub-linear manner, which reveals that the main channel DoFs (i.e. the rays carrying the major energy) are resolved with a BW less than the total BW of 20 MHz. The curves here show non linear behavior with respect to results reported in Fig. 3-9, since the deterministic channels are more complex, so that the ray powers are not distributed over all the delay bins according to the simple exponential power decay profile. Moreover, some PDPs may show clustering effect.



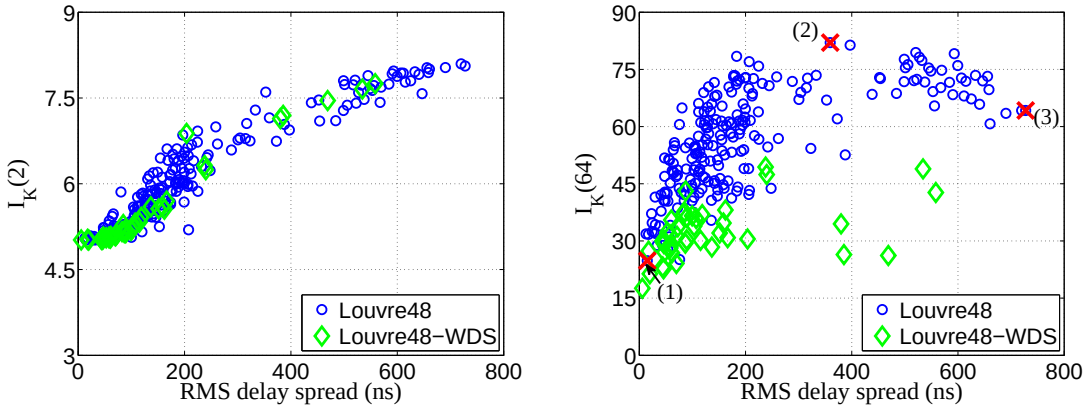
(a) The scenario Louvre48

(b) The scenario Louvre4

(c) The scenario Versailles19

Figure 5-8: Evolution of I_K with respect to the number of frequencies N_f.

We stress that the variance of I_K within each environment and for each value of N_f is relatively large, especially for high values of N_f. This means that the length of the common generated key is strongly related to the channel characteristics of each measured location. Among the relevant parameters to characterize a channel CIR is the delay spread. Hence, we plot in Fig. 5-9 the variation of I_K with σ_τ, for both two and 64 sub-carriers. Obviously, I_K increases as the RMS delay spread increases, revealing that more random information can be shared between Alice and Bob, especially when the used BW is large. These results are consistent with those presented in Fig. 3-12, except for the decreasing behavior in I_K beyond $\sigma_\tau = 400$ ns and for $N_f = 64$. Such a behavior results from the clustering aspect seen in the PDP, where e.g. two far clusters yield an increase in the apparent coherence bandwidth while the channel DoFs or the number of independent paths is small. For a clear illustrative scheme, we can relate the PDPs plotted in Fig. 5-5, according to their labels, to their corresponding (σ_τ,I_K). In particular, if we compare the PDPs labeled (2) and (3), although (2) is characterized by a smaller delay spread, it is richer in MPCs, yielding more randomness to be exploited in SKG.



(a) $N_f = 2$ (b) $N_f = 64$

Figure 5-9: I_K as a function of the RMS delay spread (σ_τ) for the scenario Louvre48.

5.3.1.2 Impact of the radio environment

Also from Fig. 5-8, we can study the impact of the physical environment on the security. First, we notice from Table 5.1 that the percentage of Rayleigh channels (either passing or not the K-S test) is very low for these outdoor environments. This results in a smaller chance to be secure in such environments. Indeed, as seen in

Section 3.1.1, Rician channels are less favorable to SKG than Rayleigh ones, since 1) some efficient methods are required to predict and remove the nonfading part of the radio channel and 2) the remaining power of the random channel part may be very small. Then, if we compare the behavior of I_K for the different scenarios and especially for $N_f = 64$, we notice that the highest values of I_K are reached in Louvre48, where the channels seem to be the richest in DoFs, or equivalently, significant MPCs. This is consistent with the discussion made above in regards of the σ_τ statistics, presented in Fig. 5-4.

5.3.1.3 Impact of the diffuse scattering

In order to highlight the crucial role played by the diffuse scattered paths on the robustness of generated keys, we compute I_K for scenarios where the channels are predicted with and without implementation of diffuse scattering. An explicit comparison is shown in Fig. 5-9 where I_K is evaluated with respect to σ_τ for both scenarios Louvre48 (with implementation of diffuse scattering) and Louvre48-WDS (without implementation of diffuse scattering), while Fig. 5-10 plots I_K versus N_f for channels excluding from diffuse scattering contributions. First of all, we stress that the ratio of Rayleigh channels significantly decreases when any contribution from scattered paths is considered, yielding less opportunities to achieve secure communications. Secondly, we obviously see that smaller values of I_K are reached in the scenario Louvre48-WDS. However, we still have some values of I_K that are comparable with scenarios where diffuse scattering is implemented, and this relies on specific surrounding environment where an important number of specular reflections and edge diffraction exist. Briefly speaking, these results show the relevance of diffuse scattered paths to sustain wireless security in a PhySec context.

5.3.1.4 BW impact by investigating a fixed number of sub-carriers

Fig. 5-11 compares the statistics of I_K when 64 sub-carriers are uniformly distributed within a BW of either 20, 40, 80 or even 160 MHz. For a fixed number of sub-carriers, the number of resolved delay paths is the same, whatever the BW. As the BW increases, the effect of overlapped paths reduces, yielding more information to be exploited in the shared key between Alice and Bob. Hence, the key length increases with the BW, provided sufficient channel DoFs are available. More clearly, while the lower tails of the distributions correspond to lower delay spreads and consequently to poor channels, the higher tails correspond to larger delay spreads and consequently

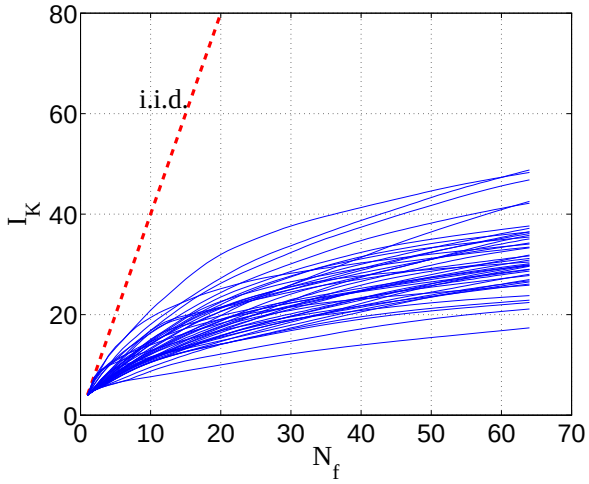


Figure 5-10: I_K as a function of N_f for the scenario Louvre48-WDS.

to richer channels in MPCs, as deduced from Fig. 5-9. As a consequence, I_K takes benefit from the increase in BW for higher tails rather than in lower tails. Moreover, for the same reasons, it seems that I_K at $BW = 80$ MHz and $BW = 160$ MHz are relatively close to each other revealing that most paths are almost perfectly resolved at a BW less than 160 MHz.

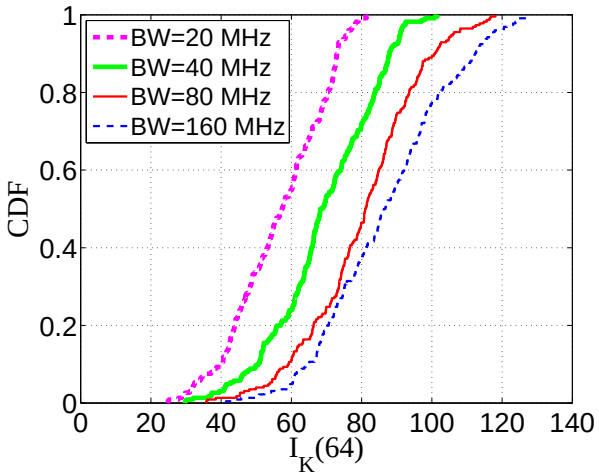


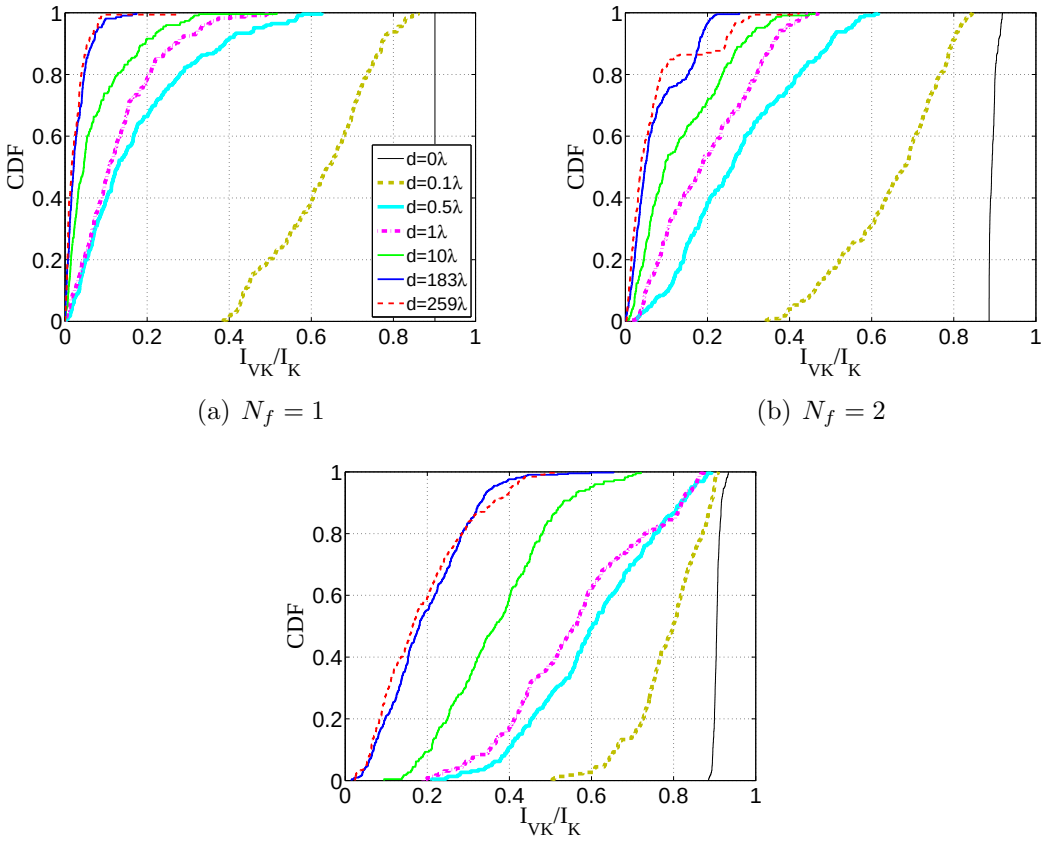
Figure 5-11: The impact of BW on I_K for $N_f = 64$ and for the scenario Louvre48.

5.3.2 Evaluation of the degree of secrecy

After discussing the behavior of I_K with respect to the channel features, it is relevant to assess the amount of information revealed to Eve according to her situation. To this end, we consider a set of Bob-Eve separation distances d, ranging from 0 up to 259λ. Since the RT computes the channel characteristics over a grid of 5 m[3] as a step, the channels at $d = 183\lambda$ and $d = 259\lambda$ are computed based on RT output while, however, it is not the case for $d \leq 10\lambda$. For $d \leq 10\lambda$, Bob and Eve are supposed to be located in the same local area, so they share the same MPCs. Hence, the channels at these distances may be computed by adopting the assumption of plane waves. Furthermore, we assume that Bob and Eve are moving, in the same manner, in the 5λx5λ centered square grid in order to build a statistical distribution over which the vulnerable key rates are computed. Moreover, only the channel measured by Eve when Alice is the transmitter is supposed to contain correlated information with the propagation channel between Alice and Bob. We note that the vulnerability is discussed by investigating merely the scenario Louvre48.

Fig. 5-12 depicts the distributions of relative vulnerable key bits I_{VK}/I_K as a function of Bob-Eve separation distance d and also with different values of N_f. It is straightforward that I_{VK}/I_K decreases as the distance increases owing to the multipath interference in small-scale distances and to the change in the structure of multipath components at larger distances, as already discussed for the disc of scatterers model in Chapter 4. Fig. 5-13 shows that the angular spread is critical for the vulnerability when a single frequency is investigated, especially for $d = 0.1\lambda$ where the variance of the relative vulnerable key bits is large. Hence, more secure key bits may be obtained in scenarios where a large angular spread is provided. On the other hand, by increasing the bandwidth, or equivalently N_f, I_{VK}/I_K increases. Hence, an increase in the channel DoFs brings an advantage for the legitimate terminals in terms of I_K, it also brings an advantage for Eve who is potentially able to capture higher portion of key bits. Similar results are shown in [58, 72], when exploiting the richness available in MIMO channels. We stress that for WB channels (i.e. large values of N_f), the vulnerability does not rely merely on the angular spread since other factors such as the delay spread are also critical in this case.

[3]A distance of 5 m corresponds to 183λ.



(a) $N_f = 1$

(b) $N_f = 2$

(c) $N_f = 32$

Figure 5-12: CDF of I_{VK}/I_K depending on N_f and on Bob/Eve distance (Louvre48).

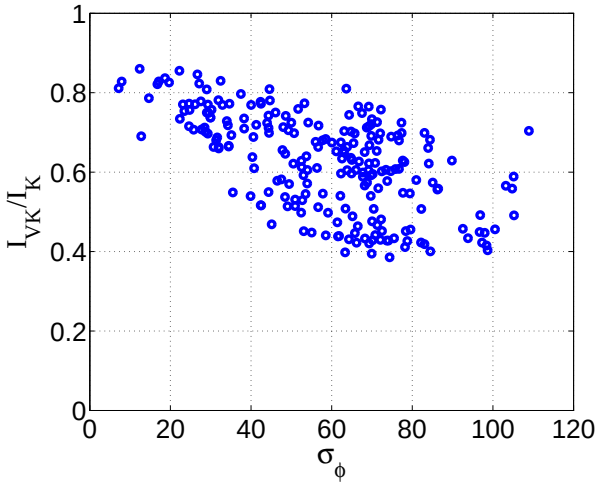


Figure 5-13: I_{VK}/I_K as a function of the angular spread, for $N_f = 1$ and $d = 0.1\lambda$.

5.4 SKG from space variability

We aim in this section to briefly assess the robustness of the extracted key, when exploiting the space variability of the channel. For that reason, we consider a 8x8 square array antennas with $\lambda/2$ inter-element spacing for Bob, while keeping Alice with a single antenna. In this case, the channel vector is constituted by these 64 space-varying parallel channels. The frequency domain, with $BW = 500$ MHz and $N_f = 2500$, is used to form the statistics over which I_K is computed. Again, we just consider the scenario Louvre48.

5.4.1 Available key bits

Fig. 5-14 shows I_K as a function of the angular spread σ_ϕ. Obviously, I_K increases with σ_ϕ until a certain value beyond which I_K decreases. Indeed, the same behavior is shown in Fig. 5-9 where the frequency variability is considered. This decrease in I_K can be ascribed to the clustering effect presented in the angular power spectrum, in relation to the APS plots in Fig. 5-7.

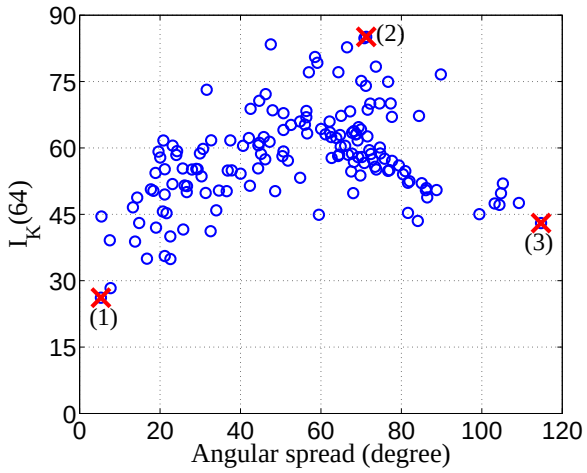


Figure 5-14: Space variability: I_K as a function of the angular spread.

The comparison between the key extraction from 64 parallel sub-channels provided in either the space or the frequency domain is displayed in Fig. 5-15. It turns out that the lower tails of the CDF correspond to small values of either the angular spread or the RMS delay spread, when dealing with space domain or frequency domain SKG, respectively, while the higher tails correspond to larger spreads. In poor channel conditions, corresponding either to small values of σ_τ or of σ_ϕ, the space

variability appears to contain more random information. Nevertheless, in general, I_K takes almost the same values in the two domains, which indicates that these sources of variability statistically offer the same amount of random information in the investigated environment.

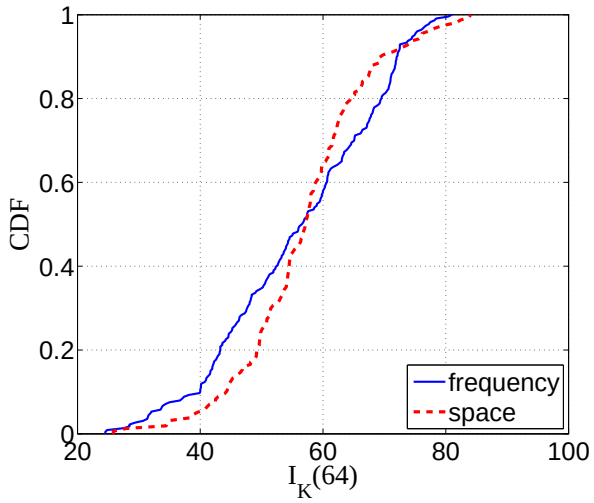


Figure 5-15: Space vs. frequency variability.

5.5 Conclusion

RT simulations constitute an interesting tool in order to assess SKG performance, especially if we want to be "site-specific". The difficulty, which is uncommon for most users of RT tools, is that high performance SKG comes with the full exploitation of the DoF, which must be properly rendered in the simulation framework. Since much of the wireless power is carried by diffuse scattered waves, this puts strong requirements on RT tools in including diffuse scattering. The problem is two-fold: (i) diffuse scattering must be correctly included from the point of view of the laws of physics ; (ii) the extra amount of computations can be expected to be large, based on the dense character of the multipaths structure.

Given these cautionary comments, the work carried out in the present chapter has concentrated on a few scenarios for which a data base of RT simulations including DS was available. A systematic evaluation of the secret key rate vs. the number of frequency sub-carriers was conducted with parameters suited to the IEEE 802.11 standard (WIFI), qualitatively confirming the results of Section 3.3.2 in terms of a sub-linear dependence of the number of available key bits vs. N_f, well below the i.i.d.

case. I_K tends to increase with the RMS-DS, up to a saturation. I_K also tends to increase with the BW, although in a sub-linear dependence.

The "spatial memory" of the channel, surmised from the results of Chapter 4, has been observed on RT results, most likely because the considered scenarios cover rather wide areas with slowly varying multipaths, resulting in a significant vulnerability (from Eve's point of view) of Alice-Bob keys. This implies the necessity to devise privacy amplification schemes that remove the common information between Eve and Alice-Bob for moderate distances.

Finally, for the parameters and considered scenarios, the spatial and frequency based SKG performance appear to be equivalent, as observed from the similarity between the statistical distributions of the secret key rate. This issue will be further investigated in the measurement results of Chapter 6.

Chapter 6

Security performance in measured channels

In this chapter, we intend to test the performance of SKG schemes, especially when exploiting the channel degrees of freedom (DoF), in indoor empirical channels, at both 2.4625 GHz and 5.4 GHz bands (i.e. the typical WIFI bands). We first start by describing the measurement campaign in Section 6.1. Then, in Section 6.2 and Section 6.3, we test the security performance relatively to the CQA algorithm as well as to the electromagnetic propagation features, in terms of key agreement between the Alice/Bob/Eve trio and in terms of key randomness. Subsequently, we exploit these measurements in Section 6.4, towards achievable key rates according to mutual information computations. Finally, the results are compared and discussed in the conclusion.

6.1 Measuring systems and scenarios

Measurements have been performed in the premises of Telecom ParisTech (TPT), which is a century old engineering education building with a highly heterogeneous internal structuring due to many refurbishing events over the years. A 4-port vector network analyzer (VNA: Agilent ENA E5071C) has been used to record channel coefficients over 4 GHz of bandwidth (2-6 GHz) with 2.5 MHz as frequency step. This step, which translates into a maximum channel response delay of 400 ns, is enough to avoid aliasing, given the instrument noise floor and the typical delay spread of multipaths in the concerned environments. Table 6.1 presents the VNA setup

parameters. One port of the VNA has been devoted to Alice, as transmitter, whereas the three remaining ports have been devoted to Bob/Eve, as receivers. Each port was equipped with an identical UWB bicone antenna with 2 dBi gain, specifically designed for the frequency stability of the radiation pattern [122]. The VNA has been calibrated with a "full 4 ports" method including the (highly phase stable) cables, resulting as output at each frequency in the full 4x4 matrix of the complex channel coefficients including all antennas.

Furthermore, the measurements were carried out in a time period where few people were present (mainly during week-ends, evenings etc.). Given the conditions of the measurements (in particular the time duration of a frequency sweep), there is no doubt that the time stability was largely sufficient for the needs of the experiments.

Table 6.1: VNA setup parameters.

Parameter	**value**
Start frequency	2 GHz
Stop frequency	6 GHz
Frequency points	1601
IF Bandwidth	5 KHz
Transmitted power	10 dBm
Dynamic range	96 dB
Typical noise floor	-86 dBm

The measurements have been carried out in classrooms and in an auditorium, in order to have indoor scenarios of sufficiently different characteristics, including identical or different heights for the terminals; LOS or NLOS propagation condition and also different room sizes. Fig. 6-1 shows the floor plans of both classrooms and auditorium where the environment is mainly constituted of concrete, plywood and partition walls. In the classroom scenario, the terminals have been placed at the same height (1.3 m from the ground) whereas in the auditorium they have been placed at different heights as seen in Fig. 6-2 and Fig. 6-3. The location of Alice was fixed for each of the two environments whereas the remaining three antennas have been moved across the area in a set of irregular locations, mostly within the room but also in the adjacent corridor or in an adjacent room. More clearly, the antennas representing Bob had 51 different positions in the classrooms scenario and 42 positions in the auditorium scenario, where only 25 total positions were in NLOS condition with respect to Alice. The NLOS scenario encompasses either room-to-room or room-to-corridor propagation conditions, as shown in Fig. 6-1.

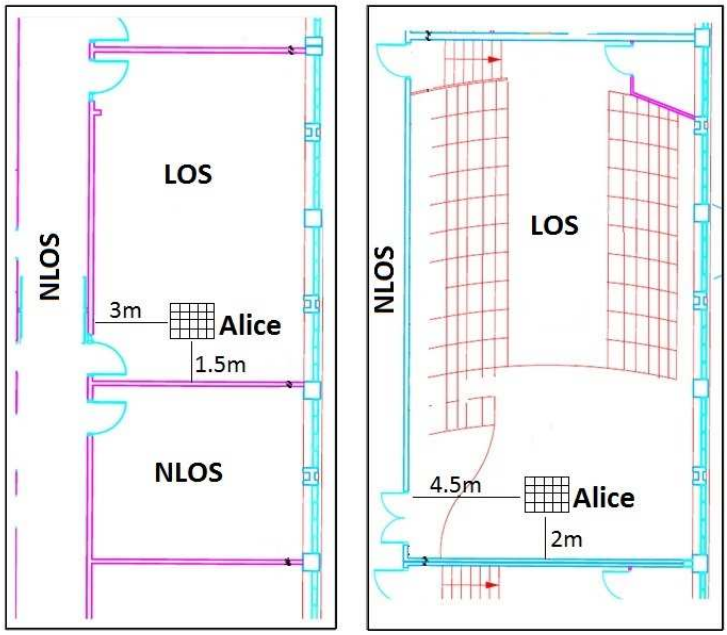


Figure 6-1: TPT measurement floor plans: (left) classrooms and (right) auditorium.

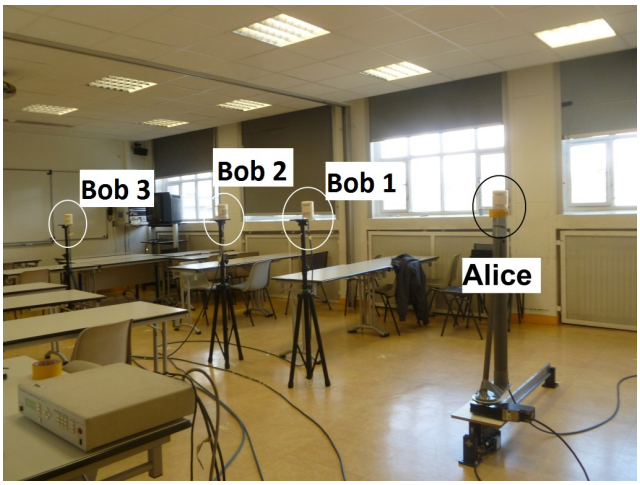


Figure 6-2: A sketch of measurement run in TPT classrooms scenario.

In each measurement run, the three receivers representing Bob are steady while the transmitter representing Alice is spatially scanned over a square grid of 11x11 points (30 cm side and 3 cm step) confined to a small area, so as to capture fast fading. More clearly, since the grid step is about half a wavelength at 5 GHz, we can expect to achieve close to statistically independent channel coefficients owing to spatial fading. The total 4 GHz bandwidth (BW) enables us to investigate in this paper the security performance of wide band (WB) channels centered at either 2.4625 GHz or 5.4 GHz (typical of the WIFI band) with e.g. a bandwidth either of

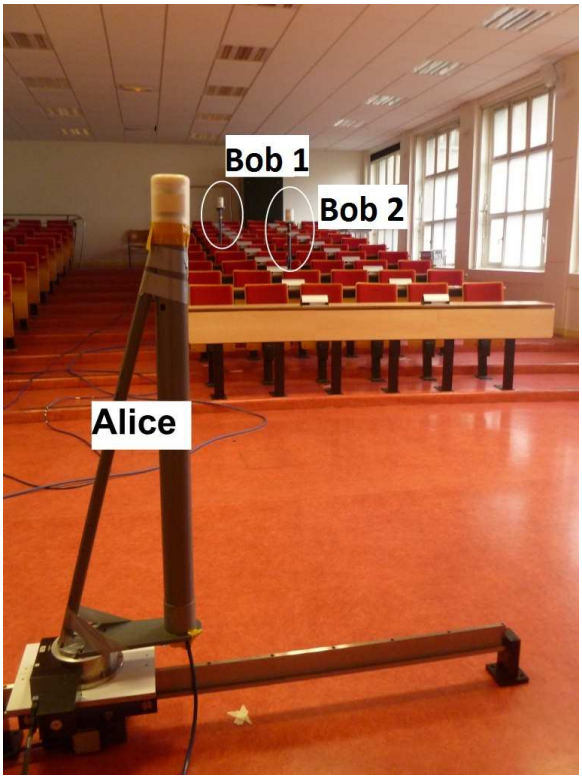


Figure 6-3: A sketch of measurement run in TPT auditorium scenario.

