Physical layer security for IoT applications

Miroslav Mitev

Supervisors: Martin Reed, Arsenia Chorti

A thesis submitted for the degree of PhD

School of Computer Science and Electronic Engineering

University of Essex

August 2020

Abstract

The increasing demands for Internet of things (IoT) applications and the tremendous increase in the volume of IoT generated data bring novel challenges for the fifth generation (5G) network. Verticals such as e-Health, vehicle to everything (V2X) and unmanned aerial vehicles (UAVs) require solutions that can guarantee low latency, energy efficiency, massive connectivity, and high reliability. In particular, finding strong security mechanisms that satisfy the above is of central importance for bringing the IoT to life.

In this regards, employing physical layer security (PLS) methods could be greatly beneficial for IoT networks. While current security solutions rely on computational complexity, PLS is based on information theoretic proofs. By removing the need for computational power, PLS is ideally suited for resource constrained devices. In detail, PLS can ensure security using the inherit randomness already present in the physical channel. Promising schemes from the physical layer include physical unclonable functions (PUFs), which are seen as the hardware fingerprint of a device, and secret key generation (SKG) from wireless fading coefficients, which provide the wireless fingerprint of the communication channel between devices.

The present thesis develops several PLS-based techniques that pave the way for a new breed of latency-aware, lightweight, security protocols. In particular, the work proposes: i) a fast multi-factor authentication solution with verified security properties based on PUFs, proximity detection and SKG; ii) an authenticated encryption SKG approach that interweaves data transmission and key generation; and, iii) a set of countermeasures to man-in-the-middle and jamming attacks. Overall, PLS solutions show promising performance, especially in the context of IoT applications, therefore, the advances in this thesis should be considered for beyond-5G networks.

Acknowledgments

I would like to thank the University of Essex, School of Computer Science and Electronic Engineering (CSEE) Doctoral Training Programme for sponsoring my studies and made my dream of pursuing a PhD degree reality.

I would like to sincerely thank my supervisors, Dr Martin Reed and Dr Arsenia (Ersi) Chorti for their valued advice and motivation while conducting this research, thanks to their continuous guidance I enjoyed a stress-free working environment. They let me freely choose my research directions and demonstrated a great deal of support. Our meetings and discussions were the key that brought this thesis to life. Furthermore, I would like to express my gratitude for the networking and travelling opportunities which greatly contributed in enhancing my research experience.

My sincere thanks also go to the members of staff of the School of Computer Science and Electronic Engineering and the Department of Mathematical Sciences, University of Essex who provided me the opportunity to become a Graduate teaching assistant and Graduate lab assistant. This was an amazing experience which helped me to advance with my PhD studies.

I would like to thank my co-authors Dr. Leila Musavian and Dr. Veronica Belmega for their insightful comments and support. I would also like to thank Dr. Mahdi Herfeh for our ongoing collaboration and future publication on the short-block length solution which is not fully contained in this thesis.

Finally, I wish to express my deepest gratitude to my family members, my friends, my colleagues, and my life-partner Elena, for their constant support throughout writing this thesis.

Contents

Notation

Variables, vectors and matrices

x	scalar
$\mathbf{x}$	vector
$\|\mathbf{x}\|$	size of a vector
$\mathbf{X}$	matrix
$\mathbf{X}^T$	transpose of a matrix
$\mathbf{I}$	identity matrix

Sets

$\mathbb{N}$	the set of natural numbers
$\mathbb{R}$	the set of real numbers
$\mathbb{C}$	the set of complex numbers
$\mathcal{X}$	the set of elements
$\{\mathbf{x_1}, \ldots, \mathbf{x_N}\}$	set with elements $\mathbf{x_1}, \ldots, \mathbf{x_N}$
$\|\mathcal{X}\|$	cardinality of set $\mathcal{X}$

Functions and operations

$\mathtt{Hash}$	hashing
$\mathtt{Es}$	encryption algorithm
$\mathtt{Ds}$	decryption algorithm
$\mathtt{Sign}$	signing algorithm
$\mathtt{Ver}$	verification algorithm
$\mathtt{Gen}$	fuzzy extractor generate function
$\mathtt{Rep}$	fuzzy extractor reproduce function
$\oplus$	bitwise exclusive-OR operation
$\{\mathbf{x}\|\|\mathbf{y}\}$	concatenation of vectors
$\lfloor\cdot\rfloor$	rounding operator returns integer less than or equal to a given number
$\lceil\cdot\rceil$	rounding operator returns integer greater than or equal to a given number

Probability

$\mathsf{Pr}(X)$	probability of event X
$\mathsf{Pr}(X\|Y)$	probability of event X under condition Y
$\mathsf{Pr}(X,Y)$	joint probability of events X and Y
$\mathsf{E}[\mathbf{x}]$	expectation of $\mathbf{x}$
$\mathsf{E}[\mathbf{x}\|\mathbf{y}]$	expectation of $\mathbf{x}$ under condition $\mathbf{y}$
$\mathsf{H}(\mathbf{x})$	entropy of $\mathbf{x}$
$\mathsf{I}(\mathbf{x}\|\mathbf{y})$	mutual information of $\mathbf{x}$ and $\mathbf{y}$
$\mathsf{cov}(\mathbf{x}\|\mathbf{y})$	covariance of $\mathbf{x}$ conditioned on $\mathbf{y}$

Acronyms and Abbreviations

0-RTT	Zero-round-trip-time
3GPP	3rd Generation Partnership Project
5G	Fifth generation mobile network
AE	Authenticated encryption
AES	Advanced encryption standard
AWGN	Additive white Gaussian noise
B5G	Beyond 5G
BF	Block Fading
BLE	Bluetooth low energy
BR	Best response
CRP	Challenge-response pair
CSI	Channel state information
DHE	Diffie Hellman
DoS	Denial of service
DY	Dolev-Yao
EAP-TLS	Extensible authentication protocol-transport layer security
ECC	Elliptic curve cryptography
EH	Energy harvesting
FE	Fuzzy extractor
HMAC	Keyed-Hash Message Authentication Code
IC	Integrated circuit

IoT	Internet of things
LDP	Large deviation principle
MAC	Message authentication code
MB logic	Mao and Boyd logic
MiM	Man-in-the-middle
NB-IoT	Narrowband Internet of things
NE	Nash equilibrium
NIST	National Institute of Standards and Technology
OFDM	Orthogonal frequency division multiplexing
PHY	Physical layer
PKE	Public key encryption
PLS	Physical layer security
POWF	Physical-one way function
PUF	Physical unclonable function
QoS	Quality of service
QPSK	Quadrature amplitude phase shift keying
RSA	Rivest-Shamir-Adleman
RSS	Received signal strength
RSSI	Received signal strength indicator
SE	Stackelberg equilibrium
SKG	Secret key generation
SNR	Signal to noise ratio
SSH	The secure shell
STEK	Session ticket encryption key
SW	Slepian Wolf
TLS	Transport layer security
URLLC	Ultra-reliable low-latency communications
V2X	Vehicle to everything

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The physical layer security (PLS) paradigm dates back to Wyner's pioneering work in the late 1970's [1]. However, after decades of advancement, it is just recently that PLS is being considered as an enabling technology in a large set of applications such as: vehicular communications [2, 3], underwater communications [4], optical fiber [5], visible light communication [6] and many more as summarised in [7, 8]. The aim of PLS [9–11] is to make use of the inherent randomness in the physical layer, including the communication channel and electronic circuits, and achieve improvements in critical security aspects. In particular, PLS can offer: i) user and message authentication [12]; ii) symmetric key generation and key agreement solutions [13, 14]; iii) countermeasures to jamming attacks [15, 16]; and, iv) confidentiality [17].

The increasing interest in PLS has been stimulated by many practical needs. Notably, many critical IoT networks require fast authentication, *e.g.*, in autonomous driving and vehicle to everything (V2X) applications, telemedicine and haptics. However, standard cryptographic schemes, particularly those in the realm of public key encryption (PKE), are computationally intensive, incurring considerable overheads and can rapidly drain

the battery of power constrained devices [18–20]. For example, the recent (2019) third generation partnership project (3GPP) technical report "Study on the Security of URLLC" [21], mentions that all aspects related to low-latency authentication remain open and no solutions have so far been standardised. In this regard, PLS has been explicitly mentioned in the first white paper on 6G [22]: "*The strongest security protection may be achieved at the physical layer*" and more importantly, it is stated as an enabling technology in the IEEE International Network Generations Roadmap [23].

A further challenge comes from quantum computing, which has seen significant progress after massive investment by companies such as Google, Intel and IBM to build prototypes with more than 50 qubits. In October 2019 Google published in the journal "Nature" their quantum computer experiments showing they have achieved quantum supremacy for a particular set of problems [24]. In this aspect, PLS, that relies upon information-theoretic security proofs, could resist quantum computers, unlike corresponding asymmetric key schemes relying on the (unproven) intractability in polynomial time of certain algebraic problems [25]. Even state-of-the-art elliptic curve cryptography (ECC) schemes [26], that require substantially shorter keys than Rivest-Shamir-Adleman (RSA) [27] or Diffie Hellman (DHE) [28] schemes, are still considerably more intensive computationally than their PLS counterparts and have not been shown to be post-quantum secure.

Another motivation in improving PLS solutions stems from the fact that a number of vulnerabilities, *e.g.*, jamming during the beam allocation in mmWave [29], arise during the establishment of the radio link. In this aspect, standard security protocols that build on the premise that the communication link has already been established, cannot offer solutions when this is not the case, whereas, PLS schemes can be seamlessly incorporated (*e.g.*, can be interwoven with channel estimation).

In conclusion, PLS provides a new opportunities of security for IoT networks. However, it has not been completely developed and efficiently integrated yet. Therefore, the proposal of novel PLS based solutions for future networks security is highly pertinent.

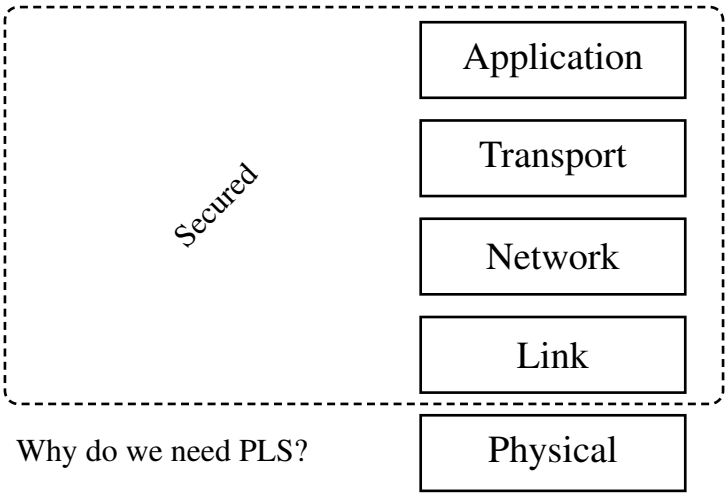

Figure 1.1: Today's security architecture

1.2 Approach of this thesis

The architecture of today's wireless systems uses layers. As it can be seen in Fig. 1.1 the layers above the physical layer enable security by various methods. In this sense, the naturally arising question is: *Why do even need PLS?*. In fact the higher layer cryptography mechanisms have numerous of advantages: i) there are no known feasible attacks; ii) they are widely employed and tested; iii) they provide trustworthy authentication. However, these methods have some disadvantages: i) they are built on unproven assumptions about the difficulty of the certain computation problems; ii) they are built on the assumption that the physical connection is already established; iii) they require significant infrastructure. Therefore, the answer to the above question is that PLS will be used in emerging technologies.

In the past years, PLS [9–11, 14, 30, 31] has been studied as a possible alternative to classic, complexity based, cryptography. Signal properties can be exploited to allow for confidential data transmission [32–34]. Therefore, work in the present thesis, proposes to move some of the security core functions down to the physical layer, exploiting both the communication radio channel and the hardware, as unique entropy sources.

Since the wireless channel is reciprocal, time-variant and random in nature, it offers

a valid, inherently secure source that may be used in a key agreement protocol between two communicating parties. The principle of secret key generation (SKG) from correlated observations was first studied in [35] and [36]. A straightforward SKG approach can be built by exploiting the reciprocity of the wireless fading coefficients between two terminals within the channel coherence time [37,38] and this thesis builds upon this mechanism and proposes an optimised SKG mechanism for delay-constrained systems. This is pertinent to many forthcoming beyond 5th generation mobile (B5G) applications that will require a strong, but nevertheless, lightweight security key agreement.