20, 40, 80 or even 160 MHz, according to the series of WIFI (IEEE 802.11) standards (Appendix B).

6.2 Evaluation of errors between Alice, Bob and Eve keys

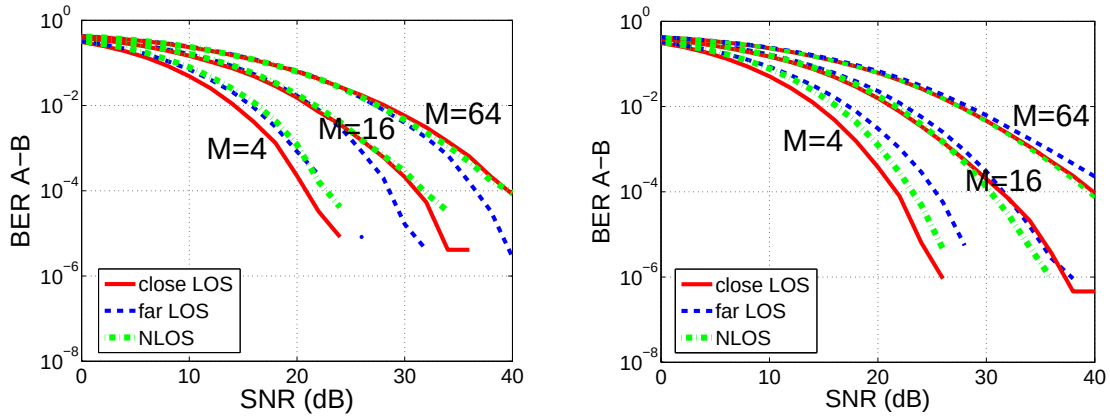
This section is dedicated to assess the quality of the key from the view point of the key similarities or differences between the Alice/Bob/Eve trio, for both NB and WB channels centered at 5.4 GHz band. The keys are derived from the noisy complex channel coefficients, through the CQA algorithm. Eve also employs the same algorithm and exploits all the public information exchanged between Alice and Bob, i.e. the quantization map QM.

The VNA measurements were carried out in excellent dynamic range conditions, owing to the VNA measurement principle. For that reason, the correlation coefficient between the Alice to Bob and Bob to Alice channels is almost always equal to 1, indicating a high degree of reciprocity (very high signal to noise ratio (SNR)). Therefore,

we added independent artificial noise in the form of a zero-mean circular complex Gaussian random variable to study the effect of noise on security performance. The BER is computed over keys extracted from both the spatial domain (i.e. the 121 spatial positions of Alice over the grid) and the frequency domain (i.e. N_f sub-carriers). The BER is then averaged over 500 channel realizations where only the noise is the random variable. All the terminals are supposed with an equal SNR.

6.2.1 Dependence on the mapping

Fig. 6-4(a) and Fig. 6-4(b) show the variation of the BER between the keys generated by Alice (A) and Bob (B) with respect to the SNR, the number of QRs M and for LOS and NLOS propagation conditions. Instead of averaging the BER over all Bob positions representing either LOS or NLOS conditions with respect to Alice, we just choose certain positions in order to reduce the simulation time. In the LOS case, two scenarios for Bob are considered. The first one, named close LOS, corresponds to a separation distance of 1.5 m with respect to Alice, while the second, named far LOS, corresponds to a separation distance of 6.5 m. In the NLOS scenario, Bob and Alice are located in two adjacent rooms belonging to the classroom environment. Fig. 6-4(a) and Fig. 6-4(b) consider, respectively, NB channels ($N_f = 1$) and WB channels ($N_f = 9$ sub-carriers separated with $\Delta f = 10$ MHz). It is clear that the BER is improved by increasing the SNR and by decreasing the number of QRs, leading to less sensitivity on noise. Moreover, the BER is almost similar for NB and WB channels since, in the latter, the sub-carriers have almost the same average SNR.



(a) $N_f = 1$ (b) $N_f = 9$

Figure 6-4: Alice-Bob BER vs. SNR and M.

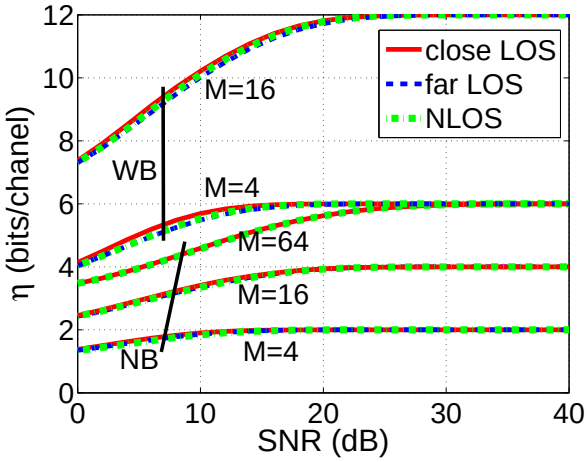


Figure 6-5: Secret key rate η for both NB and WB ($N_f = 3$) channels.

Fig. 6-5 shows the values of the efficiency η (i.e. effective key rate) for both NB and WB channels where in the latter $N_f = 3$. These values are always improved with the SNR, while an increase in M intuitively yields an increase in the efficiency. Furthermore, a WB channel is preferred since it provides the largest value of η. However, this is relevant merely in the case of decorrelated sub-carriers, from which result decorrelated bits, which is a feature not expressed by η. Therefore, in practice, Alice and Bob should be aware of the channel coherence bandwidth in order to appropriately choose the sub-carriers used as the shared source of randomness.

6.2.2 Alice-Bob disagreement vs. a simple channel model

In Fig. 6-6, we show the CDF of the bit disagreement between Alice and Bob, aggregating all locations while distinguishing between LOS and NLOS cases. This is done for different values of the SNR but only for $M = 4$ and for NB channels. We clearly see that LOS propagation is statistically more favorable, from the point of view of Alice and Bob's agreement between key bits, especially when there is a dominant component and at high SNR. This can be understood from the fact that, for a given SNR, the stochastic character of the fading channel coefficients is higher in NLOS than in LOS. This results in a larger number of channel coefficients with low instantaneous SNR in the former case.

If both Alice and Bob are in the same room, a LOS path may dominate the multipath channel between them. Since the channel variation is due to Alice movements over the square grid, the LOS phase is not constant. Given the regular movements

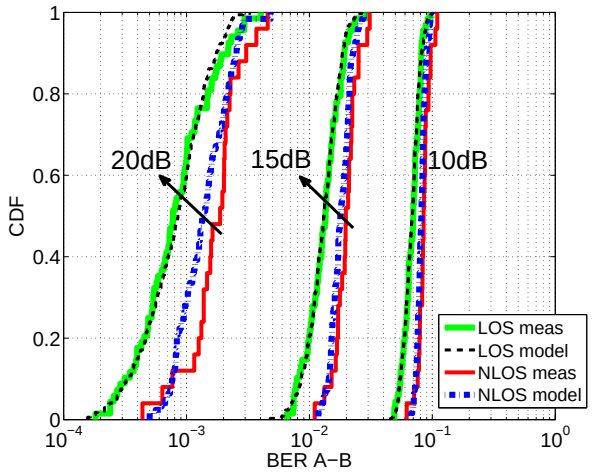


Figure 6-6: Comparison of the BER A-B computed over both the measurements and the model, for SNR=10, 15 and 20 dB.

of Alice and the Bob positions set, we are able to reproduce the channel by modeling the LOS component as follows:

$$\mathrm{h_{LOS}} = \sqrt{G_t G_r (\frac{\lambda}{4\pi d})^2} \exp j\frac{2\pi}{\lambda} d \tag{6.1}$$

where the amplitude is computed according to the Friis equation, given the antenna gains at both the transmit (G_t) and the received (G_r) sides. d is the distance between Bob and a position of Alice on the grid. The fading multipath components are then modeled as zero-mean circular Gaussian distribution (resulting in a Rayleigh multipath channel) of power equal to the total measured NLOS power and added to the LOS component. This procedure ensured that the K factor of the Rician statistics observed in the presence of a significant LOS component was correctly reproduced. Then again, the fading components model is used to model the channel when NLOS condition is provided.

Fig. 6-6 compares the behavior of the SKG from both empirical channels and the simple channel model, for NB channels. The results show a good fit between the BER distributions of the measured and modeled channels, especially in LOS. In the NLOS case, the discrepancies affect the tails of the distributions, which can be ascribed to the fact that the model (Rayleigh distribution for the multipaths contribution) imperfectly describes the limited number of significant paths in the concerned indoor environment. Similar results can be found for the WB case since the sub-carriers have the same behavior regarding the BER.

6.2.3 Security performance with respect to Eve

Eve is assumed to be a passive attacker. She is able to measure the radio channel between herself and Alice h_{ea} as well as with Bob h_{eb}. Each terminal is considered at each time as either Bob or Eve. Furthermore, Eve is assumed to also apply the CQA algorithm on her channel observation in the hope to derive a secret key close to Alice-Bob's by exploiting her own channel estimates h_{ea} and h_{eb}. The BER between the keys constructed by Alice and Bob and that derived by Eve is then evaluated. The results are shown in Fig. 6-7 for $M = 4$, $SNR = 15$ dB and NB channels.

It can be seen that, based on this simple comparaison between the keys, making use of the h_{eb} channel by Eve does not give her a strong advantage, since the BER is around 0.5. This is due to the fact that this channel is almost constant and does not reveal significant information about the key. On the other hand, the channel h_{ea} is more critical since it is a time-varying channel, depending on Alice movements. Furthermore, a LOS condition between Bob and Eve (LOS B-E) may reveal more information about the secret key than the NLOS case (NLOS B-E). Indeed, in indoor environments, LOS B-E means that Bob and Eve are in the same room and maybe very close, and they are surrounded by almost the same interacting objects, which results in a high probability that they experience correlated channels. Nevertheless, by increasing M, the leakage information to Eve decreases as shown in Fig. 6-8, where only the BER from h_{ea} is considered and for LOS B-E.

Of course, this direct comparison is not a proof of the inability for Eve to access Alice-Bob's common key. It just shows that Eve's job will not be simple and she must call for more sophisticated means and make use of any extra information she can get to guess the legitimate users' key. The number of vulnerable key bits, deduced from information theory, will be presented in Section 6.4.3 below.

6.3 Channel degrees of freedom for SKG

Since the efficiency η does not consider the randomness character of the key, we assess in this section this key quality through the statistical NIST tests described in Section 3.1.4. We also investigate the impact of the space, time, frequency and joint space-frequency degrees of freedom on the SKG performance, given that these DoFs may provide more available shared bits. The relevance of such an investigation is also discussed in Section 3.3.

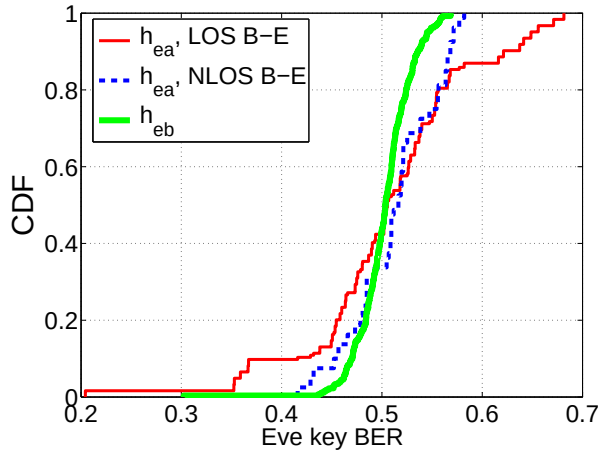


Figure 6-7: BER of Eve key bits for $M = 4$ and for NB channels.

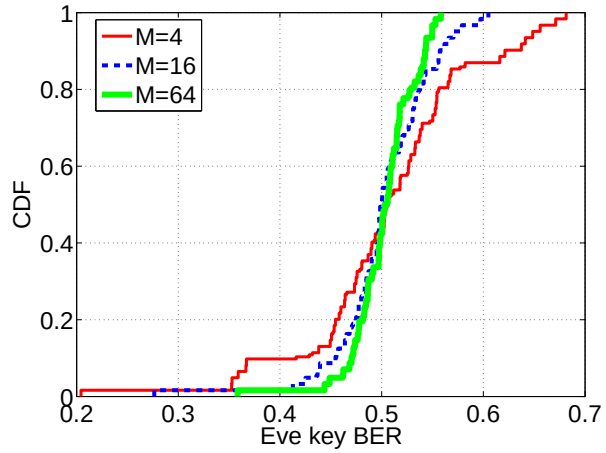


Figure 6-8: BER of Eve key bits from h_{ea}, for LOS B-E and for NB channels.

6.3.1 Space vs. time variability

In TPT measurements, the spatial variability is provided by the movement of Alice over the 11x11 square grid as explained in Section 6.1, which is equivalent to the first type of time variability. These 121 antenna positions allow testing the SKG performance provided by spatial/time degrees of freedom, where Alice's antenna can take random positions over the grid. In this way, we obtain N_S complex channel coefficients, exploited to construct a $(N = N_S \log_2 M)$-bits key at a given frequency. Hence, we randomly choose to construct 60 sets of random Alice positions for each Bob's position and for each available frequency in the 20 MHz-bandwidth. A statistical distribution can then be computed over Bob's positions, over the frequencies in the 20 MHz-bandwidth and over the 60 random sets of Alice positions.

6.3.2 Frequency variability

In order to investigate SKG performance in frequency variant channels, the data has been processed almost consistently with the 802.11a/g/n/ac standard, i.e. in order to obtain complex channel coefficients at the required number of sub-carriers for each BW. For that purpose, TPT campaign measurements were frequency interpolated, which was feasible since the effective maximum delay was well below the inverse of the BW (in other words, the frequency gap between sub-carriers is well below the coherence BW). Moreover, for the same WIFI standard consistency, we discard the channel coefficients at frequencies dedicated either for guardband interval or to transmit pilot bits, and we keep only those at data transmitting frequencies. Table 6.2 shows some frequency channel characteristics for each bandwidth and according in general to the 802.11 standard, where we denote by FFT the fast Fourier transform. Nonetheless, for the sake of simplicity, we do not respect the exact features of each 802.11 standard, which are recalled in Appendix B).

Table 6.2: Frequency channel characteristics for each BW.

BW (MHz)	occupied BW (MHz)	FFT size	Number of sub-carriers	sub-carrier separation (MHz)
20	16.25	64	52	0.3125
40	33.75	128	108	0.3125
80	73.125	256	234	0.3125
160	146.25	512	468	0.3125

Given these parameters, not all the sub-carriers need be used to generate keys of enough bit length, then comes the question: how to choose the sub-carriers? Intuitively, more correlation is likely to occur when the frequency difference between two channel coefficients is reduced. Unless the ratio between the number of available and the number of required sub-carriers is integer, there is no unique and obvious way to choose the sub-carriers used in the SKG process. Hence, Alice chooses randomly a set of N_f frequency sub-carriers, from which ($N = N_f \log_2 M$)-bits key is extracted, and she sends publicly this set to Bob. Although this information is transmitted also to Eve, it is not very relevant since it does not indicate any information about the key bit value. Finally, a set of secret keys is obtained over Bob's positions, over the 121 positions of Alice and over the random sets of sub-carriers (arbitrarily taken to be 10 sets).

6.3.3 Joint space-frequency variability

Intuitively, the smaller the coherence bandwidth, the more efficient will the SKG be able to exploit frequency variability. Unfortunately, the coherence bandwidth changes from an environment to another and is out of control. SKG performance should be achieved also in environments where the coherence bandwidth is small, which is a difficulty when no sufficient spatial variability is provided. As a way of mitigation we here consider the possibility to exploit jointly the space and frequency DoF, so to relax the requirements on each of both individually. A potential use case is that of MIMO systems (such as IEEE 802.11n/ac), providing spatial variability, together with OFDM technology providing frequency variability.

Based on the features of the TPT campaign, spatial variability is provided by considering either each two consecutive Alice positions on each row of the grid or each four-square consecutive Alice positions on each two consecutive rows of the grid, as an array antenna resulting respectively in either 110 sets of 2-array antennas or in 100 sets of 4-square array antennas. More clearly, the following vector

$$V = [X_1^1, \cdots, X_1^{N_{ant}}, \cdots, X_i^1, \cdots, X_i^{N_{ant}}, \cdots, X_{N_f}^1, \cdots, X_{N_f}^{N_{ant}}]$$

is used to construct a single key of length $N = N_{ant} N_f \log_2(M)$ where N_{ant} is the number of array antennas. $X_i^1, \cdots, X_i^{N_{ant}}$ together form an N_{ant}-array antennas at the sith chosen frequency. Finally a set of keys is obtained over Bob's positions, over the sets of N_{ant}-array antennas and over the 10 sets of randomly chosen sub-carriers.

6.3.4 Results

In the following, we use a fixed key length ($N = 128$) in the key randomness quality evaluation, with the exception of the pure spatial variability case where a comparison between different key lengths is carried out. For each channel variability type, a statistical distribution over the extracted keys is formed in order to compute a mean pass rate using the NIST tests. Table 6.3 shows the number of tested keys for each type of channel variability.

Whatever the source of channel variability used to generate the key, our results show that all the keys pass the mono-bit frequency test. This is due to the statistically equal quantization intervals on each I and Q, used to transform channel coefficients into discrete sequences of bits through CQA. Consequently, all the strings (of length

Table 6.3: Size of key set vs. the variability type.

Variability		Number of keys		
		LOS	NLOS	total
Space		36720	13500	50220
Frequency		82280	30250	112530
Joint space-frequency	$N_{ant} = 2$	74800	27500	102300
	$N_{ant} = 4$	68000	25000	93000

$\log_2 \sqrt{M}$) have the same probability to occur and equivalently, the probability to have either bit 0 or bit 1 is 1/2. Therefore, we exclude the mono-bit frequency test when we compute the mean pass rate.

For $N = 128$ and according to Table 3.1, the approximate entropy (ApEnt) test can be applied with a bit string of length $m = 1$, while the serial test with both $m = 3$ and $m = 4$. Accordingly, a 128-key bits passes the ApEnt test when overlapping strings of length 2 bits are equiprobable (i.e. uniform). For the serial test, the result depends on the uniformity of overlapping strings of length 4, 3 and 2 bits. When quantizing the CSIs into M quantization regions according to CQA, 1) the non-overlapping strings of length $\log_2 \sqrt{M}$ bits, derived from the I/Q domains, are equiprobable; 2) the uniformity of the strings of length $\log_2 M$ bits, which result from the concatenation of the $\log_2 \sqrt{M}$ bit strings quantized from the I and Q domains, depends on the correlation between the I and Q domains; 3) the uniformity of the non-overlapping strings relies on the channel samples correlation. Consequently, the percentage of sequences passing the ApEnt test is very high for $M = 16$ whatever the channel variability type, while that of the serial test is smaller especially for frequency variability with $BW = 20$ MHz (starting from a percentage of 0.3533). For $M = 4$, the success to the ApEnt test depends mainly on the correlation between I and Q of the channel coefficients. It also depends on the correlation of the used subsequent channel coefficients. Therefore, the worst case is considered for LOS case exploiting frequency variability with $BW = 40$ MHz, and thus with a percentage of 0.7273. Moreover, small mean pass rates for $M = 4$ stem from the approximately complete failure of serial tests.

6.3.4.1 Spatial variability

Key length effect: Fig. 6-9 represents the mean pass rate of key sequences passing the chosen selection of NIST tests, for both $N = 128$ and $N = 242$. The spatial channel variability is used here to construct keys of N bits with $M = 4$. $N_S = 64$ and

$N_S = 121$ channel samples are needed to respectively construct a 128-bits key and 242-bits key. Whatever the used frequency, it is shown that shorter keys better profit from the channel randomness. While maintaining the same M, we need more channel samples in order to construct a longer secret key and consequently the probability to have more correlated samples increases, yielding bits with more correlations.

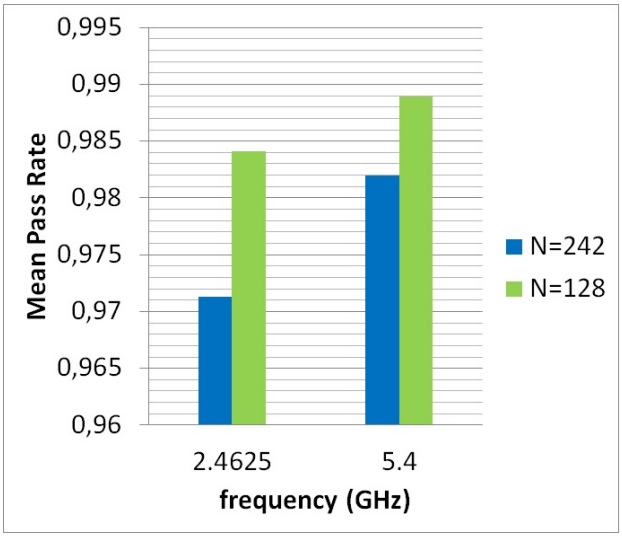


Figure 6-9: Mean pass rate exploiting spatial variability for both $N = 128$ and $N = 242$.

Carrier frequency effect: Fig. 6-10 shows the mean pass rate for $N = 128$, for both 5.4 GHz and 2.4625 GHz bands, and with respect to LOS/NLOS cases. The impact of carrier frequency is not really meaningful in Fig. 6-9 and Fig. 6-10 since the mean pass rates are very high, i.e. nearly 1, in good part owing to the random positions taken by Alice over the grid. Nonetheless this impact may be shown for the worst-case scenario corresponding to consecutive Alice positions over the regular grid, and thereby the 5.4 GHz band offers more random keys than the 2.4625 GHz band. Indeed the distance between two adjacent Alice positions on the grid corresponds almost to $\lambda/2$ at 5.4 GHz and to $\lambda/4$ at 2.4625 GHz, while $\lambda/2$ typically corresponds to the coherence distance over which channels are statistically well decorrelated in omnidirectional scenarios, resulting in extracted bits with a good level of independence.

LOS/NLOS effect: The key randomness is enhanced in NLOS propagation conditions, as shown in Fig. 6-10, due to the lack of a dominant path yielding then more fluctuation of the channel transfer function than in LOS cases.

Briefly speaking, it is noteworthy that in all cases the mean pass rate is very high,

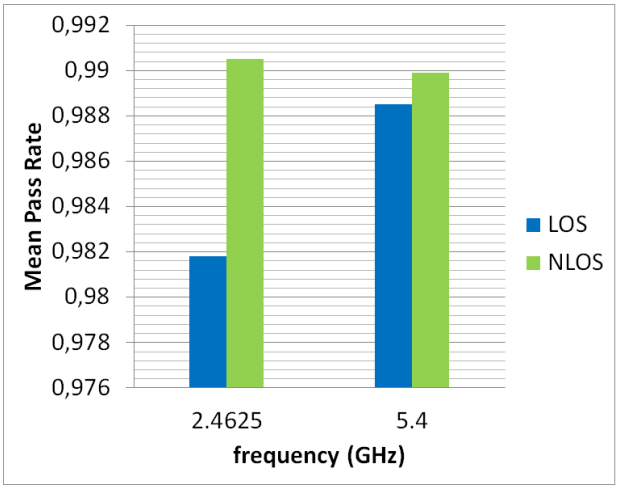


Figure 6-10: Mean pass rate exploiting spatial variability with respect to LOS/NLOS and for $N = 128$.

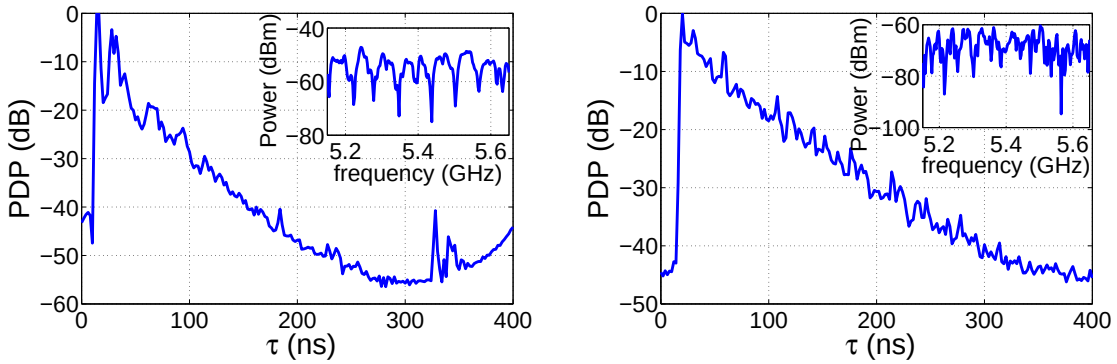
indicating that the spatial degree of freedom is suitable for random key generation. As discussed above, spatial variability can be translated into time variability through a random movement of Alice in space, providing adequate key randomness. As an extra advantage, such a time variant scheme would make it difficult for Eve to track accurately Alice's positions, reducing her ability to gather deterministic information about the channel characteristics and to guess the sequence of bits.

6.3.4.2 Frequency variability

A quantitative measure of the key randomness behavior with respect to the frequency variability domain can be found from the analysis of the root mean square (RMS) delay spread σ_τ and consequently of the coherence bandwidth, which typically varies inversely to the RMS delay spread. For each position of Alice over the square grid, a CIR is computed by taking inverse Fourier transforms of the frequency responses recorded over 500 MHz bandwidth centered on either 2.4625 GHz or 5.4 GHz band and filtered with a Hamming window. The power delay profile (PDP) $P(\tau)$ is then the average of the 121 squared CIRs computed over the spatial grid (see Eq. 2.7 where the time domain is replaced by the space domain). Only multipath components with amplitude within 20 dB from the peak of each PDP are included in the computation of σ_τ, which is given in Eq. 2.5.

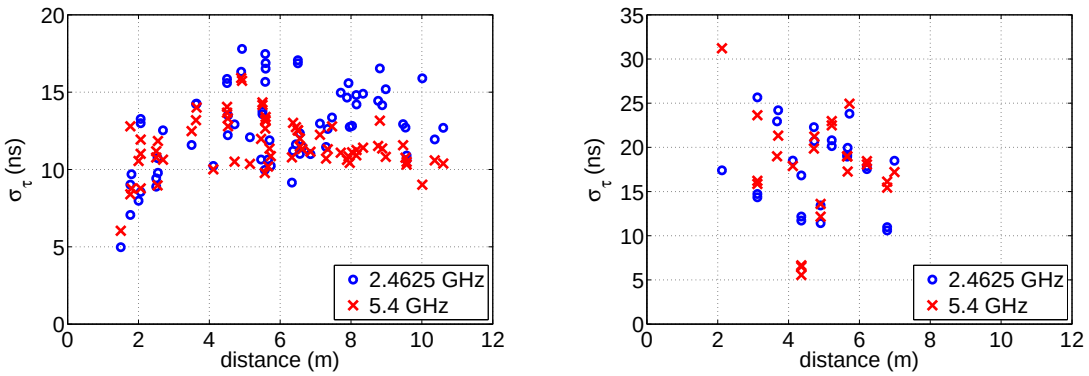
Fig. 6-11 shows two examples of normalized measured PDPs and their corresponding frequency responses for both LOS and NLOS cases. It is clear that the NLOS

PDP is richer in multipath components and thereby exhibits higher delay spreads than the LOS one, having a few dominant peaks at short delays. This observation is validated in Fig. 6-12 where σ_τ is plotted against the distance between Alice and Bob, both for LOS and NLOS cases.