However, unauthenticated key generation is vulnerable to man-in-the-middle (MiM) attacks. In this sense, physical unclonable functions (PUFs), firstly introduced in [39] (based on the idea of physical one-way functions [40]), can provide authenticated encryption and furthermore can decrease the computational cost and have a high impact on reducing the authentication latency in constrained devices [41]. Consequently, this thesis provides a full PUF-based authentication solution. Finally, SKG schemes are known to be malleable over the so called "advantage distillation" phase, during which observations of the shared randomness are obtained at the legitimate parties. As an example, an active attacker can inject pilot signals and/or can mount denial of service attacks (DoS) in the form of jamming. Therefore, this thesis investigates the impact of injection and reactive jamming attacks in SKG and presents set of countermeasures.

To conclude, conventional cryptography mechanisms are a vital element for securing wireless networks. However, they are computationally intensive and require significant infrastructure to extract high entropy keys. On the other hand, the physical layer provides a source of domain-specific information which can be used to complement and improve conventional security methods. The mechanisms proposed within the present thesis, could be easily incorporated in today's security architecture as a substitute to today's authentication protocols (Chapter 3) and session key generation mechanisms (Chapters 4, 5).

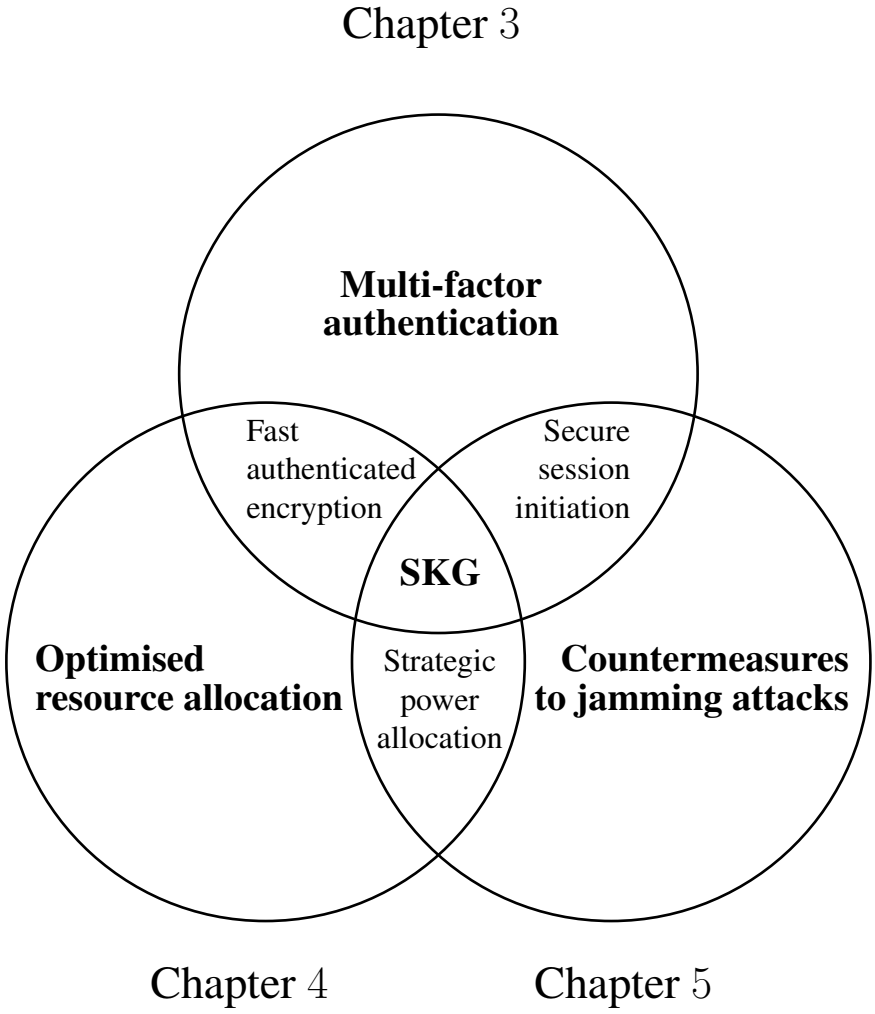

Figure 1.2: Roadmap of contributions.

1.3 Contributions

The contributions of the present thesis are illustrated in Fig. 1.2 and can be summarised as follows:

Chapter 3

Lightweight and low latency security mechanisms are becoming increasingly important for a wide range of IoT applications. Promising schemes from the physical layer include (i) PUF, (ii) localisation based authentication, and, (iii) SKG from wireless fading

coefficients. The work in Chapter 3 proposes a complete, fast, multi-factor authentication protocol that uniquely combines PUFs, proximity estimation and SKG. It focuses on short blocklengths and provides a novel closed form expression for the channel dispersion in the SKG setting. Furthermore, the SKG keys are incorporated in a forward secure zero-round-trip-time (0-RTT) resumption protocol for fast re-authentication. All schemes of the proposed mutual authentication protocol are shown to be secure through formal proofs using Burrows, Abadi and Needham (BAN) and Mao and Boyd (MB) logic as well as the Tamarin-prover [42].

Chapter 4

In computational complexity and latency constrained emerging 5G applications, *e.g.*, autonomous vehicles, haptic communications and enhanced reality, SKG at the physical layer could be considered as an alternative to currently used key agreement schemes. In this framework, the work present in Chapter 4 proposes i) a novel authenticated encryption (AE) using SKG, and, ii) a pipelining mechanism of the AE SKG and the encrypted data transfer in order to reduce latency. Implementing the pipelining at the physical layer, the work investigates a *parallel* SKG approach for multi-carrier systems, where a subset of the subcarriers are used for SKG and the rest for data transmission. The parallel approach is evaluated under power, security, rate and delay constraints. The amount of data that can be transmitted with a single key is determined by the cryptographic suites used, so that realistic key rate constraints can be identified. This allows to formulate the subcarrier allocation as a subset-sum $0-1$ knapsack optimisation problem that is solved using i) the standard dynamic programming approach, and, ii) a greedy heuristic approach of linear complexity. Numerical evaluation shows that the proposed heuristic induces virtually no loss in performance. Furthermore, a comparison with a baseline scheme in which secret key generation and data transfer are performed sequentially, shows that the proposed parallel approach offers gains in terms of efficiency. All of the proposed mechanisms, have the potential to pave the way for a new breed of latency aware security protocols.

Chapter 5

Physical layer SKG from shared randomness (*e.g.*, from the wireless channel fading

realisations), is a well established scheme that can be used for session key agreement. SKG approaches can be of particular interest in delay constrained wireless networks and notably in the context of ultra reliable low latency communications (URLLC) in B5G systems. However SKG schemes are known to be malleable over the so called "advantage distillation" phase, during which observations of the shared randomness are obtained at the legitimate parties. As an example, an active attacker can act as a MiM by injecting pilot signals and/or can mount denial of service attacks (DoS) in the form of jamming. In this sense, Chapter 5 investigates the impact of injection and reactive jamming attacks in SKG. First, it is demonstrated that injection attacks can be reduced to – potentially less harmful – jamming attacks by using pilot randomisation; consequently, a novel system design with randomised quadrature amplitude phase shift keying (QPSK) pilots is presented. Subsequently, the optimal jamming strategy is identified in a block fading additive white Gaussian noise (BF-AWGN) channel in the presence of a reactive jammer, using a game theoretic formulation. It is shown that the impact of a reactive jammer is far more severe than that of a simple proactive jammer.

1.4 Outline of thesis

This thesis is structured as follows:

Chapter 2, highlights the elements of the problems discussed in the present thesis. The following three chapters contain the main contributions of the present thesis. The organisation of these chapters is inspired by the sequence of actions taken in an actual system, *i.e.,* 1) authentication; 2) session key generation; 3) attacks during the key generation process. In this sense, Chapter 3, proposes a novel PUF-based authentication mechanism that is initialised by location confirmation approach. Furthermore, building on the SKG process it provides a forward secure resumption protocol to quickly "resume" sessions. The concept of the work presented in this chapter was published in [12], whereas the full solution is under review for publication in the "IEEE Internet of things journal". Next, chapter 4 presents a physical layer resource allocation method that jointly optimised data

rate and key generation rate. This chapter is based on works presented in [12–14, 43] and involves collaboration with Dr. Leila Musavian. Chapter 5 presents countermeasures to active attacks in wireless networks through a game-theoretic approach. This chapter is based on works presented in [15] and involves collaboration with Dr. Veronica Belmega. Finally, my perspectives for future research are presented in Chapter 6.

1.5 List of publications

Journals

M. Mitev, A. Chorti, M. Reed, L. Musavian, "Authenticated Secret Key Generation in Delay Constrained Wireless Systems", *EURASIP Journal on Wireless Communications and Networking*, June, 2020.

M. Mitev, M. Herfeh, A. Chorti, M. Reed, "Fast Multi-factor Physical Layer Security Authentication in the Short Blocklength", under review, IEEE IoT journal.

Conferences

M. Mitev, A. Chorti, V. Belmega, M. Reed, "Man-in-the-Middle and Denial of Service Attacks in Wireless Secret Key Generation", IEEE Global Communications Conference (Globecom), Dec 2019.

M. Mitev, A. Chorti, M. Reed, "Subcarrier Scheduling for Joint Data Transfer and Key Generation Schemes in Multicarrier Systems", IEEE Global Communications Conference (Globecom), Dec 2019.

M. Mitev, A. Chorti, M. Reed, "Optimal Resource Allocation in Joint Secret Key Generation and Data Transfer Schemes", IEEE International Wireless Communications and Mobile Computing Conference (IWCMC), Tangier, Morocco, Jun 2019.

Posters

M. Mitev and A. Chorti and M. Reed, "Optimal resource allocation in secure multi-carrier systems", Munich Workshop on Coding and Cryptography (MWCC), Germany, invited poster presentation, Apr 2018.

Chapter 2

Background

This chapter summarises only the general background in which this thesis was built. It provides an overview of physical layer security highlighting its applicability within IoT networks. The relative background to each study presented within the present thesis is provided later as part of the respective chapter.

2.1 Physical layer security within the 5G framework

The fifth generation (5G) communication technologies such as ultra-reliable low-latency communications (URLLC) and massive machine-type communications (mMTC) are expected to enable numerous IoT applications [44]. IoT devices are distributed in our daily life environment gathering information that ranges from temperature to location making security and privacy of critical importance [45]. However, despite the continuous improvement of security protocols, there are still open issues that have not been fully addressed [46]. In this sense, transforming the security design bottom up starting from the physical layer, will provide additional level of protection which can help overcome the security hurdles [47]. PLS mechanisms can enhance the security of IoT networks from several aspects. Authentication is central in building secure IoT networks; confirming the

identity of devices and their role in the network hierarchy eliminates the possibility of numerous attacks [48]. However, the low-latency and computational power constraints present in many IoT systems [20], renders the design of IoT authentication mechanisms a challenging task. Current solutions rely on modulo arithmetic in large fields and typically incur considerable latency, in the order of tens of milliseconds [49]; as an example, it has been reported that verifying digital signatures on a vehicle with a 400 MHz processor takes around 20 msec [50], exceeding the delays that are tolerated in vehicle to everything (V2X) communications [51]. In that direction PLS schemes exploit physical layer entropy sources, including both in the hardware, as well as in the communication medium [12–15]. With respect to the former, physical unclonable functions (PUFs) are hardware entities harnessing entropy from physically unclonable variations that occur during the production process of silicon [39, 52]. These unique and unpredictable variations allow the extraction of uniformly distributed binary sequences [53]. Due to their unclonability and simplicity, PUFs are seen as lightweight security primitives that can offer alternatives to today's authentication mechanisms [54, 55]. With respect to authentication, the research presented in Chapter 3 includes multi-factor authentication protocol with PUFs, wireless fingerprinting and localisation. The necessary background on PUFs are given in Section 3.2.2. Furthermore, a novel PLS-based resumption protocol that allows for data exchange within 0-RTT is proposed.

In numerous IoT scenarios mobility is an important factor [56]. In applications such as V2X and drone connectivity the objects of interest may leave or enter the network in a random manner [2, 57–59]. This requires strong security mechanisms with low overhead. Furthermore, the massively distributed IoT devices have limited resources and therefore, cannot employ complex cryptography [46, 60]. In this sense, IoT networks could greatly benefit from employing PLS solutions. Since the wireless channel is reciprocal, time-varying and random in nature, it offers a lightweight, inherently secure source for secret key distribution protocols [7, 8, 61]. In fact, it has been experimentally validated that mobility could highly increase the entropy of PLS generated keys [62, 63]. Furthermore, new 5G wireless methods will employ advanced channel state information (CSI) estimation

techniques towards optimal beamforming and improved reliability [64, 65]. In this regards, Chapter 4 within the present thesis, shows that the CSI measurements can be used simultaneously towards both: generation of secret keys and reliable data transfer.

The roll-out of 5G mobile networks and the forthcoming the beyond 5G (B5G) will bring about fundamental changes in the way we communicate, access services and consume entertainment [66]. With respect to the latter, the multi-fold increase in the service data rates will provide users with ultra high resolution in video-streaming, multi-media and virtual reality, offering immersive experiences [67, 68]. However to support such a large variety of services, requires novel solutions that will enable higher resource efficiency [69]. From that perspective, Chapter 4 within the present thesis proposes a novel optimised PLS resource allocation algorithm that jointly optimises data and key generation rates.

A further motivation in exploring PLS solutions stems from the fact that a number of vulnerabilities may impact the reliability of the exchanged information [70, 71]. For example, jamming attacks in URLLC may provoke re-transmission and increased overall delay [15, 72], furthermore, jamming during the beam allocation in mmWave [29] may disrupt the establishment of the radio link. A common assumption of security protocols is that the communication link has already been established, however, this may not be the case. Consequently, PLS offers several counter-jamming techniques, including: jamming detection mechanisms and counter-jamming measures in the form of i) energy harvesting (EH) [16,73], ii) channel hopping [74] and power spreading [75,76], further approaches to counter jamming attacks include iii) friendly jamming [77, 78], and iv) localisation [79]. In fact, studies [16, 73] are carried in a game-theoretic framework. The legitimate users, who have EH capabilities, act as a single player who aims at maximizing the transmission rates while the jammer aims at minimizing it. It is demonstrated that, for specific scenarios, the legitimate users could benefit from the harvested jamming energy by subsequently using it to boost their transmission. On the other hand, channel hopping [74] and power spreading [75, 76], are commonly used technique to counter jamming attacks. The legitimate users can use channel hopping in a random fashion in order to avoid most

of the jamming interference, however, depending on the scenario, it may be more beneficial to spread their power over the entire spectrum rather than concentrating it on a single channel. In this sense, Chapter 5 of the present thesis explores jamming attacks during a PLS SKG process and gives a set of countermeasures.

Overall, the physical layer properties carry important information which is directly related to security and privacy of today's systems. The intelligence of any actual system in the physical world relies on its physical infrastructure. Important questions that naturally arise are then: "how to turn the collected data into useful knowledge?"; and, "why the characteristics of the physical channel are used in enhancing the systems's reliability but are not used towards improving security?" In response to this questions, the present thesis proposes a set of techniques that could be employed independently or as a complement to existing techniques, with minimal changes in the control plane. The physical layer solutions developed within the premise of this thesis include: a lightweight PUF authentication mechanism, fast proximity detection, a resumption protocol based on physical layer SKG, an authenticated encryption (AE) SKG approach, a pipelining algorithm that interweaves data transmission and SKG, a near-optimal subcarrier allocation algorithm of linear complexity, pilot randomisation as a countermeasure to injection attacks and optimal power allocation policy in the presence of active jammers.

2.2 Key-based and key-less physical layer security

As discussed, current security solutions: are focused at the higher layers of the communication systems; are computationally intensive; and, are not suitable for resource constrained devices, such as in an IoT scenario. In this sense, PLS has proven itself as a lightweight information theoretic secure mechanism regardless of the adversary's resources [80, 81]. This section gives an introduction to the two basic PLS methods [82]: key-based PLS and key-less PLS.

2.2.1 Key-based PLS: secret key generation

Due to the nature of radio waves, wireless communication remains vulnerable to different types of attacks. Passive attacks such as eavesdropping or traffic analysis can be performed by anyone close to the area of communication. Furthermore, wireless devices can be targeted by active attacks such as jamming [15, 74] or spoofing which can cause interference or even disruption of the device (summary of attacks and countermeasures in wireless systems can be found in [83] and [84], respectively). Therefore, to ensure confidentiality, data encryption is vital for communication security. However, most of the traditional public-key encryption primitives such as RSA [27] and ECC [26] have high computational complexity and are not suitable for portable and battery-driven devices, such as in an IoT system.

In this sense, key-based PLS [85] methods are seen as a lightweight alternative for secret key generation (SKG)[1] by using the properties of the wireless medium. The reciprocity of wireless medium can be used as a tool for key generation between two parties (Alice and Bob), as shown on the block diagram in Fig. 2.1. In this case, Alice and Bob perform pilot exchanges that take place over the coherence time of the channel[2]. During this period, Alice and Bob observe highly correlated channel states, as opposed to Eve, whose observations are uncorrelated to the main channel. Following from that, Alice and Bob can use their observations to generate a shared secret key between them. Important properties that support the above mechanism are:

i) in order to reduce the key mismatch rate, each round of the pilot exchange must be within the coherence time of the channel. One round refers to one pilot sent in each direction, *i.e.*, during each round, two pilots are sent, one from Alice and one from Bob;

[1] The term SKG throughout this thesis refers to the key-based PLS mechanism introduced in this section.

[2] The coherence time corresponds to the interval during which the multipath properties of wireless channels (channel gains, signal phase, delay) remain stable [3, 86, 87]. It is inversely proportional to the Doppler spread, which on the other hand, is a dispersion metric that accounts for the spectral broadening caused by the user's mobility (for more details and derivation please see [87]).

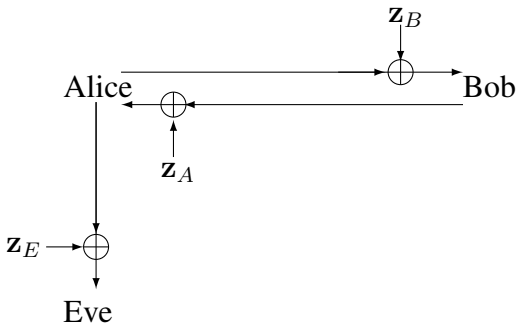

Figure 2.1: During the coherence time of the channel Alice and Bob observe highly correlated channel responses. Eve's observation is uncorrelated to their observations. $\mathbf{z}_A, \mathbf{z}_B, \mathbf{z}_E$ represent the AWGN noise variables observed by Alice, Bob and Eve, respectively.

ii) in order to have uncorrelated key bits, the time interval between rounds must be greater than the coherence time of the channel.

Overall, the SKG mechanism could greatly benefit from the time-variability property of the wireless channels [82, 88]. In fact, the security key is generated from the variations in the received signal (i.e., fading). Therefore, time-varying signals could produce secret keys with higher entropy.

However, in some static scenarios, Alice and Bob may struggle to extract enough bits from the channel. To overcome this, instead of transmitting publicly known pilot signals, one or two-way randomised signal transmission can be performed which can be used as an additional source of randomness [89]. Furthermore, Chapter 5 within the present thesis will show how randomised pilot exchange can be used as a countermeasure to injection attacks.

Commonly used parameters as source of shared randomness are the received signal strength (RSS) and the full CSI [90] (some channel estimation techniques can be found in [90]). However, RSS-based schemes are open to predictable channel attacks [86, 91] whereas CSI-based approaches, have proven to be resistant [92]. Therefore, this work assumes the usage of a CSI-based SKG, which has been practically investigated in numerous scenarios [92–95] as briefly described below.

The authors of [92] investigate the effect of non-reciprocity of the CSI caused by ad-

ditive noise, antenna gains *etc*., between two devices. To overcome this, a non-reciprocity component is evaluated during a learning phase before the SKG process. Based on this, the authors present a channel gain complement scheme which eliminates the impact of non-reciprocity in both indoor and outdoor environments.

Another important step in a practical implementation is to build a suitable unit for pre-processing in order to decorrelate the signals in the time / frequency and space domains. As an example, some recent works have shown that the widely adopted assumption [96] that a distance equal to half of the wavelength (which at 2.4 GHz is approximately 6 cm [92]) is enough for two channels to decorrelate, may not hold in reality [86].

In [93] the problem caused by the similarity of the characteristics (phase and amplitude) of neighboring subcarriers is addressed. This effect can cause a high correlation in the CSI measurements which may lead to security issues. To overcome this a three-step protocol for wireless SKG between two entities is proposed: i) CSI measurements are converted to bits; ii) universal hash function is applied to the bit strings and, iii) if the resulted bit strings after hashing differ, part of the bits on both sides are recombined until same bit sequence is observed at each of the parties. This final step helps remove the correlation of bits generated from neighbouring subcarriers.

A further measure addressing correlation between the eavesdropper and communicating devices is role reversal [94]. The party that initiates the SKG process is considered to be a leader, while the follower is trying to match the leader's bit stream. If an eavesdropper is close to the leader her channel might be correlated to the main one, thus she could obtain part of the key bits. With role reversal the legitimate parties change their roles – from leader to follower and vice versa – during the process and the obtained bit strings are concatenated at the end. This step prevents leak of key bits through eavesdropping.

To summarise, the impact of channel correlation on SKG needs to be explicitly accounted in actual implementations. However, in this thesis its effect is neglected as its impact on SKG has been treated in numerous contributions in the past [62,63,93–95,97–100] and will not enhance the problems formulated in the following chapters.

In addition to channel correlation issues, to improve error correction in the SKG pro-

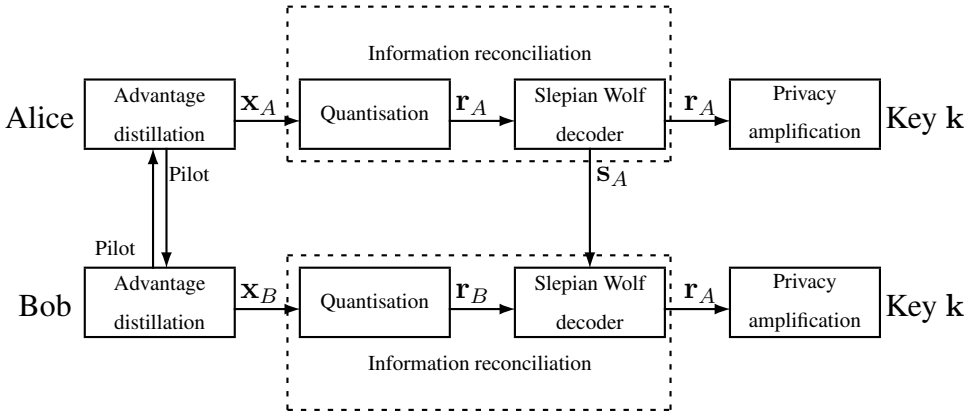

Figure 2.2: Secret key generation process between Alice and Bob.

cess, alternative quantisation schemes have been proposed [95]. In [95] two legitimate parties divide their observations into two bit streams and perform quantisation on two different levels. Next, one of the streams is used as error correction data and the other is used as a key generation string.

SKG process

The SKG process takes an important part of the work presented throughout the present thesis. In fact, Chapter 3 employs the SKG method to built a quick authentication protocol; Chapter 4 proposes an optimised SKG method for delay-constrained networks; Chapter 5 introduces a set of countermeasures to jamming attacks performed during the SKG process. The SKG system model within this thesis assumes that two legitimate parties, Alice and Bob, who wish to establish a symmetric secret key using the wireless fading coefficients as a source of shared randomness (See Fig. 2.2). Generally, the SKG procedure encompasses three phases: *advantage distillation*, *information reconciliation*, and *privacy amplification* [35, 36, 101–103] as described below:

1) Advantage distillation: This phase takes place during the coherence time of the channel. The legitimate nodes sequentially exchange constant probe signals to obtain estimates of their reciprocal CSI. Note that, the pilot exchange phase can be made robust with respect to injection type of attacks (that fall in the general category of MiM) as anal-

ysed in Chapter 5 [15, 38]. At the end of this phase, Alice and Bob obtain observation vectors $\mathbf{x}_A$ and $\mathbf{x}_B$, respectively, over a set of N subcarriers. On the other hand, an eavesdropper (Eve) observes $\mathbf{x}_E$. The work within the present thesis assumes a rich Rayleigh multipath environment, such that Eve's channel measurement remains uncorrelated to main channel.

2) Information reconciliation: At the beginning of this phase Alice's and Bob's observations, $x_{A,j}, x_{B,j}$, respectively, are quantised to binary vectors[3] $\mathbf{r}_{A,j}$, $\mathbf{r}_{B,j}$, where $j = 1, \ldots, N$ [104–106], so that Alice and Bob distill $\mathbf{r}_A = [\mathbf{r}_{A,1}|| \ldots ||\mathbf{r}_{A,N}]$ and $\mathbf{r}_B = [\mathbf{r}_{B,1}|| \ldots ||\mathbf{r}_{B,N}]$, respectively. Due to the presence of noise, $\mathbf{r}_A$ and $\mathbf{r}_B$ will differ. To reconcile discrepancies in the quantiser local outputs, side information needs to be exchanged via a public channel. Using the principles of Slepian Wolf decoding [107], the distilled binary vectors can be expressed as

$$\mathbf{r}_A = \mathbf{d} + \mathbf{e}_A, \tag{2.1}$$

$$\mathbf{r}_B = \mathbf{d} + \mathbf{e}_B, \tag{2.2}$$

where $\mathbf{e}_A, \mathbf{e}_B$ are error vectors that represent the distance from the common observed (codeword) vector $\mathbf{d}$ at Alice and Bob, respectively.