(a) LOS (b) NLOS

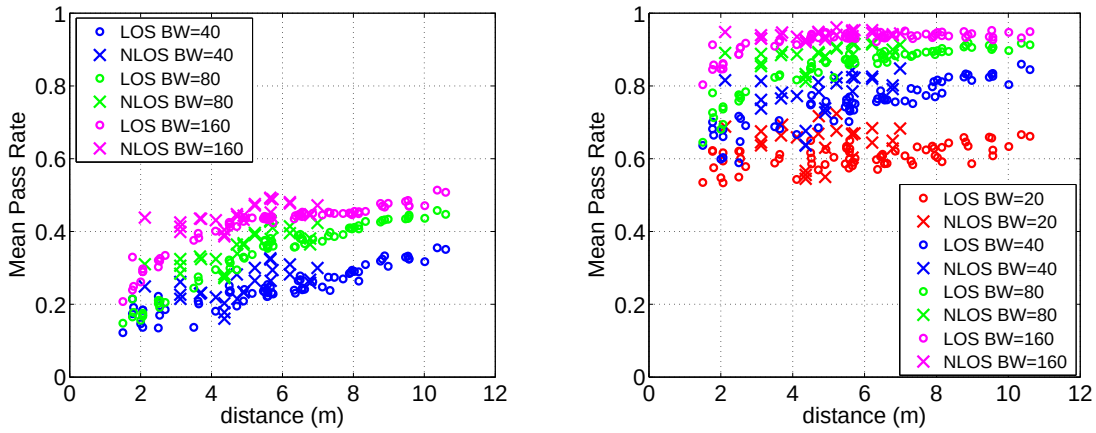
Figure 6-11: Examples of PDPs and channel transfer functions.



(a) LOS (b) NLOS

Figure 6-12: The variation of the RMS delay spread with the distance.

We assess the key randomness exploiting frequency variability by maintaining the same key length as in most analysis above, i.e. $N = 128$ bits. To this end, we fix the number of sub-carriers N_f used for SKG according to M, i.e. $N_f = 64$ for $M = 4$ and $N_f = 32$ for $M = 16$. Fig. 6-13 shows the variation of the mean pass rate as a function of the distance between Alice and Bob at 5.4 GHz band, by differentiating between LOS and NLOS for different bandwidths and both for $M = 4$ and $M = 16$. We stress that unfortunately 128 key bits cannot be extracted by exploiting the frequency variability in $BW = 20$ MHz when $M = 4$. Fig. 6-14 considers the impact



(a) M=4 (b) M=16

Figure 6-13: Mean pass rate as a function of the distance for 5.4 GHz and for frequency variability.

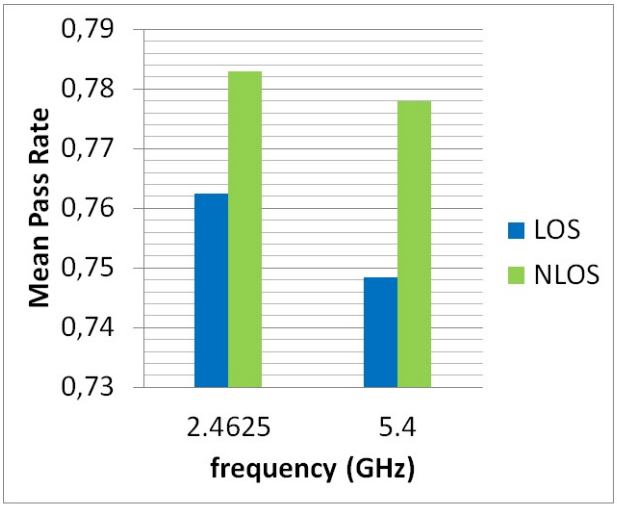
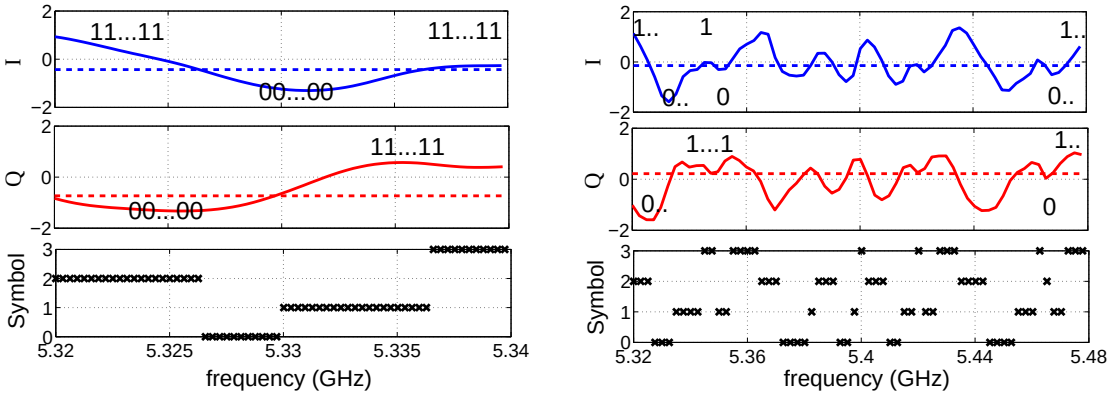


Figure 6-14: Mean pass rate exploiting frequency variability for $BW = 40$ MHz and $M = 16$.

of the carrier frequency on the key randomness behavior for $BW = 40$ MHz and $M = 16$.

Both LOS/NLOS and distance effect: Fig. 6-13 shows that the higher the separation distance between Alice and Bob, the higher the mean pass rate, especially for LOS channels or for small values of M. Moreover, NLOS channels provide statistically more random secure key bits as seen in both Fig. 6-13 and Fig. 6-14. The same behavior is noticed in Fig. 6-12 with respect to the delay spread. Hence, the improvement of the mean pass rate is explained by an increase of σ_τ indicating a

reduction in the coherence bandwidth, which yields less channel correlations for close frequency responses. Furthermore, the advantage of NLOS channels over the LOS ones in providing random keys comes from the multipaths richness of the former: the lack of proper Rayleigh fading reduces the channel variability in the frequency domain and creates insufficient randomness for a satisfactory success to NIST tests. Nonetheless, σ_τ takes relatively small values ranging from 5 ns to 30 ns, due to the open and little cluttered environment of TPT investigated locations. These values are consistent with typical ones for indoor environments, see e.g. ref. [123]. An improvement in mean pass rate is thus expected for richer scattering environments.



(a) very close channel responses

(b) far spaced channel responses

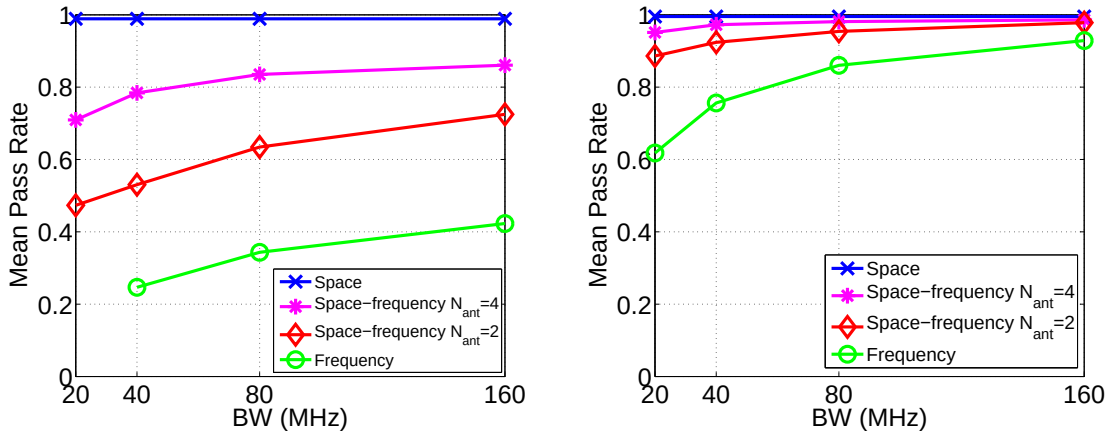
Figure 6-15: Example illustrating SKG from channel responses, for M=4.

Bandwidth effect: Larger available bandwidths yield larger separation of the sub-carriers used for SKG and consequently smaller correlations. This results in improved key randomness, as seen in Fig. 6-13. Fig. 6-15(a) and Fig. 6-15(b) illustrate two examples of key generation from respectively very close and very far spaced channel responses, for $M = 4$. It is clear that more randomness is provided by the case where the channel coefficients are very far spaced, where SKG profits from the whole bandwidth, while the efficient bandwidth is reduced in the other case yielding a key with poor randomness.

Carrier frequency effect: As seen in Fig. 6-14, the carrier frequency affects the key randomness behavior just for LOS channels, where higher mean pass rates are seen for the smallest carrier frequency (i.e. 2.4625 GHz). This is explained by the decrease in the coherence bandwidth, or equivalently by the increase in the RMS delay spread, when the frequency gets lower, as shown in Fig. 6-12(a). Furthermore, as displayed in Fig. 6-12(b), σ_τ does not change with the frequency for NLOS channels. The behavior of σ_τ with the carrier frequency is consistent with results obtained in

[124, 125]. However the difference in mean pass rates is weak, implying that there is no strong preference between the low and high WIFI band from this point of view. Still, the fact that the low band is limited to 20 MHz bandwidth while the high band reaches 160 MHz provides a clear advantage of the latter for SKG, given the above results.

6.3.4.3 Joint space-frequency variability



(a) M=4 (b) M=16

Figure 6-16: Comparison of key randomness in different channel variability at 5.4 GHz.

Fig. 6-16 compares the mean pass rate for the three types of channel variability in the 5.4 GHz band for both $M = 4$ and $M = 16$. The full space variability provides the most robust keys and thereby the most suitable source for SKG. However, such a scheme would either require the terminal mobility over all the scanned positions before generating a key or would need that many antennas in a stable scenario. Therefore we now assess the security provided by joint space-frequency variability, for both $N_{ant} = 2$ and $N_{ant} = 4$, which relaxes such requirements. Indeed, this scheme provides more random keys, especially with $N_{ant} = 4$, than the pure frequency domain variability, which stems from the larger average difference between frequency channels and the resulting reduced correlations between channel coefficients. Simply stated, for fully decorrelated antenna signals, the increase in N_{ant} reduces the bandwidth requirements. This is a very encouraging result for the effectiveness of SKG toward physical layer security. Since many wireless devices (for 3G, 4G, WIFI...) tend to be multi-antenna systems, such a solution will certainly be more and more feasible

in the short term future. We also stress the importance of increasing M, which well improves the key randomness, despite the requirement of a higher SNR.

6.4 Information theoretic bounds on key length

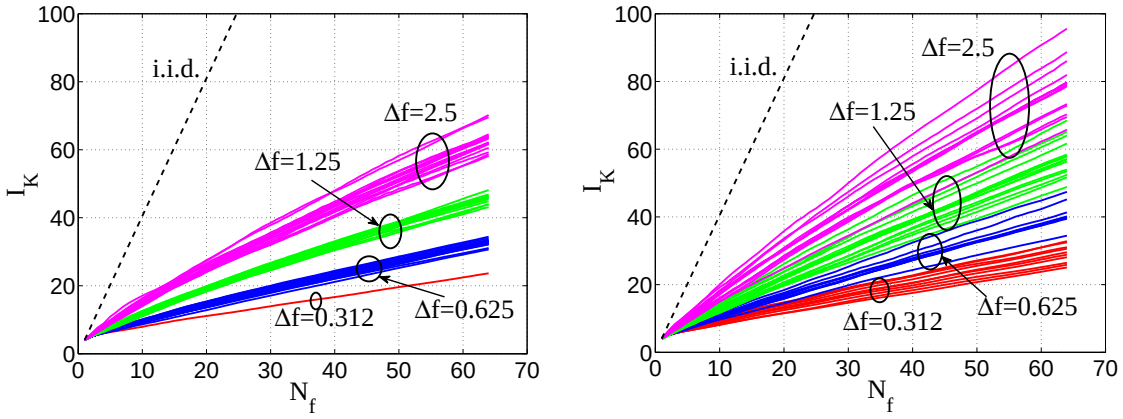
In order to account for both the key agreement and the key randomness, it is crucial to evaluate the mutual information between channels. To that aim, the radio channels are required to be jointly Gaussian. However, in real channels, the central limit theorem may not hold owing to limitation in the number of dispersive paths. Therefore, we test the Gaussianity of the channels with the Kolmogorov-Smirnov (K-S) test and then, we just compute available key rate I_K and vulnerable key rate I_{VK} for the channel coefficients that passes the test.

In this section, the results are obtained for $SNR = 15$ dB and for a central frequency of 5.4 GHz.

6.4.1 SKG from frequency variability

The frequency variability scheme adopted in this section is the same as that represented in Fig. 3-8(a) in Chapter 3. By increasing the number of sub-carriers N_f, the BW increases linearly while the frequency separation Δf is fixed. The 121 positions of Alice over the square grid form the statistics over which mutual information is computed for each position of Bob. Moreover, the channel vector results from the aggregation of N_f frequencies around the central frequency, as already done for outdoor environments. The impact of investigating N_f sub-carriers with different Δf is depicted in Fig. 6-17, by distinguishing between LOS and NLOS conditions.

As already shown in the previous chapters and also here above, investigating more sub-carriers with larger Δf yields an improvement on the SKG process. In contrast with results shown for outdoor deterministic channels, we notice here that the curves are more linear, which could be explained by the fact that the empirical CIRs are very dense in multipaths, and subsequently, we need a BW much higher than 160 MHz in order to exploit the most random information contained in the CIR. Furthermore, I_K is greater in NLOS cases than in LOS ones, which is an intuitive behavior heavily related to that of the delay spread shown in Fig. 6-12. These results are consistent with those presented in Fig. 6-13 and Fig. 6-5, where respectively, the mean pass rate and the efficiency increase with the BW.



(a) LOS (b) NLOS

Figure 6-17: I_K as a function of N_f and the BW.

6.4.2 SKG from space variability

Regarding the space variability, the channel vector is constructed by Alice positions over the grid, providing a virtual antennas-array. We consider three specific cases: a linear antenna array formed with two successive positions over the grid ($N_{ant} = 2$), and two square antenna arrays with both $N_{ant} = 4$ and $N_{ant} = 64$. The statistics necessary to compute the statistical quantity I_K are the frequency variability computed over a BW of 500 MHz with a separation of 2.5 MHz. For each position of Bob, a distribution of I_K values may be obtained by shifting the array in order to cover all the possibilities over the Alice's grid. The results are shown in Fig. 6-18, also by distinguishing between LOS and NLOS cases.

Again, increasing the channel DoFs yields an increase in the shared available key bits between Alice and Bob. Furthermore, NLOS channels provide greater values of I_K than LOS, although the channels in both cases are Rayleigh fading. This relies on the angular spread of the paths which seems to be greater in NLOS channels. In the same context, the authors in [58] proved that NLOS channels are prone to provide more randomness than LOS ones undergoing Rician fading, where the dominant path should be removed, which will yield a decrease in the power (i.e. SNR). Moreover, the authors in [58] tested SKG performance by performing indoor measurement campaigns at 2.55 GHz band and using MIMO channels. Specifically, at the same SNR of 15 dB, I_K in [58] attains values (i.e. ranging from 40 to 100 bits) larger than that reached in our measurements (I_K varies between 40 to 80), even with smaller degrees of freedom (32 instead of 64 DoFs). This may be ascribed to the fact that

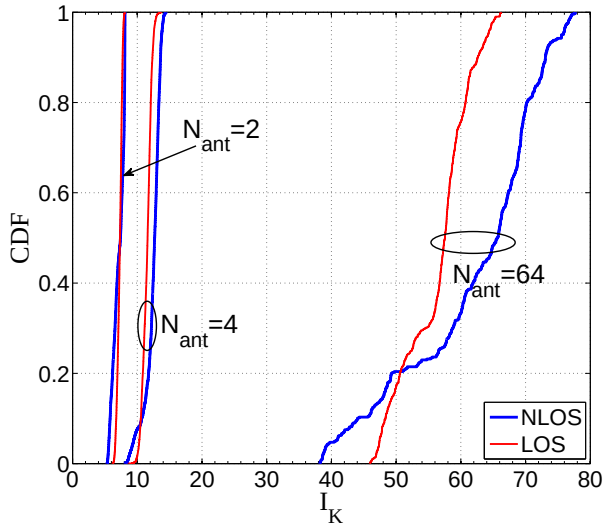


Figure 6-18: I_K vs. N_{ant} from space variability.

they use MIMO channels, exploiting the angular spread at both the transmitter and the receiver sides, while in our case, we just exploit the channel diversity at a single side of the channel, i.e. at Alice (multiple input single output: MISO).

Although the randomness provided in the space domain is always better than that provided in the frequency domain (fig. 6-16), according to NIST, the two domains contain, according to I_K, the same amount of randomness if higher BWs, i.e. 40 or 160 MHz are exploited.

6.4.3 Relative vulnerable key rate

We intend here to evaluate the relative vulnerable key rate I_{VK}/I_K for NB single antenna channels. Each terminal is considered as either Bob or Eve. Mutual information is computed over the 121 Alice's positions. I_{VK} is evaluated by considering just the channels measured by both Bob and Eve when Alice is the transmitter since the latter is responsible of the time channel variability. Fig. 6-19 depicts I_{VK}/I_K against the separation distance between Bob and Eve. It is obvious that the relative vulnerable key rate decreases as the distance between Bob and Eve increases. Eve may still obtain little information correlated with the channel measured by Bob even at large distances with respect to λ (e.g. 55 λ). These results are consistent with those obtained above in this dissertation, however they are in contrast with the ideal Clarke scenario. In fact, this means that some paths may be correlated between Bob and Eve and they are of relatively high power.

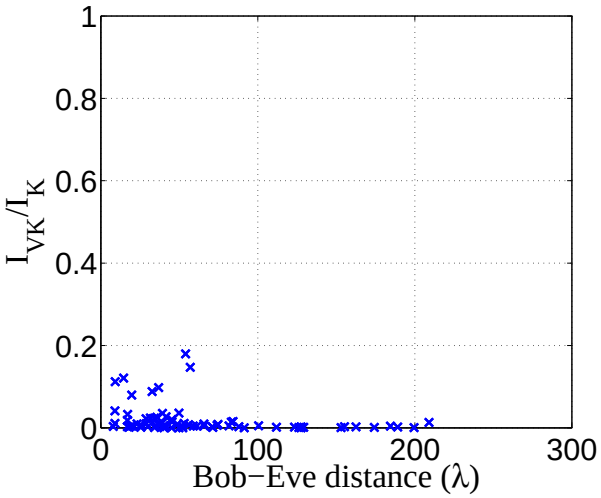


Figure 6-19: I_{VK}/I_K vs. Bob-Eve separation distance.

6.5 Conclusion

In this chapter, we presented a study of the SKG approach in indoor radio channels at the WIFI bands.

Regarding the quantization CQA protocol, Alice and Bob should choose the appropriate mapping scheme, given the SNR, in order to agree on the largest key size while having a sufficient number of identical key bits. When this is done correctly, the BER between Bob and Eve's key bits after CQA is close to 0.5 and the proportion of cases for which the BER deviates significantly from this target value is small enough that downstream techniques such as privacy amplification should be sufficient to guarantee a very good degree of confidentiality for Alice-Bob communication, unless of a very close or very powerful Eve. While we focus in this work on the channel features by just using the CQA algorithm, the complete scheme of SKG, including the reconciliation and the privacy amplification, is targeted in the PhyLaws project either for indoor or outdoor channels and using either WIFI or LTE signals.

Notably, while the amount of random information is almost the same when exploiting either the space or the frequency variability in the outdoor environment, at least from the RT results of Chapter 5, such an amount is poor when merely exploiting the frequency domain in indoor environments. Indeed, the security performance when exploiting the frequency variability relies on the statistics of the delay spread, which is higher in outdoor environments than in indoor ones and reveals that channels in outdoor environments are more frequency-selective and contains more random infor-

mation. In other words, indoor environments such as those measured here will need to use at least two antennas in order to capture a sufficient number of decorrelated key bits (e.g. 128).

Chapter 7

Complete Secret Key Generation Scheme

While the previous chapters evaluate the robustness of the generated keys directly after the channel quantization phase, we implement here the whole SKG strategy (Fig. 2-6), including information reconciliation and privacy amplification. The information reconciliation phase allows Alice and Bob to agree on the same key bits, by using error correcting codes such as LDPC codes [7, 61, 62, 65] and BCH codes [55, 126]. The last step, i.e. privacy amplification, consists of randomizing the key and removing information leaked to the eavesdropper.

We discuss the robustness of the proposed PhySec scheme by using secure sketch based on BCH codes for information reconciliation and hash functions for the privacy amplification, as proposed in [126]. To this end, we consider the indoor measured channels, presented in Chapter 6. We restrict the analysis to WIFI signals centered at 5.4 GHz band, while in the PHYLAWS project, both WIFI and LTE signals are processed through both indoor and outdoor measurements [126].

Furthermore, we discuss how the definition of the quantization maps may affect the generated key robustness. More precisely, since the CQA algorithm statistically defines the maps, how much statistical data (relatively to the key length) is required in order to achieve high key quality in terms of reliability and randomness.

7.1 Quantization

As admitted throughout this dissertation, the channel coefficients are quantized into bit streams through the CQA algorithm, where two alternative maps are defined in the I/Q domain. The maps are statistically determined after aggregating certain channel samples. Thus, the question is: how many channel samples are necessary to obtain reliable random keys?

We propose, relatively to the key length, two different sizes/densities of channel samples that may be used in the learning phase (the quantization maps definition):

1. The first method, called **single key-single map** (SKSM), consists in merely considering channel observations that are used in the construction of the key. In other words, SKSM just uses the same channel observations as those used to construct a given key. For example, quantization maps may be formed over just e.g. 64 channel coefficients when the number of quantization regions is 4 ($M = 4$), and for a key of 128 bits. The benefit of such a method is that it is fast and nearly simultaneous with the construction of the key. The drawback is an anticipated lesser robustness, since few observations imply more disagreement between Alice and Bob maps.

2. On the contrary, the second way to build a map, called **multiple key-single map** (MKSM), consists of accumulating a lot of channel samples, where most of them are not used in a unique key construction. In other words, several keys may be constructed from different observations, while all observations are used in the definition of an unique map. The drawback of such an approach is that the learning phase may be slow, however we may expect a better agreement between Alice and Bob maps, owing to the larger number of observations. Such a scheme is here implemented by accumulating all the channel samples available for the 121 Alice positions and the 108 sub-carriers, thus for each position of Bob[1].

7.2 Information reconciliation

Information reconciliation aims to suppress the remaining mismatches between Alice and Bob keys, by employing an error correcting code. In the PHYLAWS project,

[1]The SNR is averaged over all these channel samples, which are used to define the quantization map.

keys reconciliation is implemented through the use of secure sketch based on error correcting codes [127]. Without loss of generality, we suppose that Alice is the leader so that her computed key (K_a) is considered as the secret key, while Bob attempts to retrieve Alice's key using the key (K_b) he extracts from his own channel measurements. The protocol can be described as follows:

Alice

- selects a random codeword c from an error correcting code
- computes the secure sketch $s = K_a \oplus c$
- sends s to Bob over the public channel

Bob

- subtracts s from its computed key K_b: $c_b = K_b \oplus s = K_b \oplus K_a \oplus c$
- decodes c_b to recover the random codeword c and gets $\hat{c}$
- estimates K_a by computing: $\hat{K_a} = \hat{c} \oplus s$

Full reconciliation occurs when Bob perfectly retrieves the random codeword chosen by Alice ($\hat{c} = c$). As a result, Alice and Bob agree on exactly the same key ($\hat{K_a} = K_a$). The public transmission of the secure sketch, which allows the exact recovery of the key, may reveal some information to Eve, even though it does not disclose the exact value of the key. Indeed, similarly to Bob, Eve uses the secure sketch to retrieve the key K_a. Therefore, some amount of information may be revealed to Eve during this phase. In order to suppress this information leakage and randomize the key, a final step is required. This may be done through the Privacy amplification, as described in the next section.

7.3 Privacy amplification

As already mentioned, we employ privacy amplification [67, 128] in order to achieve a high key randomness, while decreasing the amount of vulnerable information. This may be achieved through either hash functions or randomness extractor, which may reduce the final key length by erasing redundant information. The efficiency of such

a process may require some information about Eve's knowledge in order to maximize the conditional entropy of the generated key, as discussed in [67].

We made use of the privacy amplification scheme developed in PHYLAWS [126], through the two-universal family of hash functions [67, 128]. We note that a class of hash functions $\{g\}$ is defined as two-universal if, from two distinct strings (i.e. $x_1 \neq x_2$), it is nearly impossible to obtain the same result (i.e. $g(x_1) \neq g(x_2)$) [67, 128]. What is particularly important here, is that with a suitable hashing scheme, this property holds even though x_1 and x_2 (i.e. the keys) are extremely close (e.g. differing by a single bit). In other words, even though Eve succeeds in obtaining nearly the same key as Alice/Bob, after the privacy amplification step, her key becomes totally different from Alice/Bob's one. Thus, Alice and Bob publicly agree on a random function (g), which, even known from Eve, reveals nearly negligible information about the final key. This hash function, defined as $g : \{0,1\}^n \rightarrow \{0,1\}^r$, transforms a partially secret key of length n into a highly secret but shorter key, with a length $r \leq n$.

Without loss of generality, we implement the privacy amplification by choosing a specific function (g) from the two-universal family of hash functions. For that, we interpret the shared key K as an element of the Galois field $GF(2^n)$, where n is the number of bits of K. For $1 \leq r \leq n$ and for a random element $a \in GF(2^n)$, g assigns to the key K the first r bits of the product $a.K \in GF(2^n)$, i.e. $g : (K)_n \rightarrow (a.K)_r$. Let us recall that the multiplication in the Galois field is a polynomial product, which can be expressed as a convolution between bit sequences. Such an operation is performed modulo an irreducible polynomial of degree n over $GF(2^n)$. Thus, a little difference between two initial binary strings, even over just one bit, yields to a wide gap between the resulting strings.

In practice, at each new key computation, Alice randomly choose $a \in GF(2^n)$ and sends it to Bob over the public channel. Both Alice and Bob compute the galois field product $a.K \in GF(2^n)$, which is considered as the final secret key.

More clearly, the hash mechanism spreads any bit error all over the final key $(a.K)_r$ (i.e. first r bits of $a.K$), thus:

- When Eve applies the privacy amplification on her key (rather her reconciled key), any remaining disagreement between her key and that shared between Alice and Bob after the reconciliation step, implies quite different computed keys $(a.K)_r$, which make her key unusable.

- From Bob point of view, he agrees on the same key generated by Alice once the keys are perfectly reconciled during the information reconciliation phase. Thus, the key is useful for legitimate users but not for Eve.