Numerous practical information reconciliation approaches using standard forward error correction codes (*e.g.*, LDPC, BCH, *etc*.,) have been proposed [37], [108]. As an example, if a block encoder is used, then the error vectors can be recovered from the syndromes $\mathbf{s}_A$ and $\mathbf{s}_B$ of $\mathbf{r}_A$ and $\mathbf{r}_B$, respectively. Alice transmits her corresponding syndrome to Bob so that he can reconcile $\mathbf{r}_B$ to $\mathbf{r}_A$. It has been shown that the length of the syndrome $|\mathbf{s}_A|$ is lower bounded by $|\mathbf{s}_A| \geq \mathsf{H}(\mathbf{x}_A|\mathbf{x}_B) = \mathsf{H}(\mathbf{x}_A, \mathbf{x}_B) - \mathsf{H}(\mathbf{x}_B)$ [36], note that $\mathsf{H}(\cdot)$ denotes entropy. This has been numerically evaluated for different scenarios and coding techniques [105, 109–111]. Following that, the achievable SKG rate is upper bounded by $\mathsf{I}(\mathbf{x}_A; \mathbf{x}_B)$, where $\mathsf{I}(\cdot)$ denotes mutual information.

3) Privacy amplification: The secret key is generated by passing $\mathbf{r}_A$ through a one-

[3]Note that each observation can generate a multi-bit vector at the output of the quantiser.

way collision resistant *compression* function *i.e.*, by hashing. Note that this final step of privacy amplification, is executed locally without any further information exchange. The need for privacy amplification arises in order to suppress the entropy revealed due to the public transmission of the syndrome $\mathbf{s}_A$. Privacy amplification produces a key of length strictly shorter than $|\mathbf{r}_A|$, at least by $|\mathbf{s}_A|$. At the same time, the goal is for the key to be uniform, *i.e.*, to have maximum entropy. In brief, privacy amplification *reduces the size of the sequence* while at the same time *increases its entropy per bit* – compared to the input.

The privacy amplification is typically performed by applying either cryptographic hash functions such as those built using the Merkle-Damgard construction, or universal hash functions and has been proven to be secure, in an information theoretic sense, through the leftover hash lemma [112]. As an example, [86, 113] use a 2-universal hash family to achieve privacy amplification. Summarising, the maximum key size after privacy amplification is [12]:

$$|\mathbf{k}| \leq \mathsf{H}(\mathbf{x}_A) - \mathsf{I}(\mathbf{x}_A; \mathbf{x}_\mathbf{E}) - \mathsf{H}(\mathbf{x}_A|\mathbf{x}_B) - r_0, \tag{2.3}$$

where $\mathsf{H}(\mathbf{x}_A)$ represents the entropy of the measurement, $\mathsf{I}(\mathbf{x}_A; \mathbf{x}_\mathbf{E})$ represents the mutual information between Alice's and Eve's observations, $\mathsf{H}(\mathbf{x}_A|\mathbf{x}_B)$ represents the entropy revealed during information reconciliation and $r_0 > 0$ is a safety parameter that ensures uncertainty on the key at Eve's side. For details and estimation of these parameters in a practical scenario please see [114]. Some practical implementations can be found in [8, 61].

SKG rate

During the advantage distillation phase, illustrated on Fig. 2.3, Alice and Bob communicate over a BF-AWGN channel that comprises N orthogonal subcarriers. The fading coefficients $\mathbf{h} = [h_1, \ldots, h_N]$, are assumed to be independent and identically distributed (i.i.d), complex circularly symmetric zero-mean Gaussian random variables $h_j \sim \mathcal{CN}(0, \sigma_h^2), j = 1, \ldots, N$. The legitimate nodes sequentially exchange constant probe

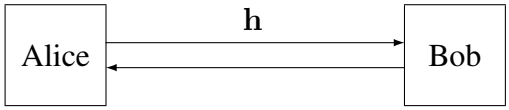

Figure 2.3: Alice and Bob exchange pilot signals over a Rayleigh fading channel with realisation $\mathbf{h} = [h_1, \ldots, h_N]$ in order to distill a shared secret key.

signals with equal power P on all subcarriers[4], to obtain estimates of their reciprocal CSI. Alice and Bob obtain observation vectors $\mathbf{x}_A = [x_{A,1}, \ldots, x_{A,N}], \mathbf{x}_B = [x_{B,1}, \ldots, x_{B,N}]$, respectively, so that:

$$\mathbf{x}_A = \sqrt{P}\mathbf{h} + \mathbf{z}_A, \tag{2.4}$$

$$\mathbf{x}_B = \sqrt{P}\mathbf{h} + \mathbf{z}_B, \tag{2.5}$$

where $\mathbf{z}_A$ and $\mathbf{z}_B$ denote zero-mean, unit variance circularly symmetric complex AWGN random vectors, such that $\mathbf{z}_A \sim \mathcal{CN}(\mathbf{0}, N_A\mathbf{I})$ and $\mathbf{z}_B \sim \mathcal{CN}(\mathbf{0}, N_B\mathbf{I})$. Finally, as discussed in the previous section the SKG rate can be expressed as $\mathsf{I}(\mathbf{x}_A; \mathbf{x}_B)$. While the final result is standard, the full derivation is not clearly given elsewhere, therefore, it is added here as follows:

$$\begin{aligned}
&\mathsf{I}(\mathbf{x}_A; \mathbf{x}_B) = \mathsf{H}(\mathbf{x}_A) + \mathsf{H}(\mathbf{x}_B) - \mathsf{H}(\mathbf{x}_A, \mathbf{x}_B) \qquad (2.6)\\
&= \frac{1}{2}\log_2 2\pi e(\sigma^2 + N_A) + \frac{1}{2}\log_2 2\pi e(\sigma^2 + N_B) - \frac{1}{2}\log_2(2\pi e)^2|K_{\mathbf{x}}|\\
&= \frac{1}{2}\log_2(2\pi e)^2(\sigma^2 + N_A)(\sigma^2 + N_B) - \frac{1}{2}\log_2(2\pi e)^2(N_A N_B + N_A\sigma^2 + N_B\sigma^2)\\
&= \frac{1}{2}\log_2 \frac{(\sigma^2 + N_A)(\sigma^2 + N_B)}{N_A N_B + N_A\sigma^2 + N_B\sigma^2}\\
&= \frac{1}{2}\log_2 \left(\frac{\sigma^2(N_A + N_B) + N_A N_B}{N_A N_B + N_A\sigma^2 + N_B\sigma^2} + \frac{\sigma^4}{N_A N_B + N_A\sigma^2 + N_B\sigma^2} \right)\\
&= \frac{1}{2}\log_2 \left(1 + \frac{\sigma^4}{N_A N_B + N_A\sigma^2 + N_B\sigma^2} \right)\\
&= \frac{1}{2}\log_2 \left(1 + \frac{\sigma^2}{N_A + N_B + \frac{N_A N_B}{\sigma^2}} \right). \qquad (2.7)
\end{aligned}$$

[4]An explanation of the optimality of this choice under different attack scenarios is discussed in [38].

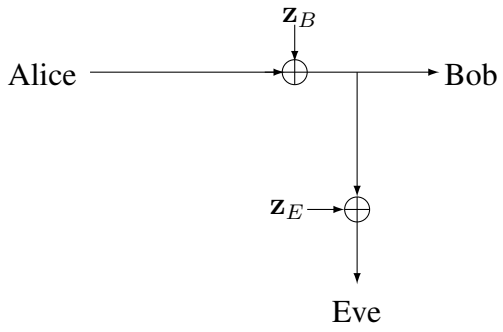

Figure 2.4: Wyner's wiretap channel model. Alice sends information to Bob, who observes channel with noise $\mathbf{z}_B$. Eve acts as an eavesdropper and observes degraded version of Bob's channel.

Assuming a unit variance noise, the rate can be represented as function of the transmitted power as [37, 74]:

$$\mathsf{I}(\mathbf{x}_A; \mathbf{x}_B) = \frac{1}{2}\log_2\left(1 + \frac{P\sigma^2}{2 + \frac{1}{P\sigma^2}}\right). \tag{2.8}$$

This is a key result that is used later in Chapter 4 and Chapter 5.

2.2.2 Key-less PLS: secrecy capacity

Key-less PLS is based on the information theoretic concept known as secrecy capacity. The secrecy capacity is the metric which defines the achievable rate for reliable and secure transmission. It was firstly introduced in the seminal works of Wyner [1], where he introduced the wiretap channel model, illustrated in Fig. 2.4. The model assumes: Alice (transmitter) and Bob (receiver) exchange data over a main channel while Eve observers degraded version of the same channel (*i.e.*, Eve's channel outputs are noisier that Bob's). Given Bob's "advantage", Alice can encode the data in a way that Eve is unable to decode. Later, this concept was generalised for AWGN channels in [115], where the secrecy capacity is estimated as the difference between Bob's and Eve's achievable rates. Given that, the secrecy capacity C_S can be expressed as:

$$C_S = \begin{cases} \log(1 + \text{SNR}_B) - \log(1 + \text{SNR}_E), & \text{if SNR}_B > \text{SNR}_E \\ 0, & \text{if SNR}_B \leq \text{SNR}_E, \end{cases} \tag{2.9}$$

where SNR_B and SNR_E, denotes the SNR level at Bob's and Eve's sides, respectively. Building on that premise, the work in [116] evaluates the outage secrecy capacity for different practical scenarios versus the SNR and the distances Alice – Bob and Alice – Eve. Today, various techniques such as artificial noise at Eve's side [117–119] and wiretap coding schemes [120–122] are used in order to achieve positive secrecy capacity.

Nevertheless, key-less PLS has several drawbacks such as reduced rates and increased overhead with respect to the security feature, plus until recently there were no existing coding scheme exist only for finite blocklength [82] (note that recent advances show promising result, however, these are still in the domain of very large blocklengths, therefore not suitable for IoT applications [123, 124]). Furthermore, a major limitation of the secrecy capacity compared to the SKG presented in the previous section, is that it requires knowledge of the adversary channel and capabilities. In this regards, using the time-variability of the physical channel, the SKG process allows the legitimate users to quickly generate secret keys and use them in existing encryption algorithms. Due to the above, the rest of this thesis focuses on key-based PLS approach in the form of SKG presented in the previous section.

2.3 Possible deployment scenario for SKG: Narrow-Band IoT

Narrow-band Internet of things (NB-IoT) is a standard developed by 3rd Generation Partnership Project (3GPP) in 2015 [125]. It main focus is on low cost indoor applications for battery-driven devices. The NB-IoT physical resource block (subframe) consists of 12 subcarriers with 15 kHz spacing, resulting into 180 kHz transmission bandwidth. In order to have efficient use of the spectral resources, NB-IoT can use three different modes, *i.e.*, stand-alone, in-band and guard-band. While stand-alone mode is intended to replace GSM carriers with NB-IoT ones, in-band and guard-band modes uses LTE carriers, *i.e.*, allowing LTE and NB-IoT traffic to coexists in the same frequency band.

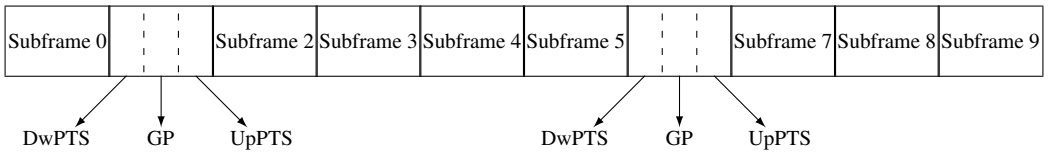

Figure 2.5: One radio frame of the NB-IoT TDD standard

The NB-IoT standard has two frame structures: one based on frequency division duplex (FDD) and the other based on time division duplex (TDD). This makes NB-IoT suitable for the physical layer SKG method described in Section 2.2.1. While FDD frames use different frequency bands for uplink and downlink, the links in a TDD frame are separated in time and not frequency. As discussed earlier, the SKG approach within the present thesis is based on the reciprocity of the wireless channel, *i.e.*, using the method described in Section 2.2.1 Alice and Bob extract a shared secret from their channel responses. The responses are highly correlated during the coherence time of the channel which allows Alice and Bob to exploit the reciprocity of the wireless channel, using the TDD based frame and perform the SKG process. On the other hand, the work within the present thesis assumes that there is no correlation between neighbouring channels, therefore, Alice and Bob would not be able to extract shared secret using the FDD based frame, making this approach not suitable.

The NB-IoT frame type 2 (based on TDD) is given in Fig. 2.5. It can be seen that there is a special subframe divided into three parts: Downlink pilot time slot (DwPTS), Guard period (GP) and Uplink pilot time slot (UpPTS). Therefore, the obtained information during the pilot exchange in DwPTS and UpPTS can be used towards both: generation of secret keys and channel estimation.

This ends the general background of this thesis. As discussed earlier the relative background for each chapter is provided at the beginning of the respective study. Next, Chapter 3 will show how the key-based PLS approach can be employed in a lightweight authentication protocol.

Chapter 3

Multi-factor authentication

This chapter introduces a novel PHY-based authentication mechanism. The mechanism provides provably secure and lightweight session establishment. The concept of the work presented in this chapter was published within the "Springer EURASIP Journal on Wireless Communications and Networking" [12]. The full solution is under review for publication within the "IEEE Internet of things journal".

3.1 Introduction

This study introduces the joint use of PUFs and SKG in a secure authentication mechanism. Furthermore, these mechanism are combined in a 0-RTT [126, 127] approach, allowing to build quick resumption mechanisms with forward security. The resumption protocol is important as it significantly reduces the use of the PUF to the initial authentication, thus, overcoming the limitation of a PUFs' challenge response space [128, 129].

The contributions of this chapter are as follows:

1. Proposal of a fast PUF-based authentication mechanism.
2. Combination of an initial proximity-based check with the PUF authentication as a counter measure to impersonation attacks.
3. Proposal of a forward secure 0-RTT-type of resumption protocol to quickly "resume" sessions.
4. Verification of the security properties of the protocol are provided using Tamarin prover and BAN logic.
5. Derivation of a closed form expression for the information theoretic bound of the SKG rate at the finite blocklength

This chapter proposes a novel use of the SKG and does not discuss improvements of the SKG itself. Next, in Chapter 4 the SKG mechanism is considered in detail and optimised.

Note that the following assumptions are made for the purpose of this study: i) the assumed network topology is a 1-hop communication between Alice and Bob who want the authenticate each other; ii) the main part of this study assumes that initial communication between Alice and Bob has already been established. During this initial communication, named enrolment phase, they have exchanged set of parameters, such as RSSI_0, challenge-response pairs etc (more details on the exchanged parameters are given in Section 3.4.1). The important assumption for the success of the proposed methods, within

this study, is that the enrolment phase must be carried out over a secure channel. This can be achieved by using traditional cryptographic primitives, however, this falls out of the scope of the current thesis and therefore is assumed that Alice and Bob have already performed this step; iii) Furthermore, without loss of generality, the system model is simplified assuming that neighboring subcarriers of the communication channel are assumed to be independent.

3.2 Respective background

3.2.1 Cryptographic primitives

The cryptographic process of transforming a message (plaintext) into an encoded ciphertext is called encryption. The reverse process where plaintext is obtained from a ciphertext is called decryption. The processes of encryption and decryption in this thesis are denoted as `Es` and `Ds`, respectively. To perform the transformation from plaintext to ciphertext and vice versa both algorithms rely on the cryptographic keys. The secrecy properties of the algorithms rely on the secrecy of the keys. Cryptographic systems are divided into two types: symmetric and asymmetric. In detail, asymmetric encryption uses pair of keys (public, private), one for each transformation, symmetric encryption assumes the usage of a single key in both transformations `Es` and `Ds`. The work presented in this thesis assumes the use of symmetric encryption, as the largely used block cipher advanced encryption standard (AES) is considered post quantum for key lengths of 256 bits. On the other hand, no asymmetric key encryption cipher is considered to be post-quantum. Another motivation for using symmetric encryption algorithms throughout this thesis is the fact that, they require less processing power for computations and are "*almost* 1000 *times faster than asymmetric algorithms*" [130]. This makes the use of symmetric encryption suitable for lightweight and latency aware IoT applications. The following gives a brief introduction to the cryptographic primitives used throughout this section.

Symmetric encryption

As mentioned the work presented in this thesis assumes the usage of symmetric encryption. Hence, two legitimate parties (Alice and Bob) posses identical and uniformly distributed key sequences $\{0,1\}^{|\mathbf{k}|}$ – where $|\mathbf{k}|$ is chosen such that a simple brute force attack should be impractical. An example for symmetric encryption mechanism considered to be post-quantum secure is $|\mathbf{k}| = 256$ [131]. Given that, a symmetric encryption algorithm, is denoted by $\texttt{Es} : \mathcal{K} \times \mathcal{M} \rightarrow \mathcal{C}_{\mathcal{T}}$ where $\mathcal{K}$ denotes the key space, $\mathcal{M}$ denotes the message space and $\mathcal{C}_{\mathcal{T}}$ denotes the ciphertext space. The corresponding decryption algorithm is $\texttt{Ds} : \mathcal{K} \times \mathcal{C}_{\mathcal{T}} \rightarrow \mathcal{M}$.

One of the mostly used symmetric cryptographic scheme used today is the AES [132–134]. AES was established by the U.S. National Institute of Standards and Technology (NIST) in 2001 by the name Rijndael [135]. Several sub-types of AES were standardised, *i.e.*, AES-128, AES-192, AES-256. Each of these symmetric-key mechanisms takes as input a 128-bit (16-byte) block size, but they use different key lengths: $128, 192$ and 256 bits, respectively. The key size determines the number of rounds used to transform the plaintext to ciphertext. The output size is determined by the length of the plaintext. The AES algorithms ensures that all blocks are with 16-byte size even if the length of plaintext is not exact multiple of 16 bytes. This is achieved by using padding mechanisms and it is done to prevent adversaries from knowing the length of the message [136].

Message authentication code

A message authentication code (MAC) is additional information sent with ciphertext to verify the authenticity of the message [137]. The MAC can be constructed by a pair of algorithms, one for signing, `Sign`, and one for verification, `Ver`. The two algorithms rely on a pre-shared secret key $\mathbf{k}_{\texttt{MAC}}$ between the communicating parties and ensures integrity and authenticity of the message $\mathbf{m}$. While data integrity assures, the accuracy and consistency of data, *i.e.*, it has not been modified by a malicious third party, authenticity verifies the identity of the sender. This is achieved by applying the MAC scheme as

follows: first, the sender signs the message using $\mathbf{k}_{\mathtt{MAC}}$, then the intended receiver must confirm that the signed message corresponds to $\mathbf{m}$. A reason for verification failure could be the manipulation of the message while in transit or receiving a message not signed from a malicious node.

Following from the above a pair of message authentication code (MAC) algorithms is denoted by $\mathtt{Sign} : \mathcal{K}_{\mathtt{MAC}} \times \mathcal{M} \rightarrow \mathcal{T}$, where $\mathcal{T}$ denotes the space of signed messages, with a corresponding verification algorithm $\mathtt{Ver} : \mathcal{K}_{\mathtt{MAC}} \times \mathcal{M} \times \mathcal{T} \rightarrow (yes, no)$. Further information about standardised MAC algorithms can be found in [138].

Hash function

Hash functions are used to transform messages with various length into fixed output hash value. The study presented in this thesis assumes the use of one-way hash functions [139, 140]. One-way hash functions are irreversible, *i.e.*, having the initial message one can obtain the corresponding hash value, however, if only the hash value is available one cannot obtain the original message. Furthermore, a one-way hash function should satisfy collision-free property. This means two different inputs should not produce the same hash value. The collision-free property is denoted as follows:

$$\mathtt{Hash}(x) \neq \mathtt{Hash}(y). \tag{3.1}$$

The output of one-way hash functions is pseudorandom meaning that outputs are independent and one should not be able to distinguish a hash value from a random sequence. Finally, one-way hash functions are used to protect data from modification. For example they are used in the Keyed-Hash Message Authentication Code (HMAC) [141] which is part of security protocols such as SSH [142].

3.2.2 Physical unclonable functions

The idea of physical one-way function (POWF) was introduced in [143]. The authors proposed a simple challenge-response authentication protocol. It assumes that an identi-

fier (*e.g.* server, that will authenticate a user) generates a set of challenges which are run on a user's equipment. These challenges are then stored by the identifier along with the corresponding set of responses, *i.e.*, as challenge-response pairs (CRPs). Next, when the user sends an authentication request to the identifier chooses one of the previously stored CRPs and "challenge" the user to produce a response. The authentication is successful, if the response is identical to the one previously stored in the identifier's database. The structure proposed in [143] is based on the usage of a laser beam as input of their system. This beam propagates through a three-dimensional micro-structure and the final resulting pattern determines the output of the system. A challenge to this scheme could define the angle of the transmitted beam.

Later, the authors of [144] brought this idea to integrated circuit (IC) structures by introducing the silicon physical unclonable function (PUF). Its idea is to utilise the fact that every integrated circuit differs to others due to manufacturing variability [52, 145] and cannot be cloned [54]. A challenge to this scheme can refer to delay at each gate, power-on state and other variable features. Nowadays, due to its randomness and unclonable properties, PUFs continues to attract the researches' attention and since the introduction of PUFs in [144], numerous PUF architectures have been proposed many of which suitable for IoT applications. A few of these architectures are: arbiter PUF [146], ring oscillator PUF [39], transient effect ring oscillator PUF [147], static random-access memory PUF [148], hardware embedded delay PUF [149] and more [53, 128].

As mentioned above, each PUF can be used as a challenge – response scheme, meaning that a PUF takes a challenge as input, runs it and outputs the corresponding response. Depending on the number of CRPs that each PUF can support, two types of PUFs can be differentiated: i) strong PUFs, suitable for authentication and ii) weak PUFs, which can be used for secret key generation. Intuitively, a weak PUF can process a small number of challenges, while a strong PUF is capable of processing a large set of CRPs. Further requirement for a strong PUF is backward secrecy, meaning that if some of the previous CRPs are revealed, an attacker should not be able to predict the response of any future challenges. Throughout this thesis PUFs will refer to strong PUFs from this point on-

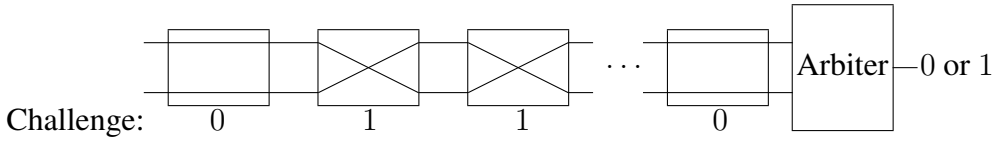

Figure 3.1: Arbiter PUF.

wards.

To familiarise the readers with the basics of PUFs, two structure examples are presented. The structure of *Arbiter* PUF was presented in [146]. In this scheme the authors propose the usage of the delay variations of wires and transistors within integrated circuits. The idea is based on transmitting rising edge signals through two "identical" delay paths which consist of several switching elements. Due to variation properties the delay will slightly differ at each path. Finally, an arbiter determines the corresponding "winner", *i.e.*, the signal that arrived first at the end of the trace. The produced response (output bit) depends on the "winner". A challenge to this scheme is composed by choosing the configuration of each switching element as illustrated on Fig. 3.1. Another PUF scheme, called *recombined oscillator* PUF, was proposed in [150]. The scheme calculates frequency differences between pairs of oscillators which produces set of random variable. The challenge to this scheme determines the sign in front of each random variable. To produce the response to a specific challenge the scheme calculates the sum of all random variables and compares it to a pre-defined threshold. Similarly, [151] proposed an improved scheme which increases the secrecy of the output by choosing a different subset of the oscillators each time (defined by the challenge).

Following from the above, a typical PUF-based authentication protocol consists of two main phases, namely *enrolment phase* and *authentication phase* [152–156].