7.4 Results

In this section, we analyze the robustness of the key, in terms both of terminals disagreement and key randomness, and after each phase of the secret key generation process: 1) At the quantization phase, keys of 128 bits are generated using $M = 4$, using either SKSM or MKSM; 2) A secure sketch based on (127, 92, 11) BCH code[2] is then performed for information reconciliation, which requires to keep only the first 127 bits of the key. 3) Finally, the privacy amplification is applied with $r = n = 127$. Although more privacy would be achieved with $r < n$, we made this choice since it allows the comparison between key randomness quality before and after privacy amplification, while using the same NIST tests.

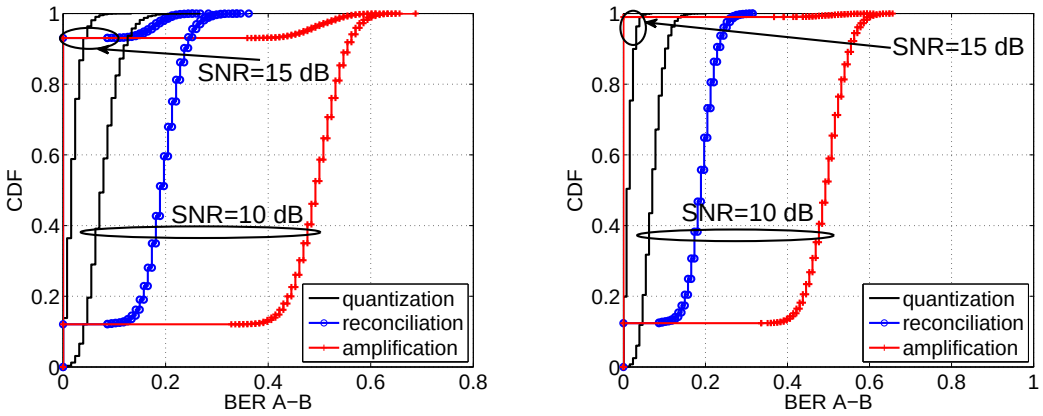
According to the previous chapters, the keys are generated by exploiting different channel degrees of freedom, e.g. in the space and the frequency domain, where the parallel channels are randomly selected from the available set of channels. For the frequency DoF, we investigate channels with a BW of 40 MHz. In addition to the three phases of SKG, the results are discussed according to the way used to define the quantization maps (i.e. SKSM vs. MKSM).

7.4.1 Alice-Bob disagreement

A successful SKG scheme allows Alice and Bob to agree on the same key, which consequently may be useful for encrypting data. Thus, we intend here to verify the agreement between Alice and Bob after considering the whole SKG scheme as well as considering each step alone, as presented in Fig. 7-1. We note that these keys are extracted by exploiting the space variability, while key statistics are constructed over both Bob positions and different sub-carriers. Fig. 7-1 also compares the BER between Alice and Bob keys between SKSM and MKSM, for different SNR (10 and 15 dB).

The information reconciliation phase allows Alice and Bob to perfectly reconcile a certain proportion of the quantized bits, which intuitively increases with the SNR.

[2]Perfect reconciliation is performed when the keys differ by at most 5 bits.



(a) SKSM (b) MKSM

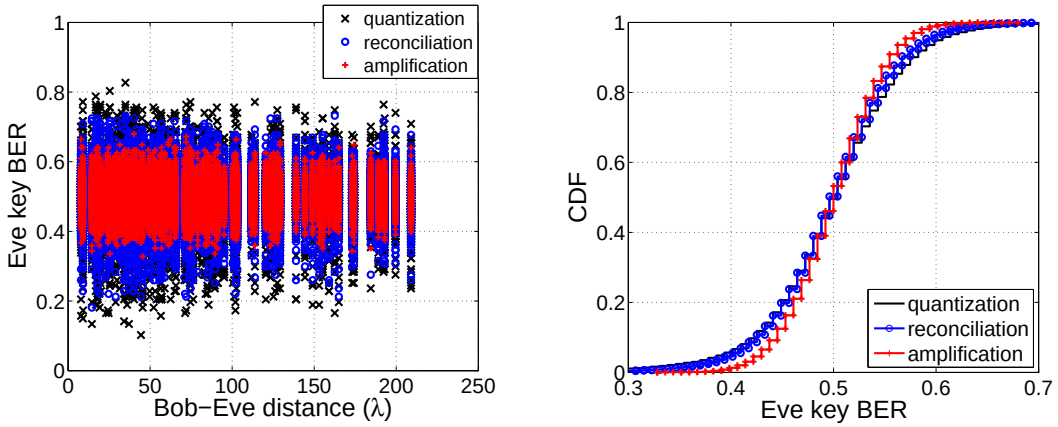
Figure 7-1: Disagreement between Alice and Bob keys.

More precisely, the keys are perfectly corrected once the BER after CQA is at most 4%, which results from a capability to correct at most 5 bits from 127 key bits. This is consistent with the parameters chosen for the BCH code, with a Hamming distance of 11. In order to increase the keys agreement ratio, we note that the parameters of the BCH code could be adapted according mainly to the channel SNR, and also to the required key length (e.g. increasing the Hamming distance for low SNR). Furthermore, when increasing the capability of correcting mismatches, the amount of redundant information (publicly exchanged) increases, which may reveal some information to Eve. Such a vulnerable information may be removed during the privacy amplification step, which outputs an exactly identical key for both Alice and Bob once perfect reconciliation occurs. However, the cost to pay is the reduce in the key length. Therefore, potential schemes require to perform perfect reconciliation with the least amount of public messages.

By comparing the key agreement for both SKSM and MKSM shown respectively in Fig. 7-1(a) and Fig .7-1(b), we observe that the proportion of perfect reconciliation increases with the latter scheme, i.e. MKSM. For $SNR = 15$ dB, 99% of keys are perfectly corrected when using MKSM, while it is 93% for SKSM. Indeed, when building maps from a relatively big number of observations, Alice and Bob are able to obtain approximately the same maps.

7.4.2 Bob-Eve disagreement

Similarly to Alice-Bob disagreement, we present in Fig. 7-2 the BER between the keys constructed by Bob and Eve after each step of SKG, for space DoF and for the SKSM approach. Obviously, owing to space decorrelation, the BER is small after quantization and becomes closer to 0.5 after the privacy amplification. Indeed, as previously mentioned, the privacy amplification spreads the errors all over the key. More precisely, the BER after privacy amplification is varying in a range centered around 0.5, i.e. in $[0.4; 0.6]$. This may be ascribed to the statistical independence between the keys generated by Bob and Eve, with a uniform distribution of bits. In other terms, this may be described by keys following the binomial law, i.e. $B(127, 0.5)$, as shown in Fig. 7-3. We note that the results are similar for MKSM.



(a) BER vs. Bob-Eve separation distance

(b) The CDF of the BER

Figure 7-2: Disagreement between Bob and Eve keys for SKSM.

7.4.3 Key randomness

The randomness of the generated keys was firstly assessed using the statistical NIST tests in Chapter 6, where the keys were derived merely from CQA and no privacy amplification was considered. We extend this randomness assessment in this section, by considering the privacy amplification step and different definitions of the quantization map (SKSM ans MKSM). We note that we consider a key length of 127 bits, even after the privacy amplification ($r = 127$). However, in practice, the key length r should be reduced in order to erase most of the leaked information.

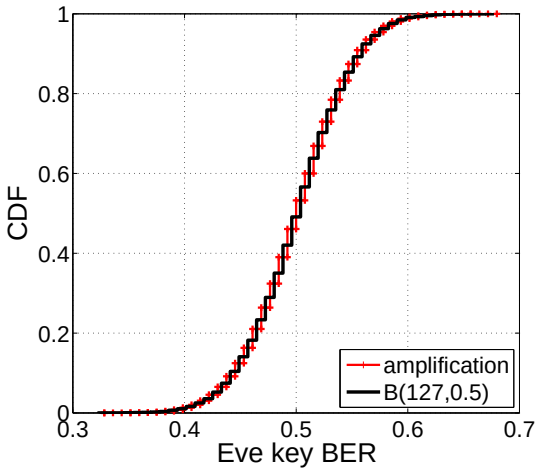


Figure 7-3: Eve key BER after privacy amplification vs. the BER for keys following B(127,0.5).

7.4.3.1 Results for the SKSM approach

We first consider the SKSM approach where the maps are defined from the channel samples directly used to build each 128 bits key. We take back some results presented in the previous chapter and compare key randomness right after the quantization scheme to that after the final privacy amplification phase. Hence, Fig. 7-4 compares the mean pass rates for keys in both cases, for $M = 4$ and $BW = 40$ MHz. Obviously, the privacy amplification improves the key randomness, whatever the channel variability. It is observable also that the mean pass rates are not equal to 1, since according to NIST, each key is characterized random if the proportion that passes the test fall into a range of acceptable proportions, without the necessity of reaching 1.

As discussed in Section 3.1.4 and Section 6.3, the NIST tests are not able to judge a perfect randomness of the key. They are rather able to evaluate a specific defect according to a specific criteria, e.g. the tests described in Section 3.1.4. Therefore, we resort to a visual interpretation of the key randomness, as admitted in [75]. More clearly, two dimensional keys are plotted, as in Fig. 7-5.

Graphical representations of keys extracted by exploiting the spatial DoFs, the frequency DoFs, the joint space-frequency DoFs for both $N_{ant} = 2$ and $N_{ant} = 4$ are plotted respectively in Fig. 7-5, Fig. 7-6, Fig. 7-7 and Fig. 7-8. These keys are shown for different Bob positions, using SKSM, and more importantly right after quantization or after privacy amplification. It is clearly shown that a much better randomness is provided by the complete scheme, either over each generated key (vertical axis) or

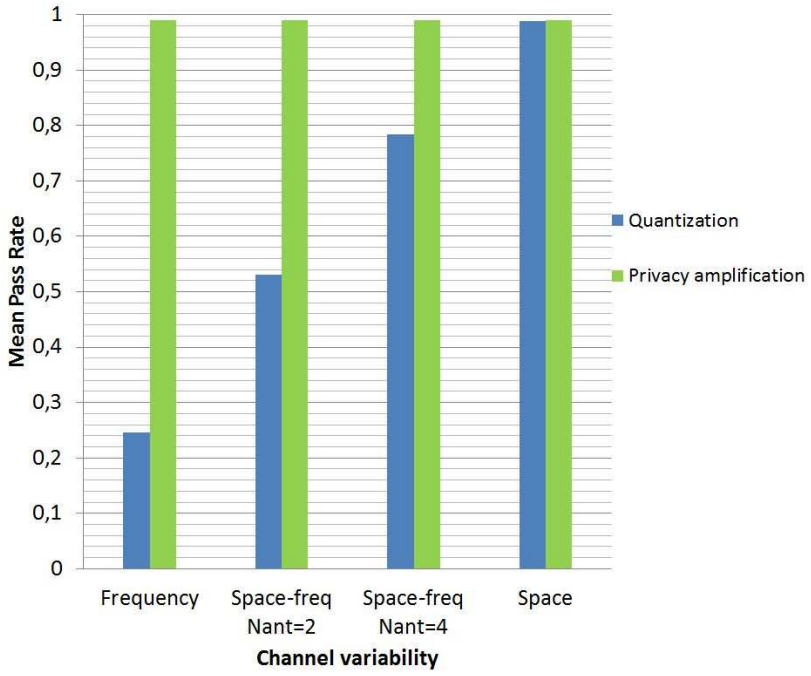
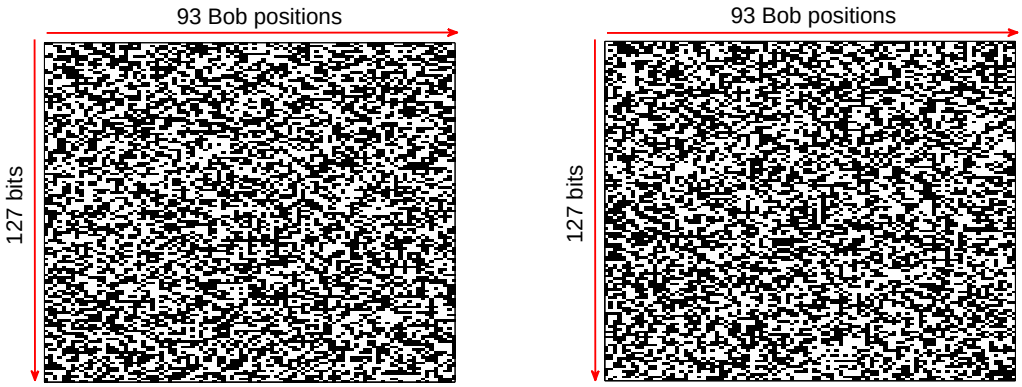


Figure 7-4: Mean pass rates before and after privacy amplification, for SKSM.

over different keys at different Bob positions (horizontal axis). Let us now focus on the randomness provided by the quantization since the more vulnerability, the shorter the key after privacy amplification (although simulations are performed with $r = n$). More clearly, Eve intelligence is exploited at the quantization phase in order to maximize the mutual information between the channels, and consequently, reduce the final key length (after the last step of SKG).

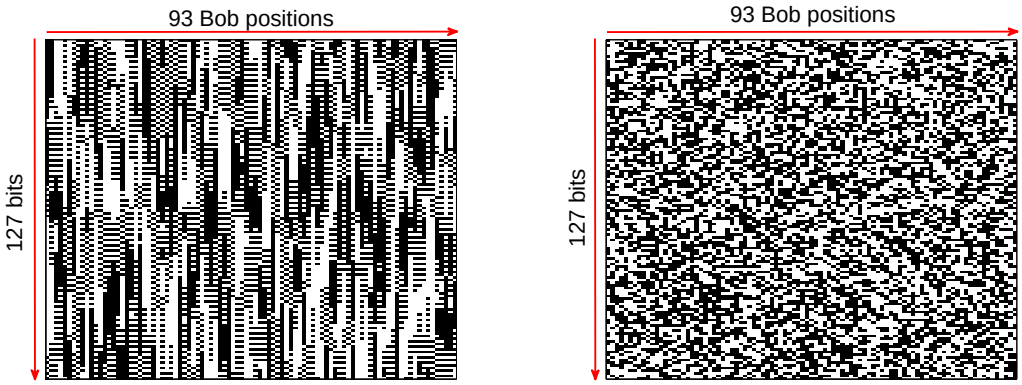
Consistently with Fig. 7-4, we see that the randomness is improved from frequency variability to joint space-frequency variability, then to spatial variability. However, the graphical representations still show some deterministic behaviors that would be useful to Eve. In particular, in Fig. 7-6, the vertical keys show some long sequences with the same bit value, which is due to the small channel coherence bandwidth or, in other words, to the high channel correlations. Moreover, in Fig. 7-7 and Fig. 7-8, deterministic behaviors are shown through the repetition of some string patterns of certain length. Furthermore, we can notice that the keys are almost uncorrelated from one position of Bob to another one. This means that it is possible for different terminals to generate different keys.

Fig. 7-9 plots key matrices generated using the frequency variability, for each Alice position over the grid, and for a given position of Bob. Thus, we consider both LOS and NLOS radio channels between Alice and Bob. It is apparent, for keys generated



(a) Quantization (b) Privacy amplification

Figure 7-5: Keys from spatial DoFs, for different Bob positions and for SKSM.

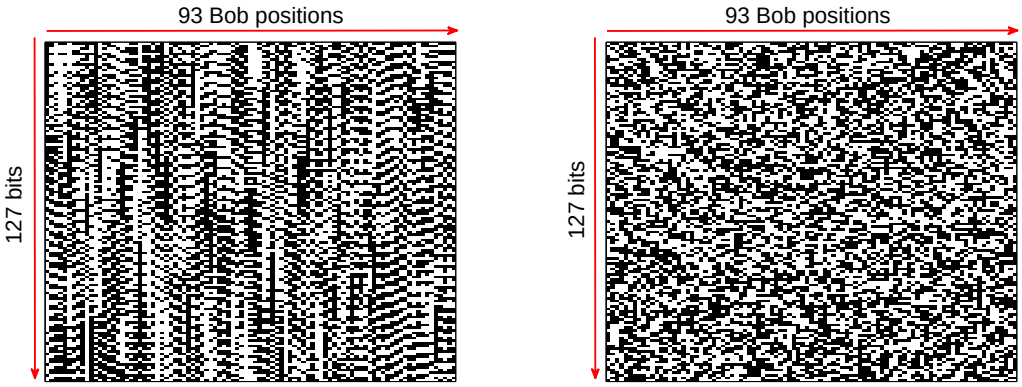


(a) Quantization (b) Privacy amplification

Figure 7-6: Keys from frequency DoFs, for different Bob positions and for SKSM.

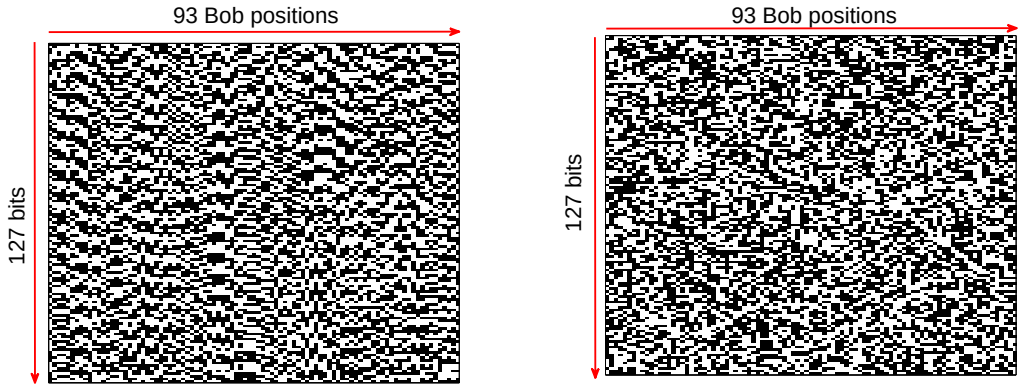
right after CQA, that slightly more randomness is present in the NLOS channels. Nevertheless, both channels are very deterministic owing to the long sequences of bit e.g. 1. In the LOS case (Fig. 7-9(a)), the variation of the bits from one key to another is very periodic, such that the bit values is changing with respect to the LOS path phase. After privacy amplification, it turns out that a very good randomness is achieved in both cases, although a sharp eye would likely be able to capture subtle structures characterizing weak randomness imperfections.

To summarize, for raw keys (i.e. right after CQA), most of the randomness is provided by the spatial variability. The joint space-frequency variability may be useful since it leaks less information about the radio channel and it follows less deterministic



(a) Quantization (b) Privacy amplification

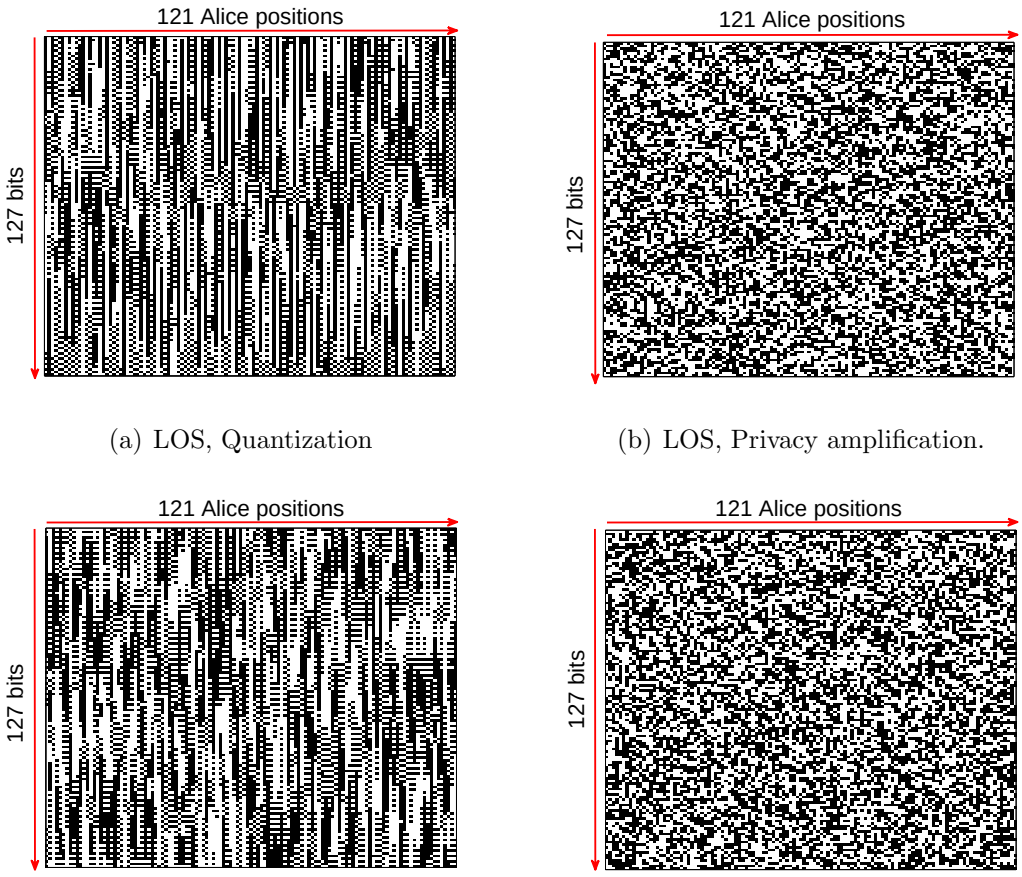
Figure 7-7: Keys from joint space-frequency DoFs ($N_{ant} = 2$), for different Bob positions and for SKSM.



(a) Quantization (b) Privacy amplification

Figure 7-8: Keys from joint space-frequency DoFs ($N_{ant} = 4$), for different Bob positions and for SKSM.

behavior, especially when increasing the number of investigated antennas. Furthermore, privacy amplification significantly increases the randomness of generated keys. Despite that, the price to pay is in principle a small key length $r < n$ (even it is not considered in simulations), especially when more information is revealed to Eve during the reconciliation step. Therefore, it is interesting to assess the amount of randomness available on raw keys (just after CQA), since Eve may exploit any deterministic behavior to guess the key.



(a) LOS, Quantization

(b) LOS, Privacy amplification.

(c) NLOS, Quantization

(d) NLOS, Privacy amplification.

Figure 7-9: Keys from frequency DoFs, over Alice positions and for SKSM.

7.4.3.2 Results for the MKSM approach

We here generate keys by defining once the quantization map using all the channel samples provided by the spatial variability of the Alice grid and the frequency variability, resulting in a single map for each position of Bob. This is explained above as the MKSM approach.

Fig. 7-10 shows the mean pass rates for keys generated using MKSM and when exploiting both the space and the frequency variability. As expected, the random character of the keys is enhanced after the privacy amplification phase. However, the randomness before this step is very poor, especially when comparing with that provided by SKSM (Fig. 7-4). This may be ascribed to the fact that using MKSM, the number of bits equal to 1 and 0 in each key is not identical, which results in keys with single value for example, especially for the frequency variability. This idea may

be clarified through Fig. 7-11 and Fig. 7-12.

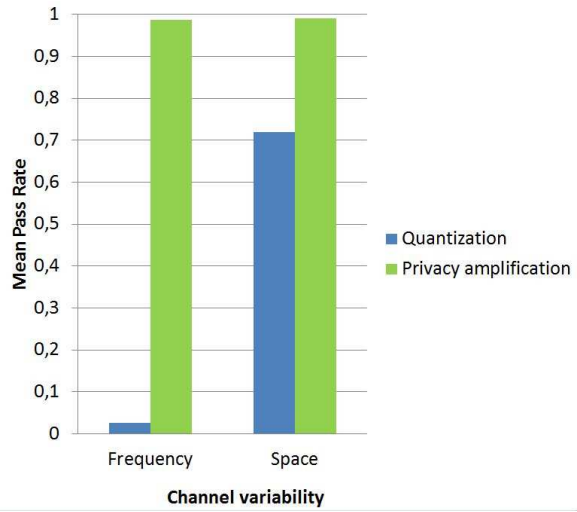
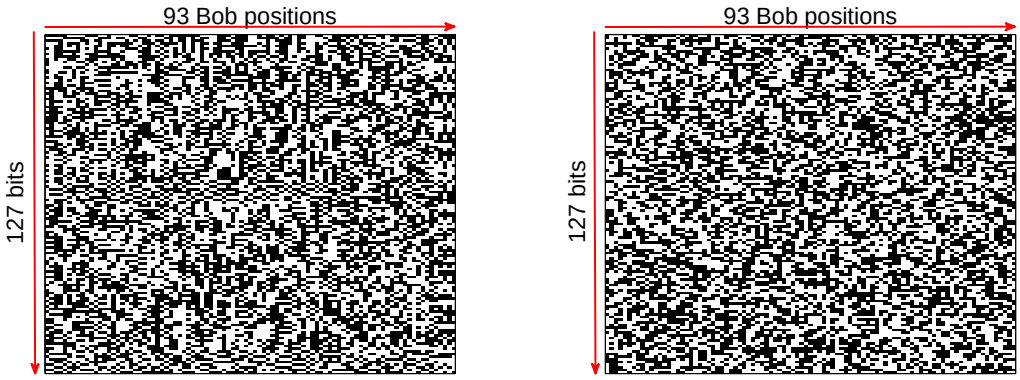


Figure 7-10: Mean pass rates before and after privacy amplification, for MKSM.

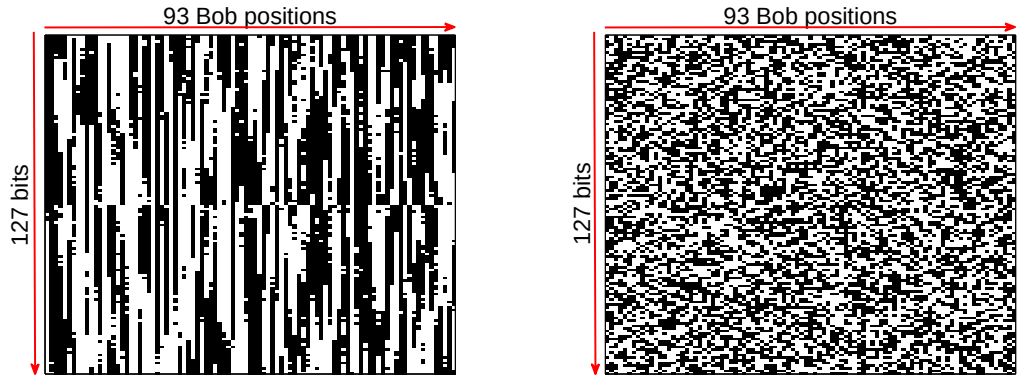
Fig. 7-11 and Fig. 7-12 illustrate key matrices computed respectively over the space and the frequency variability. Such keys are quantized from maps computed based on the MKSM approach. We still see high randomness degree for the spatial variability, even slightly less than that shown for the SKSM. Also similarly to the results presented above, the keys generated from the frequency variability are almost deterministic owing to the long stream of identical key bits. However, the difference is that the number of bit 0 and bit 1 is not identical for each generated key, as we can remark for all the keys generated from the SKSM approach. In particular, we can notice some keys formed with the same value of the bits. We note that the mid line in Fig. 7-12(a) is ascribed to the way of concatenating the bit streams, i.e. here the bits generated over the real values are first concatenated, then those of the imaginary part.

Fig. 7-13 displays the key matrices generated for two specific positions of Bob, both in LOS and NLOS conditions, where the frequency variability is exploited. Again, for quantization, these channels are very poor in randomness. Moreover, the frequency variability when using the MKSM approach is the worst. One point to notice, indeed, is that MKSM is not perfectly suited to each generated key. While the map is optimal as a whole, it is sub-optimal for the individual generated keys, for which each set of channel observations does not necessarily respect the equal probability for sub-intervals, as targeted by CQA (see Section 3.2.2).



(a) Quantization (b) Privacy amplification

Figure 7-11: Keys from spatial DoFs, over Bob positions and for MKSM.



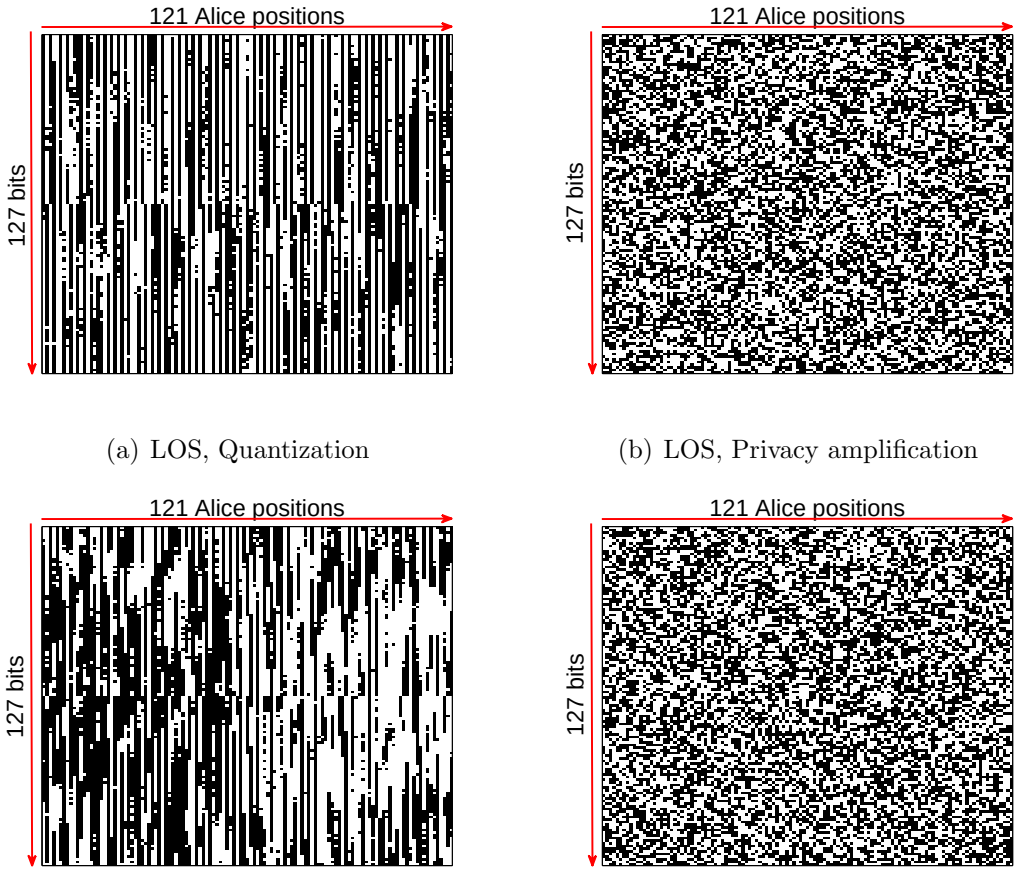
(a) Quantization (b) Privacy amplification

Figure 7-12: Keys from frequency DoFs, over Bob positions and for MKSM.

7.5 Conclusion

We have implemented in the present chapter the whole scheme of SKG from reciprocal radio channels. Although CQA is used throughout the dissertation to extract key bits from the channel coefficients, we proposed two different approaches to compute the maps (i.e. the SKSM and the MKSM approaches). Information reconciliation was achieved with a secure sketch based on BCH codes, while privacy amplification was applied through two-universal of family hash functions.

It has been shown that once the keys are perfectly reconciled, Alice and Bob agree on the same key length after the privacy amplification. With respect to Eve,



(a) LOS, Quantization

(b) LOS, Privacy amplification

(c) NLOS, Quantization

(d) NLOS, Privacy amplification

Figure 7-13: Keys from the frequency DoFs, over Alice positions and for MKSM.

the discrepancies between her key and Alice/Bob key increase after the privacy amplification. We also found that the graphical representation of the key matrices was easier to interpret in terms of randomness imperfections than the performance figures provided by NIST tests, especially for MKSM.

Furthermore, we have seen that the MKSM approach allows Alice and Bob to easily agree on the same key bits, with a proportion higher than that enabled by the other approach. However, more key randomness may be provided by the SKSM approach. Consequently, we remark that a trade-off between key reliability and key randomness results from the use of these approaches.

Chapter 8

Conclusion and perspectives

8.1 Conclusion

In this dissertation, we have addressed the role of the radio channel in terms of its impact on SKG performance. The work has been carried out as regards:

- The identification of important features of the radio channel regarding SKG.
- The selection of a few metrics intended to evaluate the performance quantitatively.
- The application of these metrics to modeled and measured channels.

Among the results, we have shown that, subject to an efficient exploitation of the channel richness with an adequate algorithm able to capture the details of the complex channel coefficients, suitably long keys could be obtained, either from the sole frequency domain or from the combination of several domains, notably in the space or Doppler domains.

Indoor environments are less prone to SKG, since such environments are naturally rather time stable (low Doppler) and with large coherence bandwidths. However, by exploiting also the spatial domain (e.g. 2 or 4 antennas), and with a wide enough BW (such as 40 MHz and up to 160 MHz, as in WIFI variants), much can be done to obtain suitably long keys.

In outdoor environments, the channel richness and the longer delay spreads favor large degrees of freedom and small coherence bandwidths. This implies that it is

relatively easy to construct suitably long keys, even though the effective DOF (seen the secret key rates) are less than for i.i.d. sub-channels. This is even better, since outdoor channels are expected to be more time variant, due to the Doppler incurred by moving vehicles, a moving terminal or any other short or long distance time variant disturbance.

Generally speaking, it is interesting to increase the number of sub-channels, even when they are correlated, since this increases somewhat the total secret key rate. However, this also means to devise clever schemes, able to produce long keys made of independent bits from very long ones made of correlated bits. This is typically what can be done in privacy amplification.

Nevertheless, it has been shown that information reconciliation and privacy amplification allow Alice and Bob to reliably share a key with a high randomness degree, whatever the channel DoF. On the other hand, the privacy amplification may increase the Eve key BER although some information are leaked during the earlier phases of SKG. Furthermore, a trade-off between reliability and randomness has been shown in relation to the statistical map computation.

8.2 Perspectives

Several important issues could not investigated in this task, owing to the specific PHYLAWS objectives and, above all, to the available resources. Some of them would require further efforts in order to consolidate the fundamental basis on SKG performance. Others may be addressed according to a practical approach, based on the observed performance in specific cases. This includes:

- The evaluation of secret key capacity in non Gaussian cases, i.e. when we want to ascertain the vulnerability of key bits when Eve can take any possible position around Bob, in other words well beyond a small scale area for which Rayleigh fading ensures that the complex channel coefficients are Gaussian distributed. Such an evaluation requires going back to the basics of mutual information (see Eq. 3.5 and Eq. 3.6) and attempt to compute I_K and I_{SK} from such general formulas, directly or approximately.

- The analysis of SKG performance in terms of antenna characteristics (e.g. antenna radiation pattern). Although omnidirectional scenarios favor SKG owing to the capability to capture dispersive paths, it has been shown in literature

that specific design of antennas would be advantageous to create channel fluctuations in static environment [55, 80, 81]. Furthermore, it seems interesting to assess SKG for different kind of antennas, especially from Eve point of view.

- Deeply investigate SKG from the eavesdropper point of view. Throughout this thesis, the passive Eve merely derives the key from her own measurements, without any effort to get more insight about the Alice/Bob channel. This may be done through ray tracing algorithms, provided diffuse scattering was suitably implemented.

- Taking into account many non ideal features that would obscure the similarity of channels seen by Alice and Bob, such as deviations from reciprocity. This occurs in the radio frequency signal processing stages, in particular due to the non linearity of certain devices and their non reciprocal character by nature (such as amplifiers, which are commonly highly directional and non reciprocal). The imperfect calibration of signal acquisition blocks, their variation with time, temperature etc. are common causes of non reciprocity. While this can be modeled [32, 129], another simplest way is to assessed the differences between channels measured from both ways of the link in available devices. VNAs are extremely performant equipments and it is commonplace that the non reciprocity is vanishingly small, owing to accurate calibration. However, in commercial wireless equipments, much less is done in this direction, resulting in mismatched measurements.

- Trying to pre-process the channel data, prior to extracting keys. One particular benefit would be through the removal of dominant paths, responsible for the Rician fading rather than Rayleigh and of more vulnerability. Advanced processing might be useful, such as principal components removal [58]. However, this would come at a price: the reduced power and consequently of the SNR. So a trade-off should be sought in order to see to what extent the security would be enhanced by this method.

- As seen in Chapter 7, privacy amplification phase is responsible of ensuring security through increasing the key randomness and increasing the Eve key BER. We have also mentioned that it implies key length reduction in order to decrease the amount of information revealed to Eve during the earlier SKG phase. Therefore, it is interesting to assess how much the key should be reduced in order to ensure robust key generation.

Résumé en Français

1. Introduction

Avec l'explosion des technologies sans fil et la multiplication des services associés, on a de plus en plus besoin de sécurité et de confidentialité. Bien que la cryptographie symétrique assure la confidentialité des données, elle est pénalisée par la gestion (i.e. la génération et la distribution) de clés qui doivent être tenues secrètes.

Des études récentes indiquent que les caractéristiques intrinsèques de la propagation peuvent être exploitées afin de consolider la sécurité. Ces nouvelles techniques, agissant au niveau de la couche physique, assurent une sécurité du point de vue de la théorie de l'information. Plus clairement et contrairement à la cryptographie classique, l'espion est supposé jouir d'une capacité de calcul illimitée et supposé ne pas avoir des informations suffisantes pour lui permettre d'affaiblir la confidentialité.

En particulier, le canal radio fournit une source d'aléa commune à deux utilisateurs à partir de laquelle des clés secrètes peuvent être générées. En effet, en raison de la réciprocité du canal, typiquement dans un système à multiplexage temporel, un émetteur et un récepteur mesurent presque le même canal de propagation, ce qui implique le partage de la même source. En outre, l'aléa est assuré par la propagation multi-trajets (Figure 1), qui aboutit à des interférences constructives/destructives et implique des propriétés de décorrélation dans les domaines temporel, spatial et/ou fréquentiel. Pour ces mêmes raisons, un espion ne parviendra pas à obtenir la clé partagée par les utilisateurs légitimes.

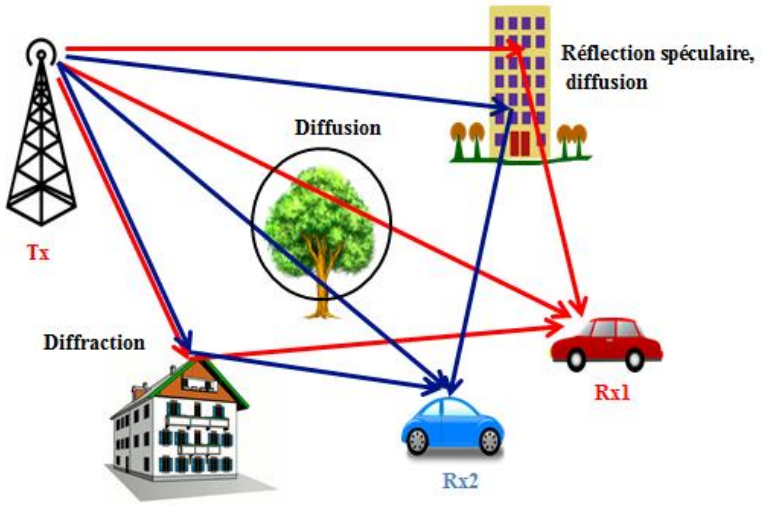


Figure 1 Propagation multi-trajets

Une clé de cryptage se caractérise par sa longueur, son caractère aléatoire ainsi que par son taux de sécurité. Une longue clé résulte de la concaténation de plusieurs symboles quantifiés à partir de plusieurs échantillons du canal. Néanmoins, les bits de la clé doivent être statistiquement indépendants afin d'assurer son caractère aléatoire. Cela dépend de la corrélation des canaux dans le domaine de variation considéré. En outre, bien que le taux de

sécurité de la clé dépende de son caractère aléatoire, il nécessite aussi le minimum d'informations divulguées à l'espion. Cela concerne le scénario relatif de l'espion vis à vis des terminaux légitimes, par exemple la distance entre Eve et Alice/Bob.

Le travail de cette thèse s'inscrit dans le cadre du projet européen PHYLAWS (PHYsical LAyer Wireless Security) qui a pour objet de renforcer la sécurité des systèmes actuels tout en exploitant les propriétés inhérentes de la couche physique. PHYLAWS vise ainsi à développer des techniques efficaces et flexibles, qui sont faciles à implémenter et consomment peu de ressources. Il concerne de nombreuses applications comme l'internet des objets ainsi que différents standards comme WIFI et LTE.

On s'intéresse dans cette thèse, en particulier, au processus de génération de clés secrètes (« Secret Key Generation », SKG) à partir de l'aléa du canal de propagation. On vise à analyser la performance de la SKG en fonction des caractéristiques réelles du canal radio et ceci à travers différents types de canaux, i.e. des canaux simulés à partir d'un modèle stochastique du canal, des canaux simulés par tracés de rayons et des canaux mesurés. Cette étude de performance repose sur la qualité des clés générées à partir de l'exploitation des degrés de liberté du canal dans les domaines temporel, spatial et/ou fréquentiel et dans différents types d'environnement. D'autre part, il s'agit d'examiner la vulnérabilité de la SKG vis-à-vis du scénario de l'espion, tel que la distance entre l'espion et au moins l'un des terminaux légitimes dans un environnement donné.

2. Méthodes d'évaluation de la SKG

Soient Alice et Bob deux terminaux légitimes qui souhaitent communiquer en toute sécurité en présence d'un espion Eve (Figure 2). Les terminaux mesurent des canaux altérés par du bruit et des erreurs d'estimation qu'on suppose complexes Gaussiens ($\widehat{\boldsymbol{h}}_{\boldsymbol{x}} = \boldsymbol{h}_{\boldsymbol{x}} + \boldsymbol{n}_{\boldsymbol{x}}$). Ces canaux peuvent être des vecteurs de sous-canaux mesurés dans le cas d'un système multi-antennaires ou OFDM.

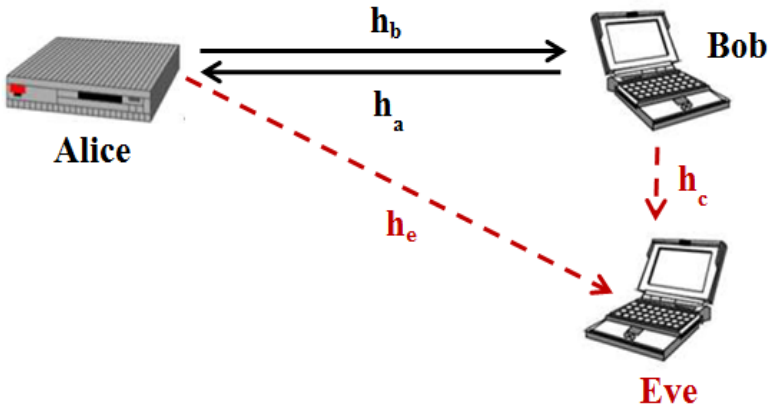


Figure 2 Scenario de communication secrète en présence d'un espion

L'approche de génération de clés secrètes s'accomplit en plusieurs étapes, comme montre la Figure 3. Après l'estimation du canal radio, les terminaux légitimes quantifient l'information

du canal (telle que la puissance, la phase ou le coefficient complexe) en un ensemble de bits constituant la clé. En raison du bruit additif et des erreurs d'estimation, les clés générées par Alice et Bob ne sont pas les mêmes. D'où la nécessité de l'étape de réconciliation qui aboutit à des clés identiques grâce à la mise en œuvre de codes correcteurs d'erreurs. En outre, les messages échangés publiquement entre les terminaux légitimes révèlent des informations à l'espion. L'amplification de confidentialité sert ainsi à réduire la vulnérabilité potentielle qui en résulte et à augmenter l'aléa de la clé, bien qu'aboutissant à des clés plus courtes. En effet, même si Eve obtient une clé très proche de celle d'Alice/Bob (par exemple avec une différence d'un seul bit), cette dernière étape amplifie l'erreur en la distribuant sur la totalité de la clé grâce à des fonctions de hachage.

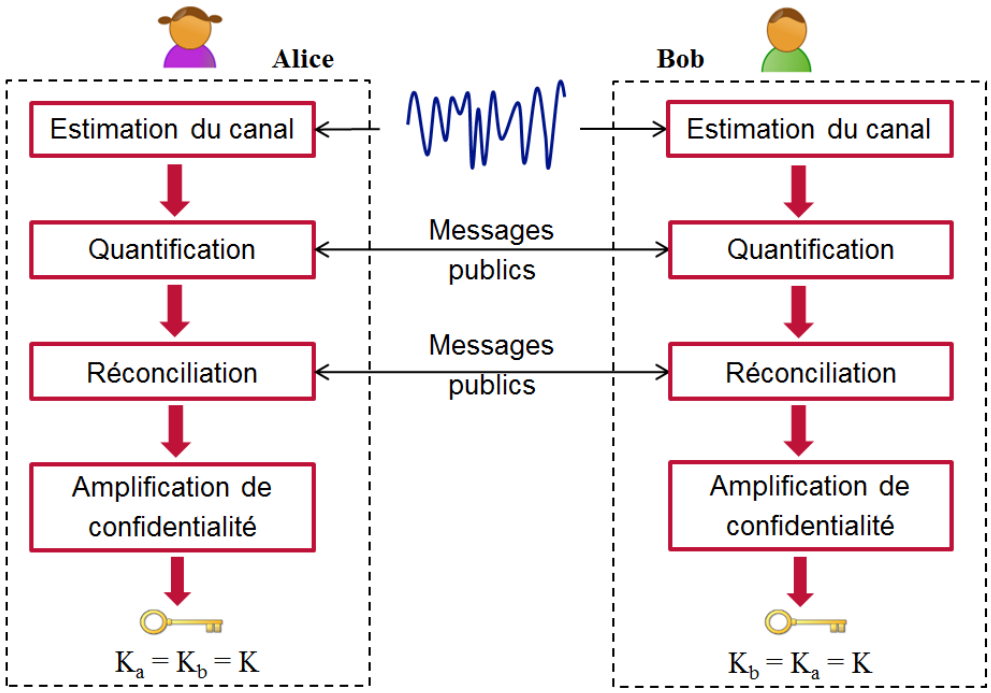


Figure 3 La stratégie complète de la SKG

2.1.Métriques d'évaluation de performance de la SKG

Théoriquement, la longueur de la clé que les différents terminaux peuvent partager se mesure en fonction de l'information mutuelle de leurs canaux. On définit ainsi la longueur de clé disponible comme étant la longueur maximale de la clé que les terminaux légitimes peuvent partager en toute fiabilité, i.e. $I_K = I(\hat{h}_a, \hat{h}_b)$. En tenant compte des observations faites par l'espion, on obtient le nombre de bits secrets, i.e. $I_{SK} = I(\hat{h}_a, \hat{h}_b | \hat{h}_e, \hat{h}_c)$, ainsi que celui de bits vulnérables, i.e. $I_{VK} = I_K - I_{SK}$. Dans le cas des canaux conjointement Gaussiens, l'information mutuelle est facilement évaluée à partir des caractéristiques du second ordre (matrices de corrélations).

Considérons le cas d'un canal Gaussien avec du bruit additif Gaussien, qui est un canal aléatoire, l'évaluation numérique montre que I_K est limité par le rapport signal sur bruit (« signal to noise ratio », SNR) (Figure 4). Par ailleurs, on considère un scénario de Clarke

pour avoir un premier accès au nombre de bits vulnérables. Le résultat montre qu'une distance d'une demi-longueur d'onde (λ/2) entre Bob et Eve est suffisante pour annuler la vulnérabilité (Figure 5). Ce scénario de Clarke est idéal et correspond à un canal où le nombre de trajets est infini. En revanche, la réalité est différente, comme on va le voir dans la suite.

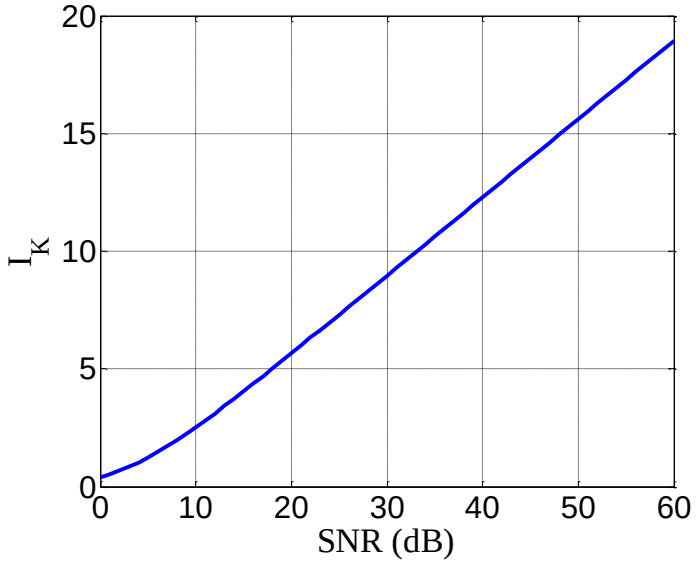


Figure 4 $\boldsymbol{I_K}$en fonction du SNR

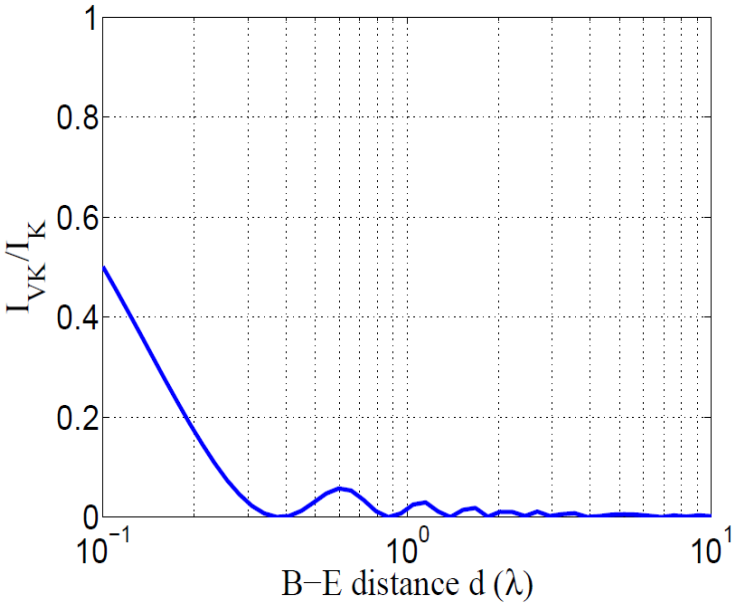


Figure 5 $\boldsymbol{I_{VK}/I_K}$ en fonction de la distance de séparation Bob-Eve

D'autre part, les métriques concrètes résultent de l'évaluation de clés binaires après l'implémentation de la stratégie de la SKG, que ce soit après la quantification ou dans les étapes postérieures. On obtient alors un taux d'erreur binaire (« bit error rate », BER) résultant de la comparaison directe entre les clés générées. En outre, le caractère aléatoire des clés est évalué à travers un ensemble de tests statistiques développés par l'Institut National de Standards et de Technologie (NIST) [94]. Chaque test identifie l'aléa selon un certain critère.

On a choisi quelques tests pertinents et on a regardé le taux moyen de succès aux tests (i.e. mean pass rate) afin de simplifier l'analyse.

2.2. L'étape de quantification

On a choisi de quantifier le coefficient complexe du canal en raison de sa richesse en information, surtout du fait qu'il contient à la fois l'information sur la puissance et sur la phase. Pour transformer cette information en un certain nombre de bits, on a choisi d'implémenter l'algorithme de quantification alternatif (« channel quantization alternating », CQA), qui utilise deux plans de quantification en alternance de telle sorte que le plan choisi soit adapté à l'observation du canal [58]. En d'autres termes, cela permet de réduire l'erreur entre les clés d'Alice et de Bob lorsque le coefficient du canal est proche d'une frontière entre deux symboles différents. L'intérêt de cet algorithme est qu'il améliore l'accord entre Alice et Bob sans perte d'information et qu'il fournit des symboles équiprobables sur I et sur Q, ce qui aide à générer des clés aléatoires.

2.3. Canal dispersif dans le temps

Comme une clé résulte de la concaténation de plusieurs symboles, sa robustesse est en partie liée à la décorrélation entre les symboles, ce qui requiert une décorrélation entre les coefficients du canal. Cela implique de disposer d'échantillons suffisamment séparés dans n'importe quel domaine et dépend des caractéristiques de l'environnement radio (surtout la richesse en multi-trajets). Par conséquent, la variabilité du canal est essentielle pour achever la SKG avec de meilleures performances.

Les plus simples variations du canal auront lieu dans le temps grâce à la mobilité des terminaux ou à des mouvements dans leur entourage. En revanche, un tel scénario n'est pas toujours possible, par exemple lorsqu'il s'agit d'une communication entre un point d'accès et un ordinateur fixe dans un environnement statique. De ce fait, on propose d'appliquer la SKG tout en exploitant autres types de variations du canal et cela à travers l'exploitation des degrés de liberté dans les domaines spatial et fréquentiel.

Cette partie est dédiée à l'analyse de la qualité des clés générées tout en exploitant le degré de liberté fréquentiel. Pour cela, on considère un modèle simple de profil de puissance-retard (« Power delay profile », PDP) où les trajets discrets sont périodiques dans le domaine temporel et avec une puissance moyenne exponentiellement décroissante. Les trajets sont indépendants et leurs amplitudes sont distribuées selon la loi Rayleigh. Par une simple transformée de Fourier, on obtient un ensemble de coefficients du canal dans le domaine fréquentiel.