During the *enrolment phase* each user runs a set of challenges $\mathcal{C}$ on his PUF and characterises the variance of the measurement noise in order to generate helper data $\mathbf{hd}$. Next, a verifier creates and stores a database of all CRPs $(\mathcal{C}, \mathcal{R})$ for each user's PUF within its network. A CRP pair in essence consists of an authentication key and related helper

data $\mathbf{hd}$. Within the database, each CRP is associated with the ID of the corresponding user. The enrolment phase is a one-time operation and it is assumed to be done on a secure channel. Later, during the *authentication phase* an user sends its ID to the verifier requesting to start a communication. Receiving the request, the verifier checks if the received ID exists in its database. If it does, the verifier chooses a random CRP (C_i, R_i) that correspond to this ID and sends C_i to the user. The user computes the response R_i by running the challenge C_i on its PUF, as follows $R_i = \texttt{PUF}(C_i)$, and sends R_i to the verifier. However, the PUF measurements at the user are never exactly the same due to measurement noise, therefore, the verifier uses the new PUF measurement and the helper data $\mathbf{hd}_{R,i}$ stored during the enrollment to re-generate the authentication key. Finally, the verifier compares the re-generated key to the one in the CRP and if they are identical the authentication of the node is successful. In order to prevent replay attacks, once used, a CRP is deleted from the verifier database.

A widely used method for helper data and key generation is by employing fuzzy extractors (FE). FE are method that allows for key derivation from nonuniform noisy sources [157]. Due to their properties they are referred as a common technique to cope the noise present in biometric data and PUF measurements [158–162]. A so called (m, l, t, ϵ) FE is built from a pair of randomised functions, namely *Generate* `Gen` and *Reproduce* `Rep`. The function `Gen` requires a single input $R \in \mathcal{R}$ and produces two outputs, helper data $\mathbf{hd}_R$ and a key $\mathbf{k}_R \in \{0,1\}^l$. The function `Rep` is deterministic reproduction and can reproduce the key $\mathbf{k}_R$ by using the helper data $\mathbf{hd}_R$ and a sequence closely associated to R, *i.e.*, R'. The correctness of the `Rep` function can be satisfies if only $D(R, R') \leq t$ (D refers to Hamming distance). On the other hand, when $\mathbf{hd}_R$ is public the key sequence $\mathbf{k}_R$ is close to uniformly random only if the min-entropy of R is at least m. This implies the following condition on the statistical distance between $\mathbf{k}_R$ and the universe of sequences with length l, whenever $\mathbf{hd}_R$ is public $SD((\mathbf{k}_R, \mathbf{hd}_R), (\mathcal{U}_l, \mathbf{hd}_R)) \leq \epsilon$. More details about FEs' parameters and bounds can be found in [157, 163, 164].

To summarise, as it is possible to construct a PUF that does not require any complex operations, it can be used as a lightweight mechanism to generate an unclonable and ran-

dom secret. However, to guarantee a secure transmission of challenges and responses one cannot send these in clear text and risk to be obtained by an adversary. Therefore, further security mechanisms have to be employed to create a secure authentication protocol based on PUF. Many examples of PUF-based authentication protocols have been proposed in the literature: some for unilateral authentication [165, 166] and some for mutual authentication [166–169]. A comprehensive survey on lightweight PUF authentication schemes is presented by Delvaux *et al.* [55].

3.2.3 0-RTT protocols

This section briefly describes the 0-RTT authentication mode introduced in the transport layer security (TLS) 1.3 protocol [126]. The use of 0-RTT obviates the need of performing a full authentication procedure for every re-authentication through the use of a resumption secret $\mathbf{z}$, thus reducing latency. The TLS 1.3 0-RTT handshake works as follows: in the very first connection between client and server a regular TLS handshake is used. During this step the server sends to the client a look-up identifier $\mathbf{k}_l$ for a corresponding entry in session caches or it sends a session ticket. Then both parties derive a resumption secret $\mathbf{z}$ using their shared key and the parameters of the session. Finally, the client stores the resumption secret $\mathbf{z}$ and uses it when reconnecting to the same server which also retrieves it during the re-connection.

If session tickets are used, the server encrypts the resumption secret using a long-term symmetric encryption key, called a session ticket encryption key (STEK), resulting in a session ticket. The session ticket is then stored by the client and included in subsequent connections, allowing the server to retrieve the resumption secret. Using this approach the same STEK is used for many sessions and clients. On one hand, this property highly reduces the required storage of the server, however, on the other hand, it makes it vulnerable to replay attacks and not is not *forward secure*. Forward secrecy assures that past sessions will remain secret even if current or/and future secret keys are compromised [170]. Due to these vulnerabilities, the work within this chapter focuses on the session cache mechanism

described next.

When using session caches the server stores all resumption secrets and issues a unique look-up identifier $\mathbf{k}_l$ for each client. When a client tries to reconnect to that server it includes its look-up identifier $\mathbf{k}_l$ in the 0-RTT message, which allows the server to retrieve the resumption secret $\mathbf{z}$. Storing a unique resumption secret $\mathbf{z}$ for each client requires server storage for each client but it provides forward security and resilience against replay attacks, when combined with a key generation mechanisms such as Diffie Hellman (or the SKG used in this thesis) which are important goals for security protocols [127].

3.2.4 Proximity detection

The IoT is becoming an important part in human machine interaction. Its applicability varies from healthcare [171] to mass-market applications [172]. A vital feature that will help enable such applications is indoor localisation. Particularly, IoT indoor localisation-based services are expected to operate without human intervention [173]. Localisation techniques can be divided into two groups: database-matching methods and geometrical methods [174]. Database-dependent methods rely on pre-stored location information, which is treated as a fingerprint map of the area. The localisation process is based on a real-time comparison between the collected information to the pre-stored one. This type of methods determines the fingerprint, within the map, in which the user belongs. Geometrical methods are based on real-time computation of the user's location. The location is determined by measuring distances or angles. Some of the widely used geometrical methods are: time-of-flight, multilateration, multiangulation and proximity detection [173, 174].

While time-of-flight, multilateration and multiangulation are capable of accurate localisation they usually require high complexity operations or high cost equipment. On the other hand, proximity detection has the lowest computational complexity, and could be easily implemented with the equipment already present in smartphones today. Proximity detection is a distance based method which does not provide exact location, instead it

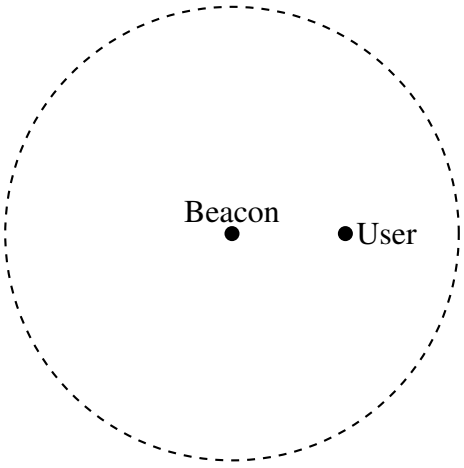

Figure 3.2: Proximity detection

can be used to estimate the distance to a transmitting beacon [175]. As illustrated in Fig. 3.2 by using this method one could measure the distance to a nearby beacon but cannot determine the direction to it.

Due to its simplicity and ease of implementation proximity detection is used for the purpose of the study within this chapter as a soft authentication mechanism along with PUFs. A widely adopted technique to obtain the distance to a user (the radius in Fig. 3.2) is by using the RSS. However, numerous factors could negatively impact the wireless communication in an indoor scenario. Such are reflection, diffraction, scattering, slow and fast fading. All of the above have direct impact on the RSS. To overcome these issues one must use a filtration algorithm to treat noisy measurements. A popular filter applied for localisation is the, so called Kalman filter [176]. It works by the assumption that the current state $\mathbf{x}_i$ has relation to the previous one $\mathbf{x}_{i-1}$, and this relation is expressed as follows:

$$\mathbf{x}_i = \mathbf{A}\mathbf{x}_{i-1} + \mathbf{B}\mathbf{u}_{i-1} + \mathbf{w}_{i-1}, \tag{3.2}$$

where $\mathbf{A}$ is a transition matrix which links the current state with the previous one, $\mathbf{B}$ is a control matrix which relates the control vector $\mathbf{u}$ to the state and $\mathbf{w}$ is an i.i.d. normally distributed process noise such that $\mathbf{w} \sim N(0, \mathbf{I}\sigma_Q)$, which represents factors such as velocity change, wind etc. Following that, the measurements of the current state are given

by:

$$\mathbf{z}_i = \mathbf{G}\mathbf{x}_i + \mathbf{v}_i, \tag{3.3}$$

where $\mathbf{G}$ is the observation matrix used to translate each state into a measurement and $\mathbf{v}$ is an i.i.d. normally distributed measurement noise such that $\mathbf{v} \sim N(0, \mathbf{I}\sigma_R)$, which denotes the noise present in the measurements, due to fading, path-loss etc. Given the above, the recursive process of the filter is presented on Fig. 3.3. It is based on two main steps: prediction and correction (time and measurement update, respectively). During the time update step: i) the next state is updated based on the previous one; and, ii) the error covariance matrix $\mathbf{P_e}$ is updated based on the previous one. In the above $\hat{\mathbf{x}}_i^-$ and $\mathbf{P_e}_i^-$ are *a priori* estimated state and *a priori* error covariance matrix, *i.e.,* predictions, which are estimated based on the previous instant. In fact the error covariance matrix is a measure of uncertainty of the estimated state $\hat{\mathbf{x}}_i^-$, due to process noise. Next, during the measurement update step the filter uses the *a priori* estimates calculated in the prediction step and updates them to find their *a posteriori* values: i) the so called Kalman gain $\mathbf{K_G}$ is computed such that it minimizes the *a posteriori* error covariance, *i.e.*, it determines the weight of the measurement $\mathbf{z}_i$ and the *a priori* estimate $\hat{\mathbf{x}}_i^-$, such that if the measurement noise is low the measurement will contribute more for the calculation of the *a posteriori* state, while if the the error in the *a priori* estimate is low it will be trusted more during the measurement update step; ii) using the Kalman gain estimates $\hat{\mathbf{x}}_i$ and $\mathbf{P}_{e,i}$ are updated following the equations from the measurement update step given in Figure 3.3.

Finally, it has been shown that the characteristics of a fading channel follow a log-normal distribution and a commonly used path loss model demonstrated through measurements is [177–180]:

$$RSSI(d) = RSSI_0 - 10n \log\left(\frac{d}{d_0}\right) + X_\sigma, \tag{3.4}$$

where $RSSI(d)$ is the path-loss (average received signal strength) at distance d, $RSSI_0$ represents the average received signal strength at some reference distance d_0, n is an attenuation factor that gives the relation between distance and received power, its values

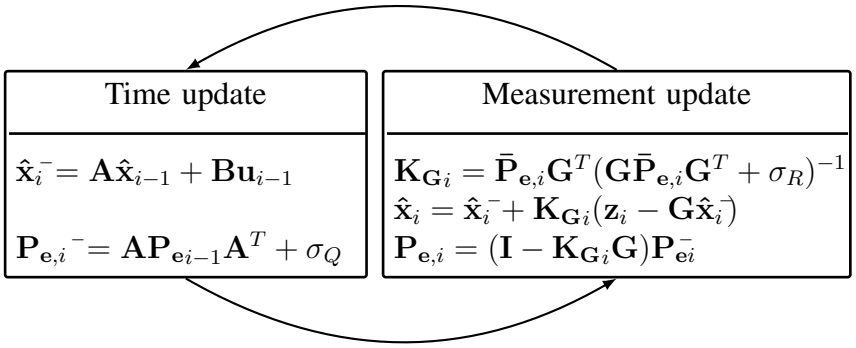

Figure 3.3: Kalman filter steps

ranges from 2 to 6, X_σ is a zero mean Gaussian random variable which captures the variations in the received power with standard deviation (smaller the deviation higher the precision of the model). Typical values of n and and the standard deviation of X_σ for different indoor environment are summarised in [87]. To simplify the model, typically, the reference distance d_0 is chosen to be 1m, hence Eq. (3.4) becomes:

$$RSSI(d) = RSSI_0 - 10n\log(d) + X_\sigma. \tag{3.5}$$

To conclude, using the Kalman filter presented on Fig. 3.3 and the path-loss model 3.5, one can build a proximity detection mechanism. Such a mechanism later in this chapter as part of an authentication process.

3.2.5 Security verification

Security protocols are used to ensure properties such as forward secrecy, anonymity, untraceability, *etc.* To avoid possible threats the designing process of such protocols must be infallible. Moreover, all security flaws cannot be identified through testing, since some may happen only in the presence of adversary [181]. Therefore, security verification is essential step before protocol deployment. The work within this chapter is verified through both BAN logic [182] and Tamarin prover [42]. For the sake of presentation the introduction to Tamarin is given in Appendix A.

BAN logic

The secrecy evaluation of security protocols ensures that an adversary cannot obtain or alter secret parameters. In this regards, the logic proposed by Burrows, Abadi and Needham (BAN) [182] is widely used secrecy verification tool. However, some weaknesses were identified and several works, such as [183–186] proposed extended and more reliable versions. For the purpose of security verification of protocols this thesis assumes the usage of Mao and Boyd (MB) logic [186].