Idéalement, le nombre de bits disponibles (I_K) croît proportionnellement au nombre de sous-porteuses (N_f) utilisées dans la génération de clés. Toutefois, à cause des corrélations entre les sous-porteuses, cette croissance se fait avec une pente plus faible. En fait, I_K augmente linéairement avec N_f jusqu'à une certaine valeur au-delà de laquelle la pente de la courbe diminue montrant une tendance à la saturation (Figure 6). Au point de déviation, les degrés de liberté sont pleinement exploités dans la SKG où l'ensemble de fréquences utilisées permet de

résoudre parfaitement les multi-trajets dans le domaine temporel par des sinus cardinaux. Puisque les trajets filtrés sont indépendants, la SKG profite du degré de liberté de chaque trajet tout en considérant son SNR. Au-delà de la première déviation, la courbe continue à augmenter, même avec une pente plus faible, en montrant plusieurs résonances. En effet, la densité spectrale de puissance de chaque sous-porteuse est constante. L'augmentation de N_f implique ainsi une augmentation de la puissance totale et ensuite une amélioration du SNR de chaque trajet résolu.

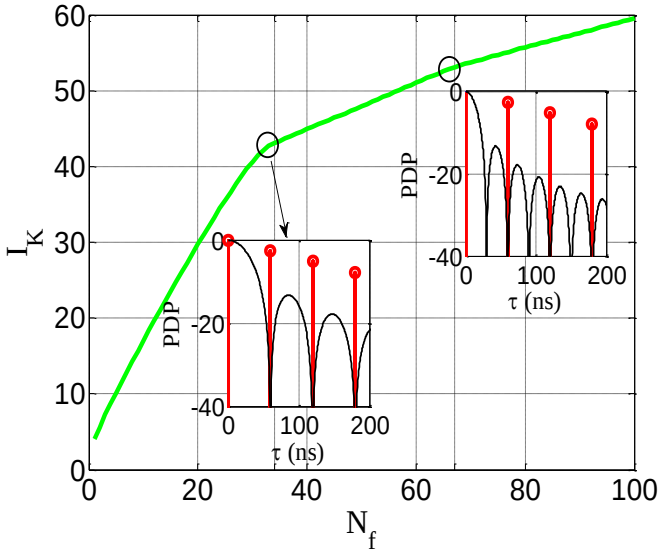


Figure 6 L'évolution de $\boldsymbol{I_K}$ en fonction du nombre de sous-porteuses

Par ailleurs, pour un SNR fixe par sous-porteuse, l'augmentation de l'étalement temporel σ_τ aboutit à une augmentation de la puissance relative de chaque trajet. Cela implique un canal plus riche en aléa et par conséquent des clés plus longues (Figure 7).

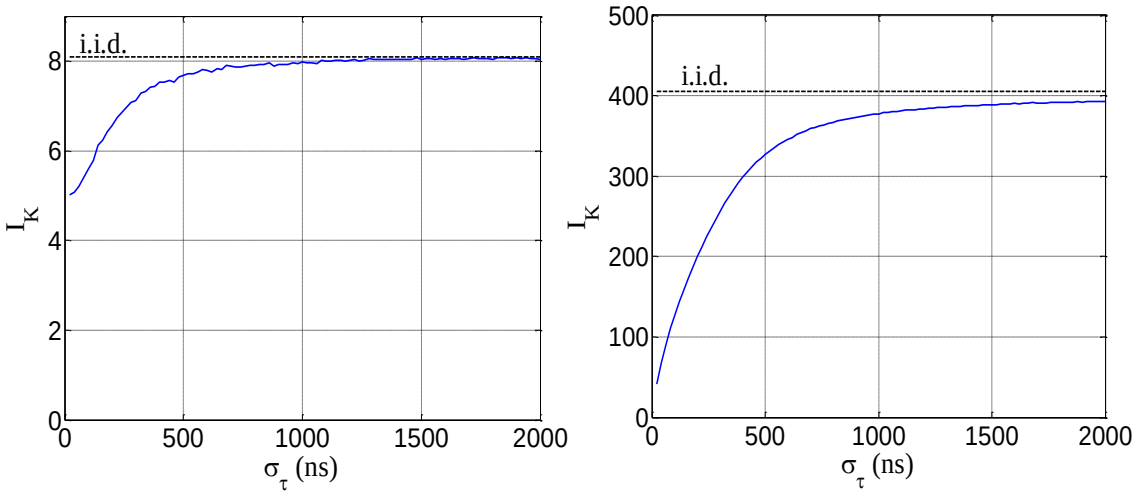


Figure 7 $\boldsymbol{I_K}$ en fonction de l'étalement temporel $\boldsymbol{\sigma_\tau}$ pour $\boldsymbol{N_f = 2}$ (gauche) et $\boldsymbol{N_f = 100}$ (droite)

3. Modèle du canal adapté à la SKG

L'objectif de cette partie est d'étudier l'impact du scénario de l'espion Eve sur la qualité des clés générées. Dans ce but, on a besoin d'un canal multi-utilisateurs qui prend en compte la corrélation spatiale entre deux utilisateurs à proximité l'un de l'autre.

Dans la littérature, la SKG est étudiée tout en considérant deux cas extrêmes : le premier correspond à un espion très lointain des terminaux légitimes, alors que le deuxième correspond à un scénario où Bob et Eve sont très proches l'un de l'autre et où ils partagent les mêmes composantes de multi-trajets. Comme ils ne sont pas co-localisés, ils partagent les mêmes puissances de multi-trajets alors que les phases changent sur une distance d'une fraction de λ. Les résultats dépendent de la richesse en multi-trajets.

D'autre part, bien que la littérature sur les modèles de canaux soit très riche, ces modèles ne sont pas dédiés à la problématique de la SKG et ne répondent pas à son besoin spécifique, tel que la nécessité d'une modélisation fidèle de la corrélation spatiale.

Pour toutes ces raisons, et afin d'évaluer la SKG en utilisant un modèle de canal simple avec un petit nombre de paramètres, on a développé notre propre modèle de type stochastique et géométrique. On vise notamment à dépasser la stationnarité spatiale entre Bob et Eve.

3.1. Description du modèle

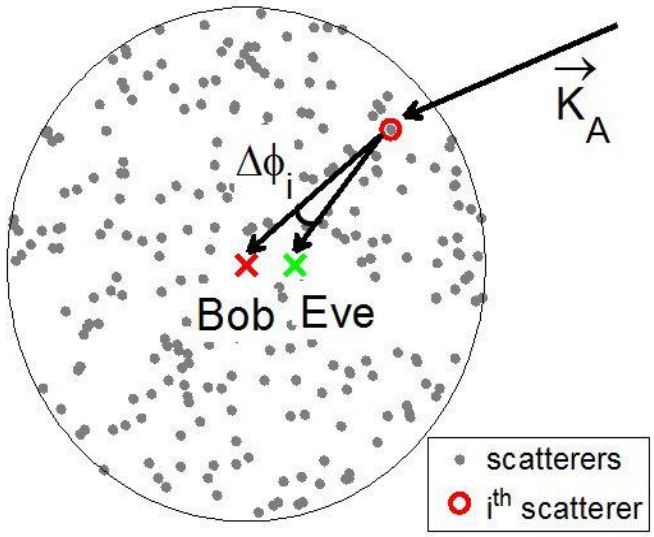


Figure 8 Modèle stochastique et géométrique

On considère un modèle stochastique et géométrique où les diffuseurs sont uniformément répartis dans un disque centré autour de Bob (Figure 8). Eve est à l'intérieur du disque et se trouve à une certaine distance de Bob. Quant à Alice, elle est supposée très lointaine, de sorte que les rayons envoyés vers le disque proviennent d'une direction unique. Les canaux vus par Bob et Eve sont calculés en utilisant les propriétés géométriques de la distribution des diffuseurs. Ainsi, on tient en compte de l'atténuation en espace libre et de la corrélation spatiale du shadowing en utilisant une loi log-normale. On suppose que les diffuseurs sont

indépendants les uns des autres, alors que les trajets émis par le même diffuseur vers Bob et Eve sont corrélés en fonction de leur écartement angulaire. Les terminaux sont équipés d'antennes omnidirectionnelles.

La génération d'une clé secrète requiert de l'aléa qui peut être fourni par une variation petite échelle (« Small Scale Fading », SSF). Pour cela, on fixe les paramètres macroscopiques et ceux du shadowing et on crée de la variabilité petite-échelle en bougeant les diffuseurs autour de leurs positions macroscopiques. Cela implique que la phase varie fortement, créant ainsi des variations SSF par le jeu des interférences et l'indépendance entre le mouvement des divers diffuseurs. La statistique petite échelle SSF permet d'obtenir une clé unique alors que la statistique macroscopique sert à fournir un ensemble de clés.

D'après la Figure **9**, on observe que lorsqu'Eve s'éloigne de Bob, les composantes des multi-trajets vus par chacun d'eux se décorrèlent, ce qui implique une décroissance de la vulnérabilité, en terme de proportion de bits vulnérables (I_{VK}/I_K) ainsi que de taux d'erreur binaire entre leurs clés (BER). En particulier, concernant I_{VK}/I_K, on constate qu'au-delà de λ/2, cette proportion ne diminue que de façon progressive avec la distance, montrant une mémoire spatiale allant bien au-delà de la longueur d'onde.

En comparant avec le scénario de Clarke, on aperçoit qu'il y a ainsi une vulnérabilité résiduelle au-delà de λ/2. Cela est expliqué par le nombre limite de trajets dont l'augmentation aboutit à moins de vulnérabilité. En revanche, le BER ne reflète pas la vulnérabilité résiduelle. En effet, l'algorithme CQA repose sur des changements majeurs du canal, ce qui ne permet pas à Eve d'exploiter toute l'information partagée avec les terminaux légitimes.

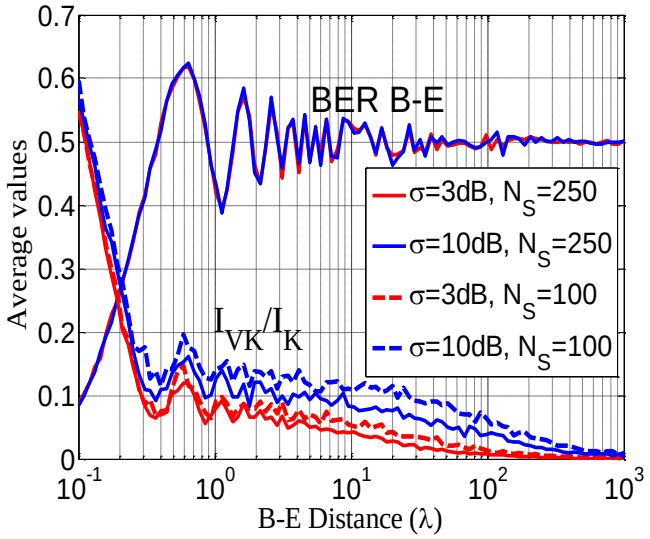


Figure 9 La vulnérabilité ($\boldsymbol{I_{VK}/I_K}$ et BER) en fonction de la distance Bob-Eve

4. Le comportement de la SKG pour des canaux déterministes simulés

On considère dans cette partie des canaux plus réalistes simulés avec l'outil de tracés de rayons (VolcanoLab de Siradel). L'intérêt est d'étudier la performance de la SKG en environnement extérieur, où il est difficile d'effectuer des mesures. On considère ainsi

l'environnement ouvert du Louvre (Paris) en mettant en relief l'implémentation du modèle de diffusion dans l'outil de simulation.

On considère un système OFDM de 64 sous-porteuses couvrant une bande de 20 MHz. La densité spectrale de puissance de chaque sous-porteuse est constante et son SNR moyen est fixe. Conformément aux résultats obtenus pour le modèle PDP (Figure 6), I_K augmente avec N_f en présentant une tendance à la saturation mais de façon sous-linéaire (Figure 10). Cette tendance indique que les principaux trajets sont résolus pour une bande inférieure à 20 MHz. En effet, puisque l'environnement est ouvert, la réponse impulsionnelle est longue et les trajets principaux sont bien étalés dans le temps. D'autre part, cela peut-être dû, en partie, aux limitations de l'outil de tracés de rayons où le nombre d'interactions rayons-obstacles est limité, ce qui implique une réponse impulsionnelle non continue (Figure 11).

On calcule les canaux en utilisant l'outil de tracés de rayons, avec et sans implémentation du modèle de diffusion. La Figure 10 met en relief la pertinence des trajets diffus sur la SKG où les valeurs de I_K atteignent le double de celles obtenues pour des canaux simulés sans le modèle de diffusion. En effet, les rayons diffus enrichissent le canal en des nouveaux trajets, ce qui implique plus d'aléa à exploiter dans la génération de clés.

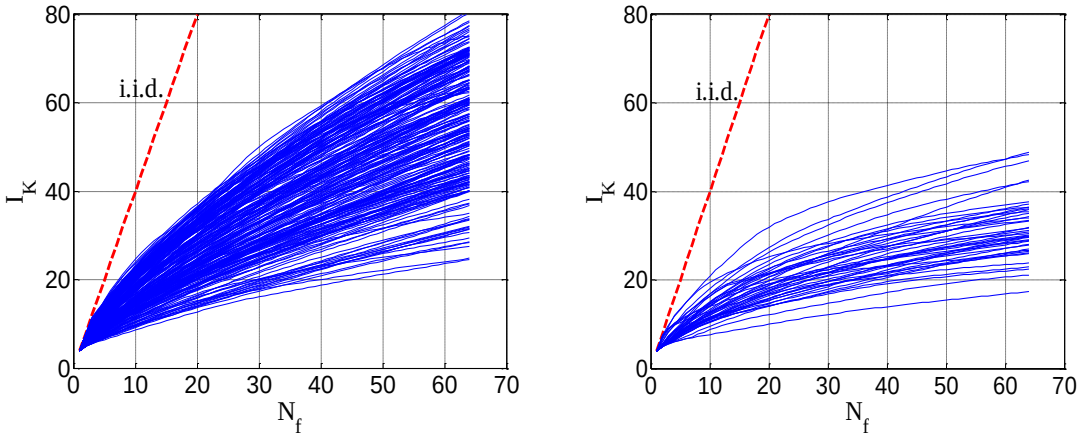


Figure 10 $\boldsymbol{I_K}$ en fonction de $\boldsymbol{N_f}$: avec diffusion (gauche) et sans diffusion (droite)

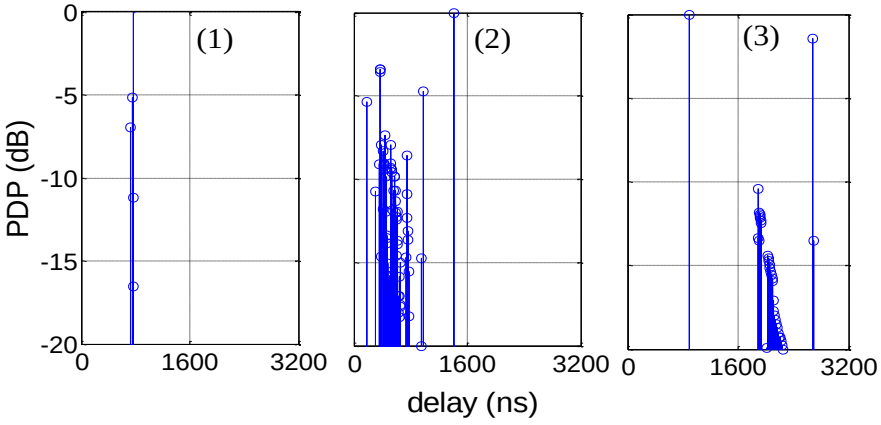


Figure 11 Des exemples de profils de puissance-délai

L'augmentation de l'étalement temporel σ_τ entraîne une réponse impulsionnelle plus étalée dans le temps, ou en d'autres termes, une bande de cohérence plus faible. Ce qui implique une croissance de I_K (Figure 12). Ce résultat est ainsi cohérent avec celui obtenu plus haut (Figure 7), sauf pour un grand N_f et pour des grandes valeurs de σ_τ. Ceci est expliqué par le fait qu'une réponse impulsionnelle plus étalée dans le temps ne signifie pas un canal plus riche en multi-trajets, mais plutôt des trajets bien espacés temporellement.

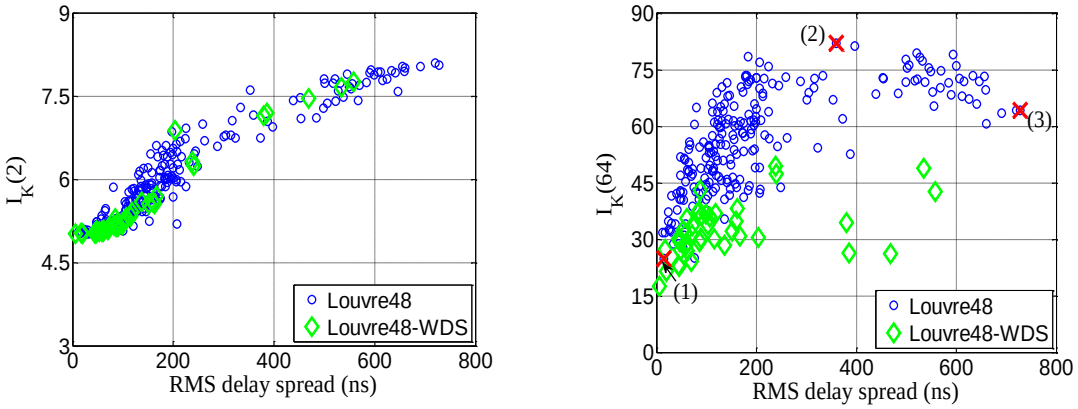


Figure 12 $\boldsymbol{I_K}$en fonction de delay spread: $\boldsymbol{N_f}$=2 (gauche) et $\boldsymbol{N_f}$=64 (droite)

La Figure 13 montre une comparaison statistique des valeurs de $\boldsymbol{I_K}$ résultant de l'exploitation des degrés de liberté spatial et fréquentiel. On note que le calcul d'une seule valeur de $\boldsymbol{I_K}$ dans le domaine fréquentiel repose sur une statistique petite-échelle en espace, alors que dans le domaine spatial, ce calcul utilise une statistique fréquentielle. Pour l'environnement étudié, il est fiable d'utiliser les deux domaines dans la SKG puisqu'ils contiennent statistiquement la même quantité d'aléa.

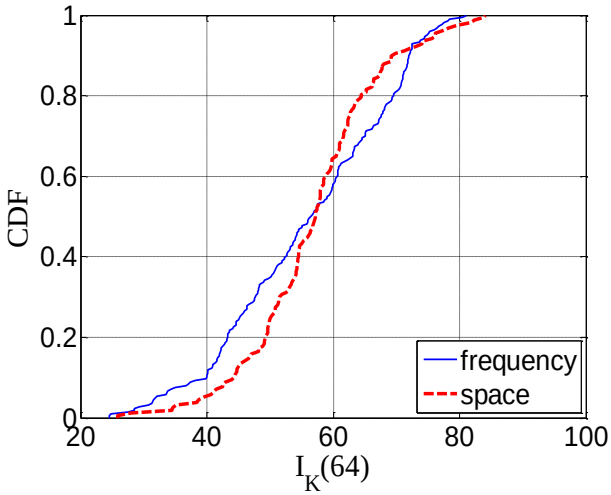


Figure 13 $\boldsymbol{I_K}$ dans les domaines spatiaux et fréquentiels

Conformément aux résultats observés dans la Figure **5** $\boldsymbol{I_{VK}/I_K}$ en fonction de la distance de séparation Bob-Eve, la proportion de bits vulnérables diminue avec la distance de séparation Bob-Eve en raison de la décorrélation des canaux. On peut toutefois remarquer l'existence

d'une vulnérabilité résiduelle, même pour de grandes distances, ce qui montre l'impact de l'environnement local sur la SKG (par exemple la richesse en multi-trajets). Par ailleurs, augmenter N_f apporte un avantage à Eve comme aux terminaux légitimes.

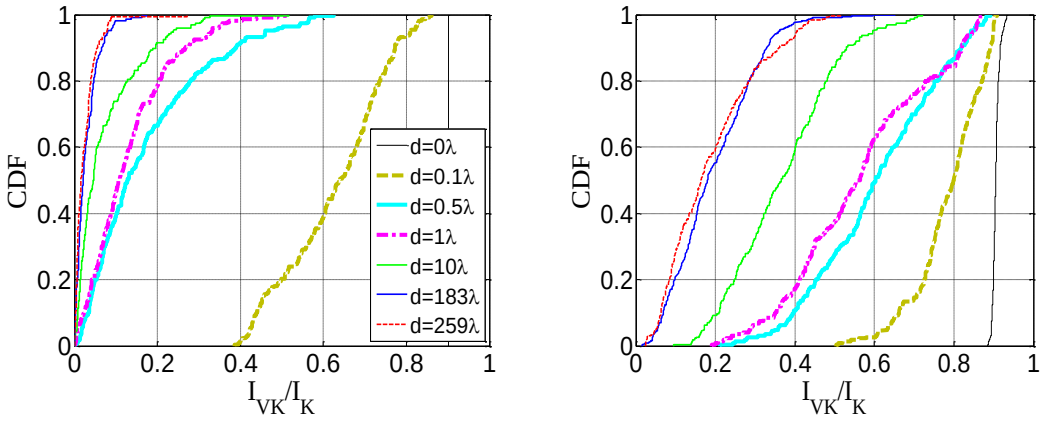


Figure 14 $\boldsymbol{I_{VK}/I_K}$ en fonction de la distance Bob-Eve et pour $\boldsymbol{N_f}$=1 (gauche) et $\boldsymbol{N_f}$=32 (droite)

5. Le comportement de la SKG pour des canaux mesurés

On vise dans cette partie à examiner la performance de la SKG pour des canaux réels et mesurés en environnement intérieur.

Les mesures ont été effectuées dans les locaux de Télécom ParisTech dans deux salles de dimensions et caractéristiques différentes (Figure 15 et Figure 16). Les coefficients complexes du canal ont été mesurés et enregistrés par un analyseur de réseaux (VNA) à 4 ports dont les paramètres de configuration sont indiqués dans le Tableau 1. Un port du VNA était dédié à Alice alors que les trois autres correspondaient à Bob et/ou Eve. Chaque port a été équipé d'une antenne ULB bicône de 2 dBi de gain.

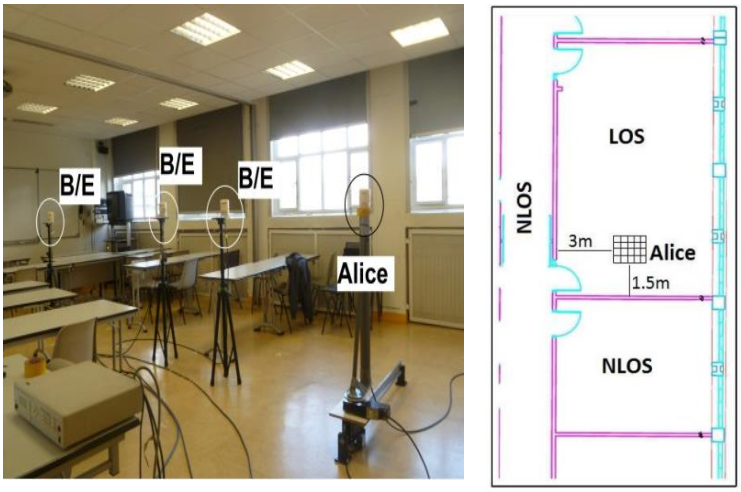


Figure 15 Mesures en salles de classe

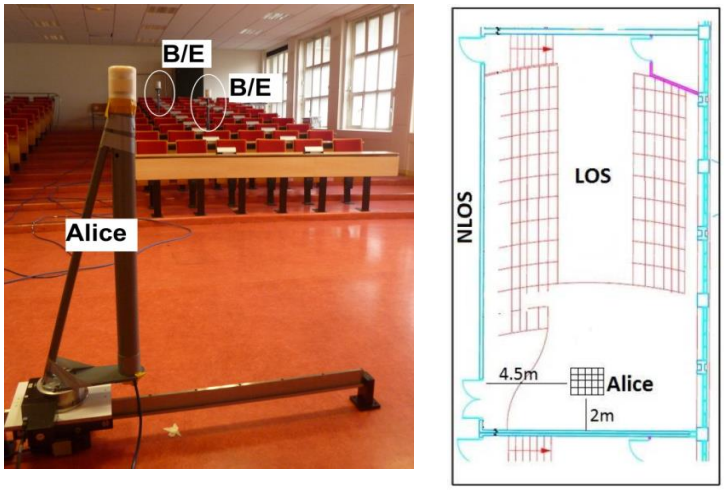


Figure 16 Mesures en amphithéâtre

Tableau 1 Les paramètres de configuration du VNA

Fréquences	2-6 GHz
Nombre de fréquences	1601
Bande IF	5 KHz
Puissance transmise	10 dBm
Dynamique de mesure	96 dB
Plancher de bruit typique	-86 dBm

Pour générer des clés, on a besoin de l'aléa. Pour ce but, lors de ces mesures, les trois récepteurs représentant Bob/Eve sont statiques alors que l'émetteur représentant Alice est spatialement balayé sur une grille carrée de 11x11 points (30 cm de côté et 3 cm de pas) confinée à une petite zone, de manière à capturer l'évanouissement petite-échelle (SSF). Plus clairement, puisque le pas de la grille est d'environ une demi-longueur d'onde à 5 GHz, les coefficients du canal sont très probablement statistiquement décorrélés.

5.1.La comparaison des clés générées par le trio Alice-Bob-Eve

Pour générer des clés, les terminaux appliquent l'algorithme CQA sur l'ensemble de coefficients complexes du canal qui constituent le plan de quantification. Ce plan est divisé en M régions de quantification, ce qui résulte en un symbole de $\sqrt{M}$ bits par coefficient du canal. D'après la Figure **17**, le taux d'erreur binaire des clés générées par Alice et Bob diminue quand le SNR augmente et quand M diminue. Par conséquent et pour une meilleure performance, le nombre de bits ($\sqrt{M}$) à générer par observation du canal doit être adapté au SNR de manière à ne pas dépasser un certain seuil de BER.

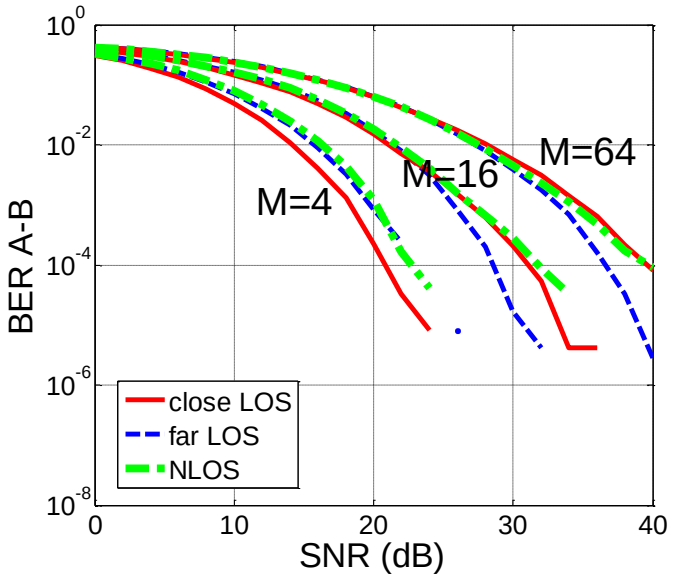


Figure 17 Le BER entre les clés d'Alice et de Bob en fonction du SNR

Eve exploite ses mesures pour générer une clé en espérant qu'elle soit très proche de celle obtenue par les terminaux légitimes. Elle applique CQA sur les canaux mesurés quand l'émetteur est Alice (h_{ea}) ou Bob (h_{eb}), et obtient ainsi deux valeurs de clés. D'après la Figure **18**, le BER est autour de 0.5, ce qui signifie que les clés générées par Eve sont presque complètement différentes de la clé légitime. Cependant, la vulnérabilité augmente quand Eve est en vue directe de Bob et quand Alice est l'émettrice puisque l'aléa mesuré par Bob et Eve est apporté par les déplacements d'Alice sur la grille.