Formal proofs are deduced using set of initial beliefs and rules and are based upon the message exchange within the protocol. The protocol analysis using MB logic consists of 3 steps: i) the protocol definition is converted to idealised form which includes syntax and logic interpretation; ii) the idealised version of the protocols must follow several definitions and rules, which will be described later; and, iii) the manipulation begins with set of initial assumptions that specify the protocol and are used to derive set of conclusions. As an example when analysing a key-distribution protocol, one may aim for the conclusion $A \overset{k}{\leftrightarrow} B$. The meaning of this expression is: k is a good shared secret between A and B. The full notation of MB logic is described later in this section.

The protocol message idealisation is used to interpret the implicit context-dependent information into explicit protocol specification. As mentioned, messages and interactions within a protocol should be idealised using several definitions and rules:

Definitions:

1. Atomic message: a piece of data within the protocol that constructs a message without using any of the following symbols: “,”, “|”, “$\mathfrak{R}$”, “()”, “{}”.

2. Challenge: a non timestamp atomic message that is sent by an agent (its originator) on one line within the protocol specification and received by the same agent on different line.

3. Replied challenge: a challenge that appears in a line intended to the challenge’s originator.

4. Response: a non timestamp atomic message and a replied challenge sent together.

5. Nonsense: a non timestamp, challenge or response atomic message.

Rules:

1. Nonsenses are removed.

2. Atomic messages that appear to be a challenge and a response in the same line is considered to be a response.

3. Challenges which are separated by comma are combined using "|" operator.

4. Responses which are separated by comma are combined into a combined response using "|" operator.

5. Challenge and its response are combined using "$\mathfrak{R}$" operator.

6. Message and a corresponding timestamp are combined using "$\mathfrak{R}$" operator.

In order to illustrate the meaning of the definitions above a simple example of 2-party communication is considered here:

1. $\mathsf{Alice} \rightarrow \mathsf{Bob} : ID_A, m, \mathsf{nonce}_A, TS_A$
2. $\mathsf{Bob} \rightarrow \mathsf{Alice} : ID_B, \mathsf{nonce}_B, \{\mathsf{nonce}_A, TS_B, k_2\}_{k_1}$
3. $\mathsf{Alice} \rightarrow \mathsf{Bob} : \{\mathsf{nonce}_B\}_{k_2}$

Note that $\{\cdot\}_k$ defines encryption using key k. In the example above, nonce_A is a challenge in the first message and it is a replied challenge in the second. Similarly, nonce_B takes the role of a challenge in the second message and the role of a replied challenge in the third. Next, k_2 is a response to challenge nonce_A. Finally, TS_A and TS_B are timestamps, whereas, m is a nonsense within this example. Following the rules, the idealised version

$A \models X$	A believes X is true
$A \overset{k}{\triangleleft} m$	A sees m using key k, if not encrypted $A \triangleleft m$
$A \overset{k}{\mid\sim} m$	A encrypts m using key k
$\#(m)$	m is of type fresh
$A \overset{k}{\leftrightarrow} B$	k is a good shared key between A and B
$A \triangleleft \mid\mid m$	m is not available to A
$\text{sup}(S)$	S is a super-principal

Figure 3.4: Notation of the MB logic for protocol analysis.

of the communication protocol defined above is given as:

1. $\mathsf{Alice} \rightarrow \mathsf{Bob} : ID_A, \mathsf{nonce}_A \mathfrak{R} TS_A$
2. $\mathsf{Bob} \rightarrow \mathsf{Alice} : ID_B, \mathsf{nonce}_B, \{k_2 \mathfrak{R} \mathsf{nonce}_A \mathfrak{R} TS_B\}_{k_1}$
3. $\mathsf{Alice} \rightarrow \mathsf{Bob} : \{\mathsf{nonce}_B\}_{k_2}$

The following differences can be seen in the idealised version compared to the initial one: the nonsense m is removed from the first message; again in the first message nonce_A is combined with the corresponding timestamp TS_A using the "$\mathfrak{R}$" operator; in the second message the operator "$\mathfrak{R}$" combines the response k_2 with its corresponding challenge nonce_A which on the other side is combined with the timestamp TS_B.

Following the above rules and definitions security properties are modelled through inference rules using the notation in Fig. 3.4. Furthermore denoting: principals as A, B; messages as m; keys as k; formulas as X and Y; the inference rules used for the purpose of this thesis are:

Authentication rule: $\frac{A \models A \overset{k}{\leftrightarrow} B \wedge A \overset{k}{\triangleleft} m}{A \models B \overset{k}{\mid\sim} m}$ meaning, if k is a good shared key between A and B (*i.e.*, it has not been leaked) and A uses key k to decrypt m, then A can believe that B is the one who encrypted m.

Confidentiality rule: $\frac{A \models A \overset{k}{\leftrightarrow} B \wedge B^C \triangleleft \mid\mid m \wedge A \overset{k}{\mid\sim} m}{A \models (A \cup B)^C \triangleleft \mid\mid m}$ meaning, if k is a good shared key between

A and B and the complement of set B cannot see m and A used k to encrypt m, then A believes that the complement of the union between A and B cannot see m, resulting in m can be seen only by A and B.

Fresh rule: $\frac{A|\equiv\#(m)\wedge A\triangleleft n\mathcal{R}m}{A|\equiv\#(n)}$ meaning, if m is of type fresh and n is a response to m follows that n is also of type fresh.

Good-key rule: $\frac{A|\equiv\{A,B\}^{C}\triangleleft||k\wedge A|\equiv\#(k)}{A|\equiv A\overset{k}{\leftrightarrow}B}$ meaning if A and B are the only agents that can see k and k is of type fresh, follows that k is a good shared key between A and B.

Nonce verification rule: $\frac{A|\equiv\#(n)\wedge A|\equiv B|\overset{k}{\sim}n}{A|\equiv B|\equiv A\overset{k}{\leftrightarrow}B}$ meaning if B used the key k in the current protocol run follows that B believes k is a good shared key.

Super-principal rule: $\frac{A|\equiv B|\equiv X\wedge A|\equiv\sup(B)}{A|\equiv X}$ meaning A unconditionally trusts B beliefs, in this example X, and this is true as far as A thinks B is the super principal w.r.t. X.

Belief axiom 1: $\frac{A|\equiv X\wedge A|\equiv Y}{A|\equiv(X\wedge Y)}$ meaning A believes X and A believes Y, therefore, A believes the set of functions (X, Y).

Belief axiom 2: $\frac{A|\equiv X\wedge A|\equiv X/Y}{A|\equiv Y}$ meaning A believes X and A believes X implies Y, therefore A believes Y.

3.3 Employed methods and system model

Having now presented the necessary background material, the rest of this chapter proposes a full authentication solution based on a range of basic primitives. Each of these primitives is introduced below together with a summary of the methods used to analyse and optimise the solution. The system model assumed in this chapter consists of two legitimate parties Alice and Bob, where Alice refers to a resource-constrained IoT device and Bob to a resourceful server.

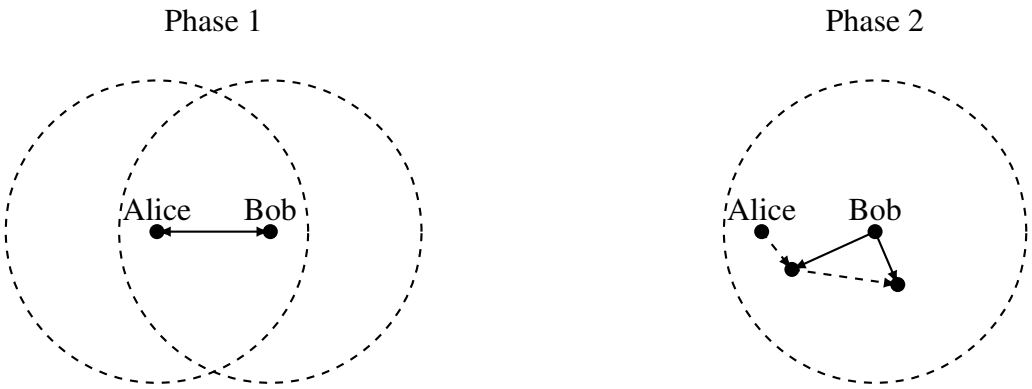

Figure 3.5: Enhanced proximity detection - phase 1: Alice and Bob perform proximity detection; phase 2: Alice performs two additional measurements at different locations.

Proximity detection

It is expected that future IoT devices will be able to obtain, or even generate a map of an indoor premise by simply entering the environment [187, 188]. This ability will enhance the performance of methods such as: signal level prediction, on-the-fly location estimation and numerous mobile applications [187, 189]. Furthermore, many mobile IoT devices in scenarios such as smart home, industry and robotics are stored in a specific location while not in use. Following from the above, this work proposes an authentication protocol that relies upon a proximity detection (discussed in Section 3.2.4) to add a second factor of authenticity. Before executing the authentication procedure, Alice and Bob roughly estimate each others location. On one hand, Bob (static server) checks whether Alice is within a expected area, while on the other hand Alice (mobile IoT node) performs an enhanced proximity detection as illustrated in Fig. 3.5. The procedure comprises of two phases: 1) This phase is based on a pre-defined distance threshold ζ. Before running the authentication procedure both parties compare ζ to online measured distance and confirm whether the other party is within the expected area or has been compromised and moved. Given the above, the verification mechanism can be built as follows:

$$\text{Position} = \begin{cases} \text{Legitimate}, & \text{if distance} \leq \zeta \\ \text{Compromised}, & \text{otherwise.} \end{cases} \tag{3.6}$$

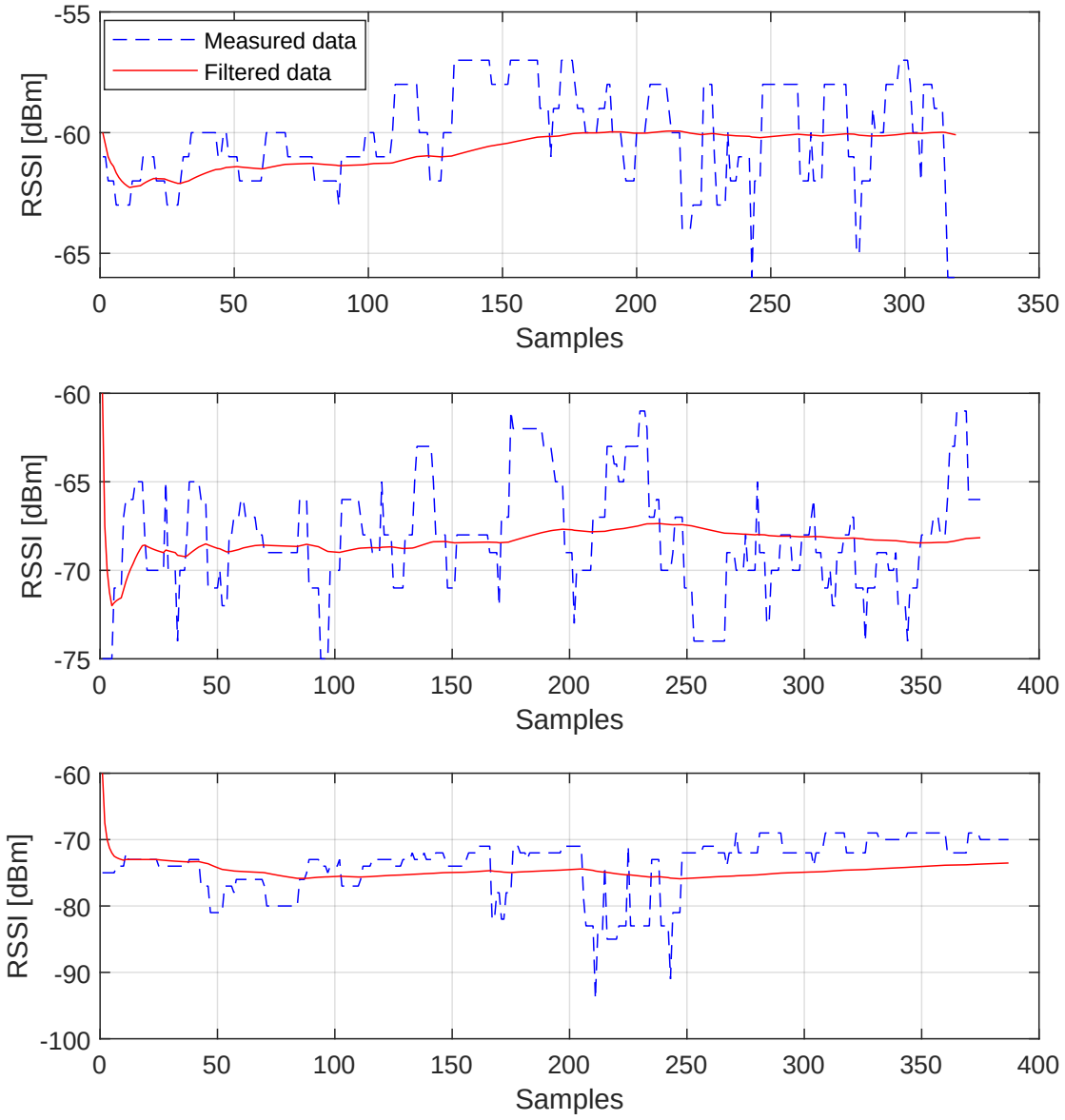

Figure 3.6: Measured RSSI data (dashed) and filtered data using Kalman filter (solid) for three different distances: TOP 1 meter, MIDDLE 3 meters, BOTTOM 6 meters. The measurement noise variance is set to $\sigma_R = 0.1$.

2) Assume that a mobile IoT device begins to move in a unpredictable manner that an adversary cannot foresee. The IoT device performs two additional RSS measurements using a downloaded map of the premise in order to confirm the position of the server. The additional measurements are performed to prevent the IoT from impersonation attacks.

Next, to validate the practicality of the proposed method, a set of experiments were

performed in an indoor environment using a Bluetooth low energy (BLE) beacon that transmits with power of -4 dBm and a smartphone to collect the signals and measure the RSS (which in the BLE protocol is denoted as received signal strength indicator (RSSI)). As described in Section 3.2.4, to ensure the received signal is reliable, noise and variations introduced of the wireless channel are filtered using a Kalman filter. Due to the fact the RSSI is a scalar variable the process does not contain matrices. Therefore, the parameters of the filter are as follows: initial prediction of the RSSI value $\hat{x}_{i-1}$, initial prediction of the error variance $P_{\mathrm{e},i-1} = 0.1$, since the process is controlled, the variance of the process noise is assumed to be nearly zero $\sigma_Q = 10^6$; the measurement noise varies from environment to environment, hence σ_R must be adjusted depending upon the environment; since no any control signals are used, u equals zero; the obtained RSSI values are in dBm and do not need to be converted implying $G = 1$; furthermore, since the measurements in the specific scenario are for a static position the transition matrix $A = 1$. Following from this, the measurements collected during three different experiments are illustrated on Fig. 3.6. The experiments differ by the distance between the two devices - 1 m (TOP graph), 3 m (MIDDLE graph) and 6 m (BOTTOM graph). Each graph illustrates the RSSI values with dashed line and the resulting filtered values with solid line. The measurement noise variance for all three experiments was estimated as $\sigma_R = 0.1$. It should be noted that during the experiment the line of sight between the two devices was not always present due to people moving in the area, additionally there were other Bluetooth and wifi devices working in the vicinity which causes further interference. To show how quickly the filter will find the pattern of the data the initial prediction is set at $\hat{x}_{i-1} = -60$ for all three experiments. It can be seen in Fig. 3.6 that the filter quickly converges to the desired value (by about the third sample). The filtered output was found to be stable after convergence in all tested scenarios, showing that the filter successfully eliminates the noise present in the measurements. Based on the results in Fig. 3.6, three regions for the filtered RSSI values, were identified: $[-62, -59]$dBm when at 1 m; $[-70, -67]$dBm at 3 m; $[-72, -76]$dBm at 6m. This allows the enhancement of the proximity detection discussed in Section 3.2.4. Note that the results show that there is a relation between RSSI and distance, however,

the proposed method in its current form cannot provide high accuracy, therefore, further experimental validation needs to be performed before employing the method in a real system setting.

Finally, the novelty within this method can be summarised as follows: 1) proposing the usage of proximity detection as a soft, second factor, authentication method 2) development of an enhanced proximity detection mechanism that can be performed by mobile IoT nodes; 3) promising experimental results of the proposed method in a real-life setting.

PUF authentication

Before establishing a shared secret key, Alice and Bob must be sure they are communicating with a trusted party. This work assumes the usage of a PLS method, and more specifically PUF authentication. As discussed in Section 3.2.2, many PUF authentication protocols have been proposed in the literature, with even a few commercially available [190, 191]. Furthermore, by eliminating the need of non-volatile memory the usage of PUFs could greatly reduce the complexity compared to existing authentication alternatives. This work does not look into developing a new PUF architecture, instead, it proposes a new fast PUF-based authentication protocol that relies on existing PUF architectures. Finally, the contributions in relation to this method are: 1) development of a new PUF-based lightweight authentication protocol; 2) using PUF in conjunction with PHY SKG.

Fuzzy extractor

In a PUF setting, a fuzzy extractor (FE) can be used to correct discrepancies in a reproduced responses. The FE allows authentication of a device, equipped with PUF, by comparing a generated key from a response, R_i, to a key generated from a noisy version of the response, R_i', as follows:

$$\texttt{Gen}(R_i) = (\mathbf{hd}_{R,i}, \mathbf{k}_{R,i}) \tag{3.7}$$

$$\texttt{Rep}(R'_i, \mathbf{hd}_{R,i}) = \mathbf{k}_{R',i} \tag{3.8}$$

By confirming the condition $\mathbf{k}_{R,i} \stackrel{?}{=} \mathbf{k}_{R',i}$ one can confirm that the responses R_i and R'_i were generated from the same PUF. Generally, the `Gen` function is executed on the server and `Rep` on a remote device equipped with the PUF. However, the `Rep` function is more computationally complex than the `Gen` and consequently a reverse FE has been proposed [192]. In reverse FE, the more complex operation is performed on the resourceful device (Bob) instead on the resource constrained device. Furthermore, this can help the prevention of a helper data manipulation attack, *i.e.* in the reverse FE, it is the server that observes the helper data and is thus better placed to detect a manipulation of the data than a resource constrained IoT device [193]. The use of FE in a PUF setting is not a novel approach, however, FE is a necessary tool to eliminate the noise present in PUF measurements. Therefore, FEs are used throughout this chapter.

Secret key generation

To ensure that their communication is private, after authenticating each other, Alice and Bob have to encrypt / decrypt the exchanged data. For this work, it is assumed the use of symmetric encryption where the same key is used for both operations. In order to obtain a shared key this work proposes to use SKG which consists of three standard steps: i) advantage distillation; ii) information reconciliation; and, iii) privacy amplification; each of these steps are explained in more detail in Section 2.2.1. This works assumes short block-length communication in an additive white Gaussian noise AWGN channel. Therefore, an initial contribution in regards to this method is the derivation of the closed form expression for the SKG rate at the finite block-length. The derived information theoretic bound was used by my colleague, Dr. Mahdi Herfeh, to compare different short block-length Slepian Wolf key reconciliation approaches. While the numerical results, which are illustrated in Fig. 3.7, were produced by Mahdi, all derivation below was the work of the author of this thesis and is included here and in Appendix B.

For the purpose of the derivation the channel model from Section 2.2.1 is assumed:

$$x_i = h_i + z_i \tag{3.9}$$

where $z_i \sim N(0,1)$ is noise variable and h_i represents the received symbols which has variance P. Following this model, the maximum secret key rate over which Alice and Bob agree on with probability greater than $1-\epsilon$ is [194]:

$$S(n,\epsilon) = C_{SKG} - \sqrt{\frac{V}{n}}Q^{-1}(\epsilon) + \mathcal{O}(\sqrt{n}). \tag{3.10}$$

with terms as described below. Here, Alice and Bob observe blocks x_A and x_B, respectively, of length n, hence, the SKG capacity is (as derived in Section 2.2.1):

$$C_{SKG} = \mathsf{I}(x_A, x_B) = \log\left(1 + \frac{P\sigma^2}{2 + \frac{1}{P\sigma^2}}\right), \tag{3.11}$$

where unit noise variance is assumed. The dispersion of the channel is given as:

$$V = \mathsf{Var}\left[\log\frac{\mathsf{Pr}(x_A, x_B)}{\mathsf{Pr}(x_A)\mathsf{Pr}(x_B)}\right]. \tag{3.12}$$

Finally, $Q^{-1}(\cdot)$ is the inverse of the Gaussian Q function, ϵ is the average block error rate, and $\mathcal{O}(\sqrt{n})$ comprises the remainder terms of order $\sqrt{n}$ which is negligible w.r.t. to the other terms. For the purpose of this study, the closed-form of the dispersion V is evaluated resulting into

$$V = 8aP(aP + a + bP + b) + 2b^2P(P+1) + 4a^2 + b^2 \tag{3.13}$$

with:

$$a = \frac{\left(-\sqrt{P+1} - \sqrt{P+1}P + 1 + 2\,P\right)}{2\,(1+2\,P)\,\sqrt{P+1}}, \tag{3.14}$$

$$b = \frac{P}{1+2P}. \tag{3.15}$$

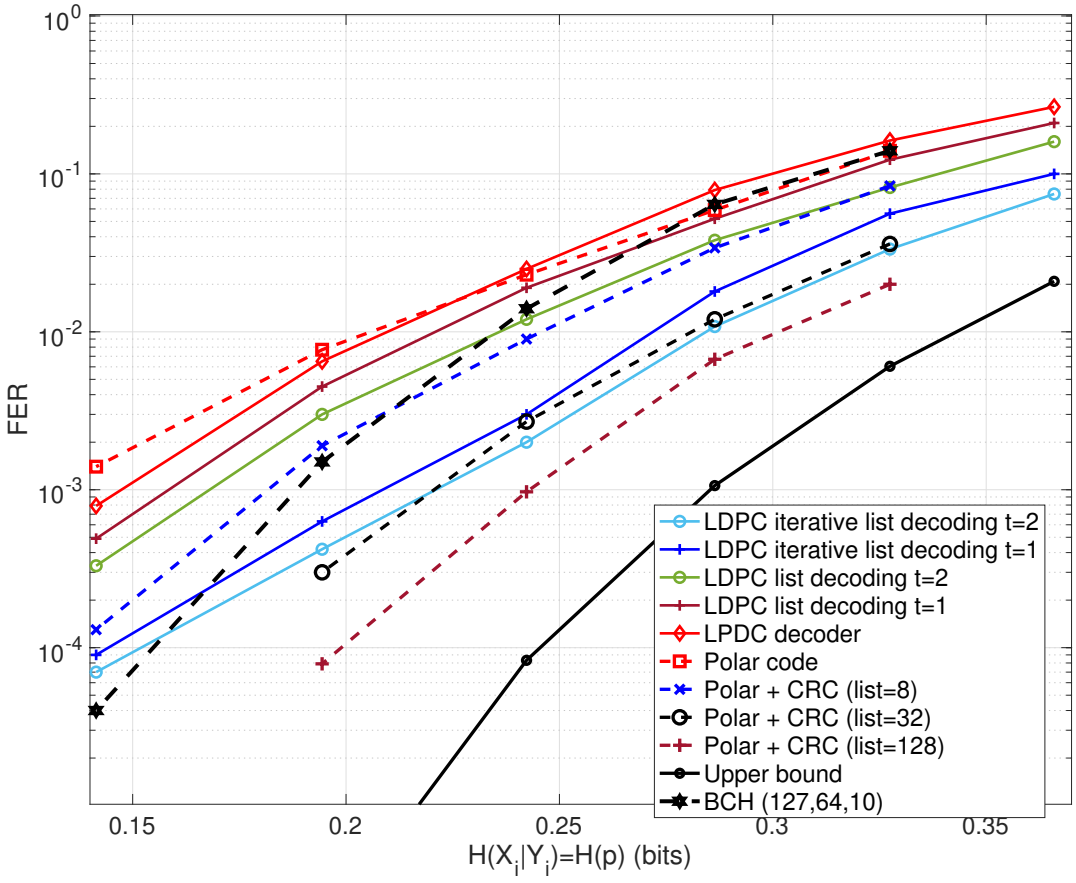

Figure 3.7: Comparison of FER performance over the reconciliation rate of different coding schemes to the upper bound given in Eq. (3.10) using the derivation of the channel dispersion as in Eq. (3.13).

All steps of the derivation from Eq. (3.12) to Eq. (3.13) can be found in Appendix B.

Next, Fig. 3.7 compare different reconciliation coding techniques used today with the estimated bound. Overall, it can seen that most of these coding techniques do not perform well for short block length and are away from the bound. It can be seen that, classical polar codes and LDPC codes do not perform well in the short block length, however, their performance can be significantly improved by using them with list and iterative list decoding, respectively. Note that, to better utilise the spectrum, future IoT