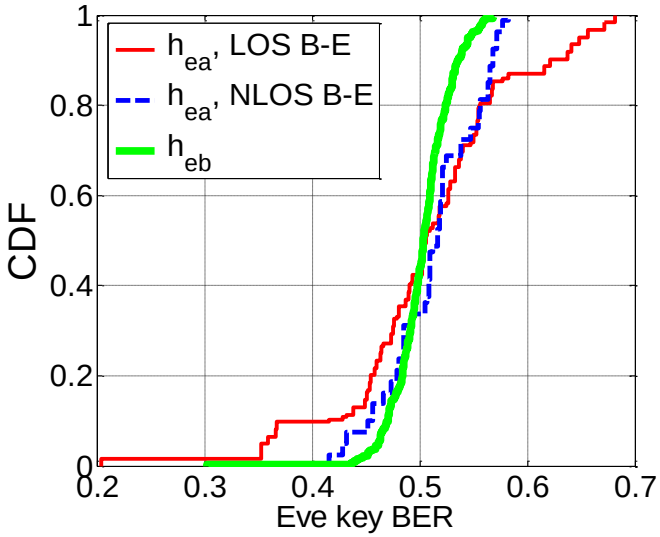


Figure 18 Taux d'erreur binaire par rapport à l'espion

5.2. Evaluation de la quantité d'aléa fournie par différents degrés de liberté

En exploitant le degré de liberté spatial, les clés sont générées en concaténant des symboles issus des coefficients complexes du canal sélectionnés de façon arbitraire parmi ceux

correspondant aux 121 positions d'Alice sur la grille carrée. Un ensemble de clés résulte d'une sélection arbitraire sur différentes sous-porteuses de la bande 20 MHz. Quant au degré de liberté fréquentiel, l'ensemble des clés résulte d'une sélection aléatoire de sous-porteuses dans une certaine bande, pour différentes positions d'Alice sur la grille. Chaque clé est fabriquée par accumulation des bits quantifiés à partir de sous-porteuses sélectionnées de façon arbitraire. En effet, pour construire une clé de 128 bits et en utilisant $M = 4$, on a besoin de quantifier 64 sous-porteuses dans les bandes de 40, 80 ou 160 MHz.

Les tests de NIST montrent que les clés dérivées du degré de liberté spatial sont suffisamment aléatoires (Figure 19). Cela est dû au fait que le canal est échantillonné avec un pas minimal de λ/2 à 5 GHz. En outre, l'exploitation du degré de liberté fréquentiel (Figure 20) montre que le caractère aléatoire s'améliore avec une bande passante plus grande où l'écart entre deux sous-porteuses sélectionnées augmente, impliquant moins de corrélation entre elles et par la suite entre les bits générés. Les coefficients du canal doivent être séparés par la bande de cohérence pour garantir leur décorrélation. En outre, le mean pass rate augmente avec la distance de séparation entre Alice et Bob, ce qui est expliqué dans la Figure 21 par le fait que l'étalement temporel augmente avec la distance et par la suite la bande de cohérence diminue. De plus, on remarque que la condition NLOS assure plus d'aléa que LOS puisque la première implique des canaux plus riches en multi-trajets.

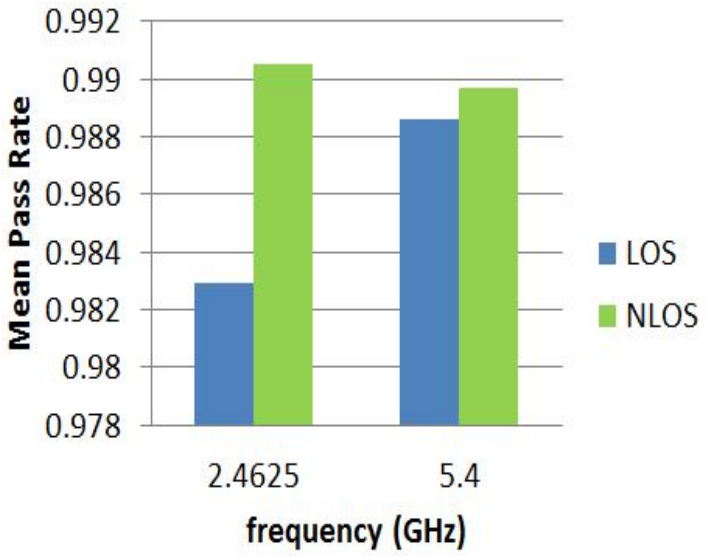


Figure 19 Evaluation de l'aléa fourni par le degré de liberté spatial

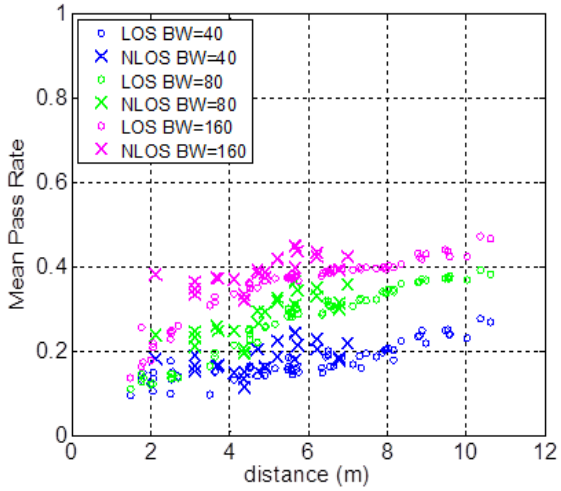


Figure 20 Evaluation de l'aléa fourni par le degré de liberté fréquentiel pour *M=4*

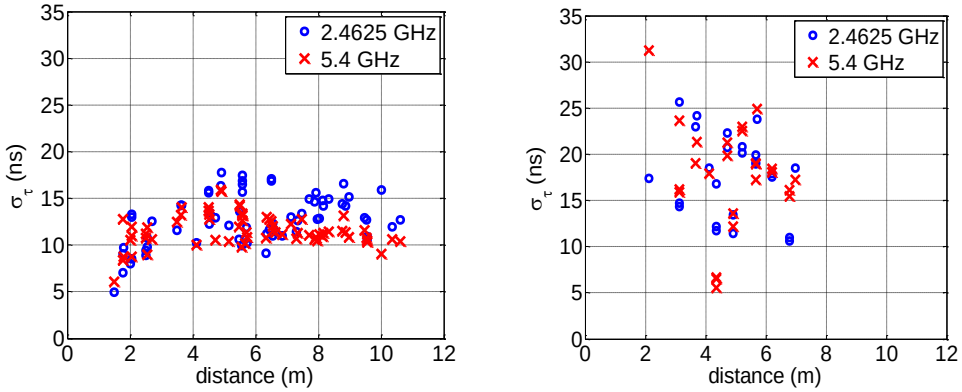


Figure 21 $\boldsymbol{\sigma_\tau}$ en fonction de la distance Alice-Bob pour LOS (gauche) et NLOS (droite)

En environnement intérieur, le degré de liberté fréquentiel fournit moins d'aléa que le spatial (Figure 22). En d'autres termes, la source d'aléa la plus adaptée à la SKG provient du degré de liberté spatial, alors que la moins bonne résulte de l'exploitation du degré de liberté fréquentiel. Comme compromis, l'exploitation conjointe des degrés de liberté spatial et fréquentiel semble être efficace, tout en considérant un système large bande avec un réseau de deux ou quatre antennes (soit typiquement 2 ou 4 degrés de liberté spatiaux). L'augmentation du nombre d'antennes réduit la demande en bande passante.

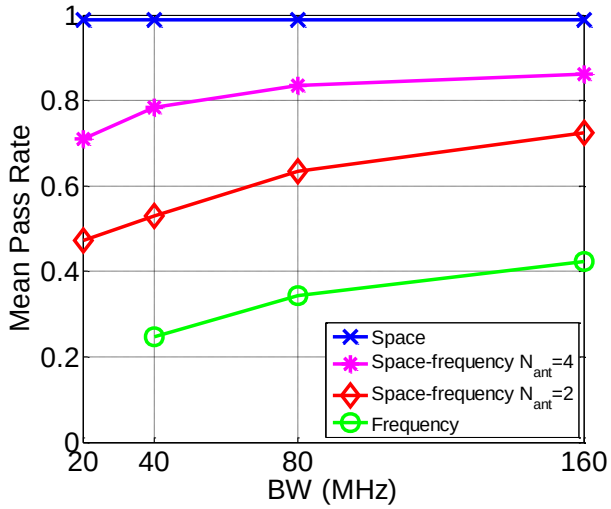


Figure 22 Comparaison de l'aléa fourni par les différents degrés de liberté

5.3. L'impact de N_f sur le nombre de bits disponibles I_K

En exploitant plus de degrés de liberté en fréquence, le nombre de bits qu'Alice et Bob peuvent partager croît de façon sous-linéaire (Figure **23**). Contrairement aux résultats déjà obtenus pour l'environnement extérieur (Figure **10**), l'évolution de I_K en fonction de N_f n'a pas tendance à saturer puisque la réponse impulsionnelle (Figure **24**) est continue et requiert une large bande pour que les trajets soient parfaitement résolus. Néanmoins, les canaux en intérieur sont moins riches en aléa que ceux en extérieur en raison du faible étalement temporel. Pour augmenter les valeurs de I_K, les sous-porteuses doivent être plus écartées en augmentant la bande passante.

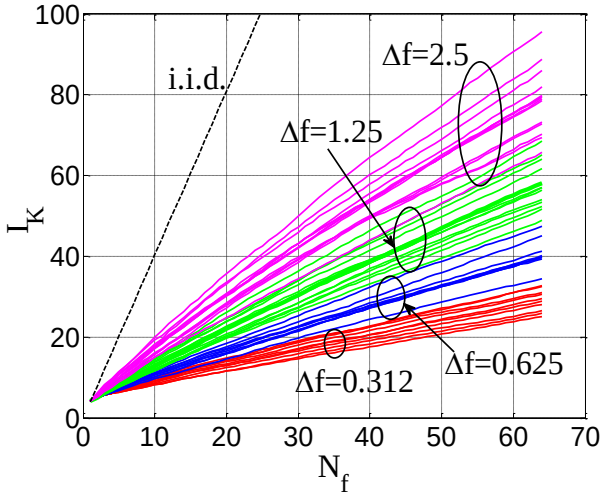


Figure 23 I_K en fonction de N_f

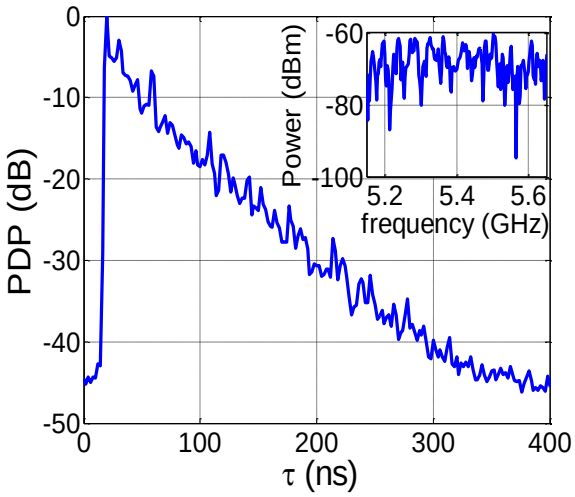


Figure 24 Un exemple de réponse impulsionnelle en environnement à l'intérieur

5.4.La stratégie complète de la SKG

Alice et Bob doivent parvenir à générer exactement la même clé (i.e. BER=0) pour qu'elle soit utilisée pour le cryptage des données. En revanche, les clés résultant de l'étape de quantification présentent quelques différences évaluées par le BER. Le rôle de la réconciliation est ainsi de rendre les clés d'Alice et de Bob complètement identiques tout en utilisant un code correcteur d'erreur. On choisit ainsi le code BCH pour mettre en œuvre cette étape. D'autre part, pour réduire l'information révélée à Eve et améliorer la qualité de la clé, on implémente l'étape d'amplification de confidentialité à travers des fonctions de hachage. Chaque erreur sera distribuée sur la clé entière grâce à des opérations de convolution, ce qui augmente la différence entre les clés d'Eve et des terminaux légitimes.

La Figure 25 montre que presque 95% des clés sont parfaitement réconciliées pour SNR=15 dB alors que seulement 12% le sont pour SNR=10 dB. On peut ainsi souligner qu'il faut bien choisir les paramètres de la réconciliation pour qu'ils soient adaptés au SNR. En outre, l'amplification de confidentialité augmente le BER des clés non réconciliées, ce qui révèle son utilité pour compliquer le travail de l'espion. Par ailleurs, cette dernière étape améliore la qualité de l'aléa des clés quel que soit le type de degré de liberté utilisé (Figure 26).

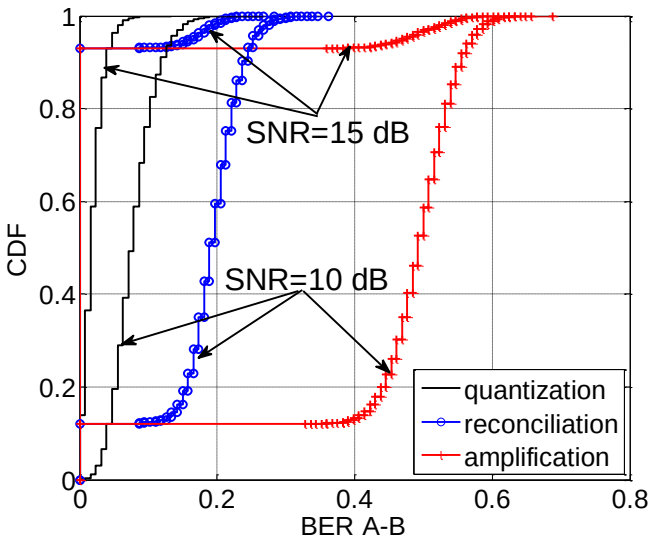


Figure 25 Le BER des clés légitimes durant les différentes étapes de la SKG

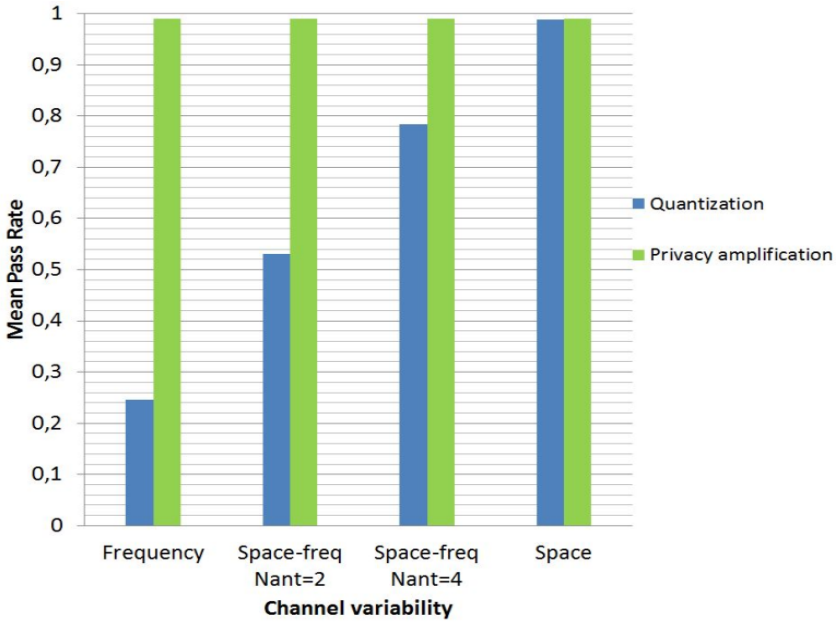


Figure 26 Le caractère aléatoire des clés durant les étapes de la SKG

6. Conclusion et perspectives

6.1.Conclusion

On a examiné dans cette thèse le processus de génération de clés secrètes à partir de l'aléa du canal de propagation. En particulier, l'analyse de performance a concerné l'impact de la propagation sur la qualité des clés, tout en considérant différents degrés de liberté du canal.

Le degré de liberté spatial est le plus adapté à la SKG quel que soit le type de l'environnement. Cela est conditionné par le fait que les antennes sont suffisamment espacées, par exemple $\lambda/2$. En outre, le degré de liberté fréquentiel ne fournit pas suffisamment d'aléa en environnement intérieur, au contraire d'en environnement extérieur. En effet, la bande de

cohérence est plus large dans le premier que le second. Toutefois, en exploitant conjointement le domaine spatial (par exemple avec 2 ou 4 antennes) et fréquentiel (avec une certaine bande), des clés robustes peuvent êtres extraites du canal.

D'une manière générale, il est intéressant d'augmenter le nombre de sous-canaux, même lorsqu'ils sont corrélés, puisque cela augmente légèrement le taux de clé secrète. Cependant, cela signifie également de concevoir des systèmes intelligents, capables de produire de longues clés de bits indépendants à partir de très longues clés de bits corrélés. Cela peut être accompli par l'amplification de confidentialité. Néanmoins, la qualité de clé est améliorée de façon remarquable tout en implémentant la stratégie complète de la SKG, i.e. la réconciliation et l'amplification de confidentialité. Les paramètres utilisés doivent être adaptés au SNR.

6.2.Perspectives

Il existe plusieurs perspectives au travail réalisé dans cette thèse.

- Il est intéressant de calculer la longueur théorique des clés, i.e. l'information mutuelle, pour n'importe quel type du canal. Cela permet d'évaluer l'efficacité de la SKG dans l'environnement considéré
- L'antenne est un élément du canal de transmission et elle peut affecter la performance de la SKG, d'où la nécessité de considérer et d'étudier son influence sur la qualité des clés générées
- En pratique, le canal n'est pas parfaitement réciproque puisque d'une part, les terminaux ne l'estiment pas en même instant et d'autre part, dans un système d'émission-réception les signaux émis ne suivent pas le même chemin que les signaux reçus. Par conséquent, il est important de considérer les sources de non réciprocité
- La SKG repose sur une sécurité du point de vue de la théorie de l'information où seule l'information peut aider l'espion à affaiblir la confidentialité des données. De ce fait, Eve peut investir des techniques lui permettant d'acquérir plus d'information ainsi que plus de corrélation avec le canal légitime. Dans ce but, Eve peut utiliser des outils puissants de calcul de la propagation pour améliorer sa connaissance du canal Bob-Eve, de type tracé de rayons. Néanmoins, la complexité de la physique de la propagation, par exemple la diffusion diffuse, rendent une telle hypothèse très peu vraisemblable ?

List of Publications

Journal papers

[J.1] T. Mazloum and A. Sibille, *Analysis of secret key randomness exploiting the radio channel variability*, International Journal of Antennas and Propagation, Septembre 2015.

Conference papers

[C.1] T. Mazloum, F. Mani, and A. Sibille, *A disc of scatterers based radio channel model for secure key generation*, EUCAP, The Hague, Netherlands, pp. 1290-1294, April 6-11, 2014.

[C.2] E. Nicollet, F. Delaveau, R. Molière, C.L.K. Ngassa, C. Lemenager, T. Mazloum, and A. Sibille, *Towards a key-free radio protocol for authentication and security of nodes and terminals in advanced waveforms*, WinnComm, San Diego, 26 March, 2015.

[C.3] T. Mazloum and A. Sibille, *Performance of secret key generation in non stationary channels*, EUCAP, Lisbon, Portugal, April 13-17, 2015.

[C.4] Vitucci, F. Mani, T. Mazloum, A. Sibille, and V.E. Espotti, *Ray tracing simulations of indoor channel spatial correlation for physical layer security*, EUCAP, Lisbon, Portugal, April 13-17, 2015.

[C.5] T. Mazloum, F. Mani, and A. Sibille, *Analysis of secret key robustness in indoor radio channel measurements*, VTC-spring, Glasgow, Scotland, May 11-14, 2015.

[C.6] R. Molière, F. Delaveau, C.L.K. Ngassa, C. Lemenager, T. Mazloum, and A. Sibille, *Tag signals for early authentication and secret key generation in wireless public networks*, EuCNC, Paris, France, June 29-July 2, 2015.

[C.7] C. Kameni-Ngassa, M. Renaud, F. Delaveau, T. Mazloum, and A. Sibille, *Secret key generation scheme from WIFI and LTE reference signals*, WinnComm 2016, Reston, Virginia, March 15-17, 2016.

COST meeting presentations

[TD.1] T. Mazloum, F. Mani, and A. Sibille, *A disc of scatterers based radio channel model for secure key generation*, TD(14) 09017, COST IC1004, Sep. 25-27 2013, Ghent, Belgium.

[TD.2] F. Mani, E. Vitucci, T. Mazloum, A. Sibille, , and V.D. Esposti, *Ray tracing simulations of indoor channel spatial correlation for physical layer security*, TD(13) 08072, COST IC1004, Feb. 5-7 2014, Ferrara, Italy.

References

[1] P. Kyosti et al., "Winner ii channel model." https://www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf, Sept. 2007.

[2] ZEIT, "Wie merkels handy abgehrt werden konnte." http://www.zeit.de/digital/datenschutz/2014-12/umts-verschluesselung-umgehen-angela-merkel-handy, Dec. 2014.

[3] A. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, pp. 1355–1367, Oct. 1975.

[4] M. Bloch and J. Barros, *Physical-layer security: From information theory to security engineering*. Cambridge University press, 2011.

[5] W. Stallings, *Cryptography and Network Security: Principles and Practice*. Pearson Prentice Hall, 3rd ed., 2005.

[6] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Trans. Inform. Theory*, vol. 22, pp. 644–654, Nov 1976.

[7] M. Bloch, J. Barros, M. Rodrigues, and S. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inform. Theory*, vol. 54, pp. 2515–2534, June 2008.

[8] A. Thangaraj, S. Dihidar, A. Calderbank, S. McLaughlin, and J. Merolla, "Applications of ldpc codes to the wiretap channel," *IEEE Trans. Inform. Theory*, vol. 53, pp. 2933–2945, Aug. 2007.

[9] X. H. He and A. Yener, "Providing secrecy with lattice codes," in *in Proc. 46th Annual Allerton Conf. on Communication, Control, and Computing*, pp. 1199–1206, Sept 2008.

[10] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inform. Theory*, vol. 57, pp. 6428–6443, Oct 2011.

[11] N. Romero-Zurita, M. Ghogho, and D. McLernon, "Physical layer security of mimoofdm systems by beamforming and artificial noise generation," *Physical Communication on Advances in MIMO-OFDM*, vol. 4, no. 4, pp. 313–321, 2011.

[12] Y.-W. Hong, P.-C. Lan, and C.-C. Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Processing Magazine*, vol. 30, pp. 29–40, Sept 2013.

[13] R. Negi and S. Goel, "Secret communication using artificial noise," in *in Proc. 62nd IEEE Vehicular Technology Conf. (VTC)*, pp. 1906–1910, Sept 2005.

[14] N. Anand, S.-J. Lee, and E. Knightly, "Strobe: Actively securing wireless communications using zero-forcing beamforming," in *in Proc. 2012 IEEE INFOCOM*, pp. 720–728, March 2012.

[15] "Phylaws." www.phylaws-ict.org.

[16] C. Shannon, "Wireless information-theoretic security," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.

[17] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. on Inform. Theory*, vol. 39, pp. 733–742, May 1993.

[18] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1121–1132, Jul 1993.

[19] C. Bennet and G. Brassard, "Quantum cryptography: public key distribution and coin tossing," in *in Int. conf. on computers, systems and signal processing*, pp. 175–179, Dec 1984.

[20] A. Ekert, "Quantum cryptography based on bell's theorem," *Physical review letters*, vol. 67, Aug 1991.

[21] A. Hassan, W. Stark, J. Hershey, and S. Chennakeshu, "Cryptographic key agreement for mobile radio," *Digital Signal Processing*, vol. 6, pp. 207–212, Oct 1996.

[22] J. Hershey, A. Hassan, and R. Yarlagadda, "Unconventional cryptographic keying variable management," *IEEE Trans. Communications*, vol. 43, pp. 3–6, Jan 1995.

[23] T. Mazloum, F. Mani, and A. Sibille, "Analysis of secret key robustness in indoor radio channel measurements," in *Proc. 2015 IEEE 81st Vehicular Technology Conference (VTC-Spring)*, May 2015.

[24] T. Mazloum and A. Sibille, "Analysis of secret key randomness exploiting the radio channel variability," *International Journal of Antennas and Propagation (IJAP)*, 2015.

[25] T. Mazloum and A. Sibille, "Performance of secret key generation in non stationary channels," in *Proc. 2015 9th Europ. Conf. Antennas and Propagation (EuCAP'15)*, April 2015.

[26] J. W. Strutt, *The theory of sound*, vol. 2. Macmillan, 1896.

[27] H. Lorentz, "The theorem of poynting concerning the energy in the electromagnetic field and two general propositions concerning the propagation of light," *Amsterdammer Akademie der Wetenschappen*, vol. 4, p. 176, 1896.

[28] J. Guey and L. Krasny, "Transceiver design and performance evaluation of mimo systems with forward link channel knowledge," in *Proc. 60th IEEE Vehicular Technology Conf. (VTC-Fall)*, vol. 2, pp. 1386–1390, Sept 2004.

[29] S. Premnath, S. Jana, J. Croft, P. Gowda, M. Clark, S. Kasera, N. Patwari, and S. Krishnamurthy, "Secret key extraction from wireless signal strength in real environments," *IEEE Trans. on Mobile Computing*, vol. 12, pp. 917–930, May 2013.

[30] S. Hamida, J. Pierrot, and C. Castelluccia, "Empirical analysis of uwb channel characteristics for secret key generation in indoor environments," in *Proc. 2010 IEEE 21st Int. Symp. Personal Indoor and Mobile Radio Communications (PIMRC'10)*, pp. 1984–1989, Sept 2010.

[31] A. Bourdoux, B. Come, and N. Khaled, "Non-reciprocal transceivers in ofdm/sdma systems: impact and mitigation," in *Proc. Radio and Wireless Conference (RAWCON'03)*, pp. 183–186, Aug 2003.

[32] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in mimo/tdd systems," in *Proc. 2010 Future Network and Mobile Summit*, pp. 1–10, June 2010.

[33] M. Petermann, M. Stefer, D. Wubben, M. Schneider, and K. Kammeyer, "Low-complexity calibration of mutually coupled non-reciprocal multi-antenna ofdm transceivers," in *7th IEEE Intl. Symp. on Wireless Commun. Sys. (ISWCS)*, p. 285289, 19-22 Sep. 2010.

[34] A. Mahmood and M. Jensen, "Assessing and removing the impact of non-reciprocal transceiver circuitry for channel-based key establishment," in *Proc. 9th European Conf. on Antennas and Propagation (EuCAP)*, April 2015.

[35] Y. Han, J. Ni, and G. Du, "The potential approaches to achieve channel reciprocity in fdd system with frequency correction algorithms," in *Proc. 5th Int. ICST Conf. on Communications and Networking in China (CHINACOM)*, pp. 1–5, Aug 2010.

[36] A. Molisch, *Wireless Communications*. Chichester, West Sussex, UK: IEEE Press-Wiley, 2005.

[37] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, pp. 2145–2146, Nov 1991.

[38] B. Fleury, "First- and second-order characterization of direction dispersion and space selectivity in the radio channel," *IEEE Trans. Information Theory*, vol. 46, pp. 2027–2044, Sep 2000.

[39] G. Durgin and T. Rappaport, "Basic relationship between multipath angular spread and narrowband fading in wireless channels," *Electronics Letters*, vol. 34, pp. 2431–2432, Dec 1998.

[40] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, and A. Paulraj, "Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels," *IEEE Trans. Vehicular Technology*, vol. 59, pp. 2253–2265, Jun 2010.

[41] P. Almers et al., "Survey of channel and radio propagation models for wireless mimo systems," *EURASIP Journal Wireless Communications and Networking*, vol. 2007, Jan 2007.

[42] S. Hamida, J. Pierrot, B. Denis, C. Castelluccia, and B. Uguen, "On the security of uwb secret key generation methods against deterministic channel prediction attacks," in *Proc 2012 IEEE Vehicular Tech. Conf. (VTC-Fall)*, pp. 1–5, Sept 2012.

[43] E. Vitucci, F. Mani, T. Mazloum, A. Sibille, and V. Degli Esposti, "Ray tracing simulations of indoor channel spatial correlation for physical layer security," in *Proc. 2019 9th Europ. Conf. Antennas and Propagation (EuCAP'15)*, pp. 1–5, April 2015.

[44] V. Degli-Esposti, D. Guiducci, A. de'Marsi, P. Azzi, and F. Fuschini, "An advanced field prediction model including diffuse scattering," *IEEE Trans. Antennas and Propagation*, vol. 52, no. 7, pp. 1717–1728, 2004.

[45] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. Doncker, "The cost 2100 mimo channel model," *IEEE Wireless Communications*, vol. 19, pp. 92–99, December 2012.

[46] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, pp. 128–137, February 1987.

[47] J. Wallace and M. Jensen, "Modeling the indoor mimo wireless channel," *IEEE Trans. Antennas and Propagation*, vol. 50, pp. 591–599, May 2002.

[48] H. Ozcelik, N. Czink, and E. Bonek, "What makes a good mimo channel model?," in *Proc. 2005 IEEE 61st Vehicular Technology Conf.*, vol. 1, pp. 156–160, May 2005.

[49] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic mimo channel model with joint correlation of both link ends," *IEEE Trans. Wireless Communications*, vol. 5, pp. 90–100, Jan 2006.

[50] M. Patzold, B. Hogstad, and N. Youssef, "Modeling, analysis, and simulation of mimo mobile-to-mobile fading channels," *IEEE Trans. Wireless Communications*, vol. 7, pp. 510–520, February 2008.

[51] A. Borhani and M. Patzold, "A non-stationary one-ring scattering model," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, pp. 2620–2625, April 2013.

[52] J. Karedal, F. Tufvesson, N. Czink, A. Paier, C. Dumard, T. Zemen, C. Mecklenbrauker, and A. Molisch, "A geometry-based stochastic mimo model for vehicle-to-vehicle communications," *IEEE Trans. Wireless Communications*, vol. 8, pp. 3646–3657, July 2009.

[53] O. Renaudin, V. Kolmonen, P. Vainikainen, and C. Oestges, "Non-stationary narrowband mimo inter-vehicle channel characterization in the 5-ghz band," *IEEE Trans. Vehicular Technology*, vol. 59, pp. 2007–2015, May 2010.

[54] F. Graziosi and F. Santucci, "A general correlation model for shadow fading in mobile radio systems," *IEEE Communications Letters*, vol. 6, pp. 102–104, March 2002.

[55] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka, "Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels," *IEEE Trans. on Antennas and Propagation*, vol. 53, pp. 3776–3784, Nov 2005.

[56] G. Tsouri and J. Wilczewski, "Reliable symmetric key generation for body area networks using wireless physical layer security in the presence of an on-body eavesdropper," in *Proc. 4th Int. Symp. on Applied Sciences in Biomedical and Communication Technologies*, ISABEL'11, pp. 1–6, 2011.

[57] K. Zeng, D. Wu, A. Chan, and P. Mohapatra, "Exploiting multiple-antenna diversity for shared secret key generation in wireless networks," in *INFOCOM, 2010 Proceedings IEEE*, pp. 1–9, March 2010.

[58] J. Wallace and R. Sharma, "Automatic secret keys from reciprocal mimo wireless channels: Measurement and analysis," *IEEE Trans. Inform. Forensics and Security*, vol. 5, pp. 381–392, Sept 2010.

[59] J. Croft, N. Patwari, and S. Kasera, "Robust uncorrelated bit extraction methodologies for wireless sensors," in *Proc. 9th ACM/IEEE Int. Conf. on Information Processing in Sensor Networks*, IPSN'10, pp. 70–81, 2010.

[60] S. Mathur, W. Trappe, N. Mandayan, C. Ye, and A. Reznik, "Radio-telepathy: extracting a secret key from an unauthenticated wireless channel," in *Proc. MobiCom*, pp. 128–139, Sep 2008.

[61] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. Mandayam, "Information-theoretically secret key generation for fading wireless channels," *IEEE Trans. Information Forensics and Security*, vol. 5, pp. 240–254, June 2010.

[62] C. Chen and M. Jensen, "Secret key establishment using temporally and spatially correlated wireless channel coefficients," *IEEE Trans. Mobile Computing*, vol. 10, pp. 205–215, Feb 2011.

[63] R. Wilson, D. Tse, and R. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *IEEE Trans. Inform. Forensics and Security*, vol. 2, pp. 364–375, Sept 2007.

[64] M. Madiseh, S. He, M. McGuire, S. Neville, and X. Dong, "Verification of secret key generation from uwb channel observations," in *Proc. 2009 IEEE Int. Conf. Communications (ICC'09)*, pp. 1–5, June 2009.

[65] Y. Liu, S. Draper, and A. Sayeed, "Exploiting channel diversity in secret key generation from multipath fading randomness," *IEEE Trans. Information Forensics and Security*, vol. 7, pp. 1484–1497, Oct 2012.

[66] R. Chou, M. Bloch, and E. Abbe, "Polar coding for secret-key generation," in *Proc. 2013 IEEE Information Theory Workshop (ITW)*, pp. 1–5, Sept 2013.

[67] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels .ii. privacy amplification," *IEEE Trans. Information Theory*, vol. 49, pp. 839–851, April 2003.

[68] A. Sibille, "Analysis of alice-bob-eve scenarios for secret key generation from random channels," in *Proc. General Assembly and Scientific Symposium (URSI GASS)*, pp. 1–4, Aug 2014.

[69] R. Moliere, D. Delaveau, C. Ngassa, C. Lemenager, T. Mazloum, and A. Sibille, "Tag signals for early authentication and secret key generation in wireless public networks," in *European Conf. Networks Communications (EuCNC)*, July 2015.

[70] F. Delaveau, A. Evestti, A. Kotelba, R. Savola, and N. Shapira, "Active and passive eavesdropper threats within public and private cililian networks - existing and potential future countermeasures - an overview," in *WinnComm 2013*, 2013.

[71] C. Ye, A. Reznik, and Y. Shah, "Extracting secrecy from jointly gaussian random variables," in *Proc. 2006 IEEE Int. Symp. Inform. Theory*, pp. 2593–2597, July 2006.

[72] T. Mazloum, F. Mani, and A. Sibille, "A disc of scatterers based radio channel model for secure key generation," in *Proc. 2014 8th Europ. Conf. Antennas and Propagation (EuCAP'14)*, pp. 1290–1294, April 2014.

[73] M. Madiseh, M. McGuire, S. Neville, and A. Shirazi, "Secret key extraction in ultra wideband channels for unsynchronized radios," in *Proc. 6th Annual*

Comm. Networks and Services Research Conf. (CNSR'08), pp. 88–95, May 2008.

[74] A. Shahzadi, M. Madiseh, and A. Shirazi, "Secret key capacity for wireless nakagami and suzuki fading channels," *Int. Journal of Computer Science and Network Security (IJCSNS)*, vol. 7, no. 3, pp. 296–303, 2007.

[75] I. Tunaru, B. Denis, and B. Uguen, "Random patterns of secret keys from sampled ir-uwb channel responses," in *Proc. IEEE Int. Conf. Ultra-WideBand (ICUWB)*, pp. 74–79, Sept 2014.

[76] A. Sayeed and A. Perrig, "Secure wireless communications: Secret keys through multipath," in *Proc. 2008 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3013–3016, March 2008.

[77] R. Sharma and J. Wallace, "Bit error rate and efficiency analysis of wireless reciprocal channel key generation," in *Proc. 2010 IEEE Int. Conf. Wireless Information Technology and Systems (ICWITS)*, pp. 1–4, Aug 2010.

[78] S. Hamida, J. Pierrot, and C. Castelluccia, "An adaptive quantization algorithm for secret key generation using radio channel measurements," in *Proc. 2009 3rd Int. Conf. New Technologies, Mobility and Security (NTMS'09)*, pp. 1–5, Dec 2009.

[79] A. Pierrot, R. Chou, and M. Bloch, "Experimental aspects of secret key generation in indoor wireless environments," in *Proc. 2013 IEEE 14th Workshop Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 669–673, June 2013.

[80] T. Ohira, "Secret key generation exploiting antenna beam steering and wave propagation reciprocity," in *European Microwave Conference*, vol. 1, pp. 9–12, Oct 2005.

[81] R. Mehmood and J. Wallace, "Experimental assessment of secret key generation using parasitic reconfigurable aperture antennas," in *Proc. 6th European Conf. Antennas and Propagation (EUCAP)*, pp. 1151–1155, March 2012.

[82] P. Huang and X. Wang, "Fast secret key generation in static wireless networks: A virtual channel approach," in *Proc. IEEE INFOCOM*, pp. 2292–2300, April 2013.

[83] L. Kang, S. Primak, and W. Xianbin, "On secret key generation from multiple observations of wireless channels," in *Proc. 2014 IEEE Int. Conf. on Communication Systems (ICCS)*, pp. 147–151, 2014.

[84] S. Primak, K. Liu, and X. Wang, "Secret key generation using physical channels with imperfect csi," in *Proc. 2014 IEEE 80th Vehicular Technology Conf. (VTC-Fall)*, pp. 1–5, Sept 2014.

[85] G. Pasolini and D. Dardari, "Secret information of wireless multi-dimensional gaussian channels," *IEEE Trans. Wireless Communications*, vol. 14, pp. 3429–3442, June 2015.

[86] G. Tsouri and D. Wagner, "Threshold constraints on symmetric key extraction from rician fading estimates," *IEEE Trans. on Mobile Computing*, vol. 12, pp. 2496–2506, Dec 2013.

[87] K. Liu, S. Primak, and X. Wang, "On secret key generation from multiple observations of wireless channels," in *Proc. 2014 IEEE Int. Conf. on Communication Systems (ICCS)*, pp. 147–151, Nov 2014.

[88] S. Vaudenay, "Secure communications over insecure channels based on short authenticated strings," in *Advances in Cryptology CRYPTO 2005* (V. Shoup, ed.), vol. 3621 of *Lecture Notes in Computer Science*, pp. 309–326, 2005.

[89] J. Wallace, "Secure physical layer key generation schemes: performance and information theoretic limits," in *Proc. 2009 IEEE Int. Conf. Communications (ICC'09)*, pp. 1–5, June 2009.

[90] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[91] M. Madiseh, S. Neville, and M. McGuire, "Time correlation analysis of secret key generation via uwb channels," in *Proc. 2010 IEEE Global Telecommunications Conference (GLOBECOM 2010)*, pp. 1–6, Dec 2010.

[92] X. Zhu, F. Xu, E. Novak, C. Tan, Q. Li, and G. Chen, "Extracting secret key from wireless link dynamics in vehicular environments," in *Proc 2013 IEEE INFOCOM*, pp. 2283–2291, April 2013.

[93] W. C. Jakes and D. C. Cox, eds., *Microwave Mobile Communications*. Wiley-IEEE Press, 1994.

[94] W. Burr, D. Dodson, and W. Polk, *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*. Information Technology Laboratory, NIST, Gaithersburg, Maryland, Tech. Rep., 2010.

[95] H. Koorapaty, A. Hassan, and S. Chennakeshu, "Secure information transmission for mobile radio," *IEEE Communications Letters*, vol. 4, pp. 52–55, Feb 2000.

[96] Q. Wang, H. Su, K. Ren, and K. Kim, "Fast and scalable secret key generation exploiting channel phase randomness in wireless networks," in *Proc. 2011 IEEE INFOCOM*, pp. 1422–1430, April 2011.

[97] Y. El Hajj Shehadeh, O. Alfandi, and D. Hogrefe, "Towards robust key extraction from multipath wireless channels," *Journal of Communications and Networks*, vol. 14, pp. 385–395, Aug 2012.

[98] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive mimo performance evaluation based on measured propagation data," *IEEE Trans. Wireless Communications*, vol. 14, pp. 3899–3911, July 2015.

[99] K. Pedersen, P. Mogensen, and B. Fleury, "Spatial channel characteristics in outdoor environments and their impact on bs antenna system performance," in *Proc. 48th IEEE Vehicular Technology Conf. (VTC'98)*, vol. 2, pp. 719–723 vol.2, May 1998.

[100] J. Andersen, J. Nielsen, G. Pedersen, G. Bauch, and M. Herdin, "Room electromagnetics," *IEEE Antennas and Propagation Magazine*, vol. 49, pp. 27–33, April 2007.

[101] A. Richter, "The contribution of distributed scattering in radio channels to channel capacity: Estimation and modeling," in *Proc. 4th Conf. Signals, Systems and Computers (ACSSC'06)*, pp. 951–955, Oct 2006.

[102] J. Kunisch and J. Pamp, "Measurement results and modeling aspects for the uwb radio channel," in *Ultra Wideband Systems and Technologies (UWBST)*, pp. 19–23, May 2002.

[103] V. Erceg and et al., "Tgn channel models," tech. rep., IEEE 802.11-03/940r4, IEEE P802.11 Wireless LANs, 2004.

[104] G. Breit and et al., "Tgac channel model addendum," tech. rep., IEEE 802.11-09/0308r12, IEEE P802.11 Wireless LANs, 2010.

[105] R. Saadane, A. Menouni, R. Knopp, and D. Aboutajdine, "Empirical eigenanalysis of indoor uwb propagation channels," in *Global Telecommunications Conf. (GLOBECOM'04)*, vol. 5, pp. 3215–3219, Nov 2004.

[106] "Deliverable 3.1 - channel based random generators." www.phylaws-ict.org.

[107] K. Borner, J. Dommel, S. Jaeckel, and L. Thiele, "On the requirements for quasi-deterministic radio channel models for heterogeneous networks," in *Proc. IEEE Int. Symposium Signals, Systems, and Electronics (ISSSE)*, pp. 1–5, Oct 2012.

[108] A. Molisch, "A generic model for mimo wireless propagation channels in macro- and microcells," *IEEE Trans. Signal Processing*, vol. 52, pp. 61–71, Jan 2004.

[109] H. Asplund, A. Glazunov, A. Molisch, K. Pedersen, and M. Steinbauer, "The cost 259 directional channel model-part ii: Macrocells," *IEEE Trans. Wireless Communications*, vol. 5, pp. 3434–3450, December 2006.

[110] L. Vuokko, V.-M. Kolmonen, J. Salo, and P. Vainikainen, "Measurement of large-scale cluster power characteristics for geometric channel models," *IEEE Trans. Antennas and Propagation*, vol. 55, pp. 3361–3365, Nov 2007.

[111] S. Szyszkowicz, H. Yanikomeroglu, and J. Thompson, "On the feasibility of wireless shadowing correlation models," *IEEE Trans. Vehicular Technology*, vol. 59, pp. 4222–4236, Nov 2010.

[112] F. Graziosi, M. Pratesi, M. Ruggieri, and F. Santucci, "A multicell model of handover initiation in mobile cellular networks," *IEEE Trans. Vehicular Technology*, vol. 48, pp. 802–814, May 1999.

[113] A. Borhani and M. Patzold, "A unified disk scattering model and its angle-of-departure and time-of-arrival statistics," *IEEE Trans. Vehicular Tech.*, vol. 62, pp. 473–485, Feb. 2013.

[114] A. Borhani and M. Patzold, "Modelling of non-stationary mobile radio channels using two-dimensional brownian motion processes," in *Proc. 2013 Int. Conf. Advanced Technologies for Communications (ATC)*, pp. 241–246, Oct 2013.

[115] O. Renaudin, V.-M. Kolmonen, P. Vainikainen, and C. Oestges, "Wideband measurement-based modeling of inter-vehicle channels in the 5-ghz band," *IEEE Trans. Vehicular Technology*, vol. 62, pp. 3531–3540, Oct 2013.

[116] F. Amoroso and W. Jones, "Geometric model for dspn satellite reception in the dense scatterer mobile environment," *IEEE Trans. Comm.*, vol. 41, pp. 450–453, Mar 1993.

[117] "Volcano lab." http://www.siradel.com.

[118] Y. Corre and Y. Lostanlen, "Three-dimensional urban em wave propagation model for radio network planning and optimization over large areas," *IEEE Trans. Vehicular Technology*, vol. 58, pp. 3112–3123, Sept 2009.

[119] V. Degli-Esposti, F. Fuschini, E. Vitucci, and G. Falciasecca, "Measurement and modelling of scattering from buildings," *IEEE Trans. Antennas and Propagation*, vol. 55, pp. 143–153, Jan 2007.

[120] Y. Lostanlen and G. Gougeon, "Introduction of diffuse scattering to enhance ray-tracing methods for the analysis of deterministic indoor uwb radio channels (invited paper)," in *Int. Conf. Electromagnetics in Advanced Applications (ICEAA)*, pp. 903–906, Sept 2007.

[121] D. Cox and R. Leck, "Distributions of multipath delay spread and average excess delay for 910-mhz urban mobile radio paths," *IEEE Trans. Antennas and Propagation*, vol. 23, pp. 206–213, March 1975.

[122] H. Ghannoum, S. Bories, C. Roblin, and A. Sibille, "Biconical antennas for intrinsic characterization of the uwb channel," in *Proc. 2005 IEEE Int. Workshop on Antenna Technology: Small Antennas and Novel Metamaterials (IWAT)*, pp. 101–104, March 2005.

[123] H. Hashemi and D. Tholl, "Statistical modeling and simulation of the rms delay spread of indoor radio propagation channels," *IEEE Trans. Vehicular Technology*, vol. 43, pp. 110–120, Feb. 1994.

[124] T. Jamsa, V. Hovinen, A. Karjalainen, and J. Iinatti, "Frequency dependency of delay spread and path loss in indoor ultra-wideband channels," in *Ultra Wideband Systems, Technologies and Applications, 2006. The Institution of Engineering and Technology Seminar on*, pp. 254–258, April 2006.

[125] S. Geng and P. Vainikainen, "Frequency and bandwidth dependency of uwb propagation channels," in *IEEE 17th Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–5, Sept. 2006.

[126] C. Kameni-Ngassa, M. Renaud, F. Delaveau, T. Mazloum, and A. Sibille, "Secret key generation scheme from wifi and lte reference signals," in *WinnComm 2016 (to be published)*, 2016.

[127] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *Advances in Cryptology - EUROCRYPT 2004* (C. Cachin and J. Camenisch, eds.), vol. 3027 of *Lecture Notes in Computer Science*, pp. 523–540, Springer Berlin Heidelberg, 2004.

[128] C. Bennett, G. Brassard, C. Crepeau, and U. Maurer, "Generalized privacy amplification," *IEEE Trans. Information Theory*, vol. 41, pp. 1915–1923, Nov 1995.

[129] Y. Zou, O. Raeesi, and M. Valkama, "Efficient estimation and compensation of transceiver non-reciprocity in precoded tdd multi-user mimo-ofdm systems," in *Proc. 2014 IEEE 80th Vehicular Technology Conf. (VTC-fall)*, pp. 1–7, Sept 2014.

[130] IEEE, "Ieee-802.11-2012: Part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications," 2012.

Appendices

Appendix A

Information theoretic bounds

For zero-mean jointly complex Gaussian channels, the capacity of available key bits is expressed as follows:

$$\begin{aligned} I_K &= I(\hat{\mathbf{h}}_a, \hat{\mathbf{h}}_b) \\ &= \log_2 \frac{|\hat{\mathbf{R}}_{aa}||\hat{\mathbf{R}}_{bb}|}{|\hat{\mathbf{R}}_{AB}|} \end{aligned} \tag{A.1}$$

Straightforward evaluation of the estimated covariance matrices reveals:

$$\hat{\mathbf{R}}_{aa} = E\{\hat{\mathbf{h}}_a \hat{\mathbf{h}}_a^H\} = \mathbf{R}_{aa} + \sigma_a^2 \mathbf{I} \tag{A.2}$$

$$\hat{\mathbf{R}}_{bb} = E\{\hat{\mathbf{h}}_b \hat{\mathbf{h}}_b^H\} = \mathbf{R}_{bb} + \sigma_b^2 \mathbf{I} \tag{A.3}$$

$$\hat{\mathbf{R}}_{AB} = E\{[\hat{\mathbf{h}_a}^H \hat{\mathbf{h}}_b^H]^H [\hat{\mathbf{h}}_a^H \hat{\mathbf{h}}_b^H]\} = \begin{bmatrix} \mathbf{R}_{aa} + \sigma_a^2 \mathbf{I} & \mathbf{R}_{ab} \\ \mathbf{R}_{ab}^H & \mathbf{R}_{bb} + \sigma_b^2 \mathbf{I} \end{bmatrix} \tag{A.4}$$

with

$$\mathbf{R}_{xy} = E\{\mathbf{h}_x \mathbf{h}_y^H\} \tag{A.5}$$

We recall that $\mathbf{h}_x$ is a stacked channel vector with N_h elements, i.e. parallel sub-channels, assuming all the terminals having the same number N_h. Accordingly, $\mathbf{I}$ is an $N_h \times N_h$ identity matrix, while $\{.\}^H$ denotes for matrix conjugate transpose (Hermitian). Moreover, we designate by σ_x^2 the average noise power at node x, with x being referring to Alice (a or A), Bob (b or B) and Eve (e or E).

Similarly, the secret key capacity is expressed as follows:

$$\begin{aligned} I_{SK} &= I(\hat{\mathbf{h}}_a, \hat{\mathbf{h}}_b | \hat{\mathbf{h}}_e) \\ &= \log_2 \frac{|\hat{\mathbf{R}}_{AE}||\hat{\mathbf{R}}_{BE}|}{|\hat{\mathbf{R}}_E||\hat{\mathbf{R}}_{ABE}|} \end{aligned} \tag{A.6}$$

while the estimated covariance matrices are explicitly evaluated as:

$$\hat{\mathbf{R}}_{AE} = \begin{bmatrix} \mathbf{R}_{aa} + \sigma_a^2 \mathbf{I} & \mathbf{R}_{ae} \\ \mathbf{R}_{ae}^H & \mathbf{R}_{ee} + \sigma_e^2 \mathbf{I} \end{bmatrix} \tag{A.7}$$

$$\hat{\mathbf{R}}_{BE} = \begin{bmatrix} \mathbf{R}_{aa} + \sigma_a^2 \mathbf{I} & \mathbf{R}_{ae} \\ \mathbf{R}_{ae}^H & \mathbf{R}_{ee} + \sigma_e^2 \mathbf{I} \end{bmatrix} \tag{A.8}$$

$$\hat{\mathbf{R}}_E = \hat{\mathbf{R}}_{ee} = \mathbf{R}_{ee} + \sigma_e^2 \mathbf{I} \tag{A.9}$$

$$\hat{\mathbf{R}}_{ABE} = \begin{bmatrix} \mathbf{R}_{aa} + \sigma_a^2 \mathbf{I} & \mathbf{R}_{aa} & \mathbf{R}_{ae} \\ \mathbf{R}_{aa} & \mathbf{R}_{aa} + \sigma_b^2 \mathbf{I} & \mathbf{R}_{ae} \\ \mathbf{R}_{ae}^H & \mathbf{R}_{ae}^H & \mathbf{R}_{ee} + \sigma_e^2 \mathbf{I} \end{bmatrix} \tag{A.10}$$

Appendix B

Some characteristics of the IEEE 802.11 standard

IEEE 802.11 [130] is a set of MAC (media access control) and PHY (physical layer) specifications for implementing WLAN (wireless local area network). It is commonly known as WiFi. It relies on a series of half-duplex modulation techniques that employ the same basic protocol. While the family of the IEEE 802.11 standards uses mainly two modulation techniques: DSSS and OFDM, we focus here on the characteristics of the OFDM waveforms.

Table B.1 summarizes the main characteristics of the IEEE 802.11 versions that use the OFDM modulation at the physical layer, where the number of data subcarriers excludes those dedicated for guardband interval.

- 802.11a: Released in 1999, it defines protocols that enable transmission and reception of data at rates of 1.5 to 54 Mbit/s at 5 GHz band. Wireless access point (cards and routers) manufacturers used the term "802.11a" to describe interoperability of their systems at 5.8 GHz, 54 Mbit/s.

- 802.11g: Released in 2003, it operates at 2.4 GHz band with a maximum physical layer bit rate of 54 Mbit/s. Owing to the desire for higher data rates as well as to reductions in manufacturing costs, consumers started adopting the 802.11g in January 2003.

- 802.11n: Released in 2009, it is an amendment which improves upon the previous 802.11 standards mainly by increasing the channel DoFs, by adding MIMO and wider channel bandwidths (40 MHz vs. 20 MHz). 802.11n operates on both

the 2.4 GHz and the lesser used 5 GHz bands, where the 40 MHz bandwidth is exclusively used at 5 GHz band. It operates at a maximum net data rate from 54 Mbit/s to 600 Mbit/s.

- 802.11ac: Approved in 2014, it is an amendment that operates at the 5 GHz band. It relies on 802.11n, but brings improvement with respect to 802.11n including wider channel bandwidths (80 or 160 MHz vs. 40 MHz), more spatial streams (up to 8 vs. 4), higher order modulation (up to 256-QAM vs. 64-QAM), and the addition of multi-user MIMO.

Table B.1: IEEE 802.11 standard characteristics.

Standard IEEE	Frequency band (GHz)	Bandwidth (MHz)	FFT size	Total sub-carriers	Data sub-carriers	MIMO
802.11a	5	20	64	52	48	N/A
802.11g	2.4	20	64	52	48	N/A
802.11n	2.4, 5	20, 40	64, 128	52(or 56), 114	48(or 52), 108	4
802.11ac	5	20, 40, 80, 160	64, 128, 256, 512	52, 114, 242, 484	48, 108, 234, 468	8

TELECOM ParisTech

EDiE

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY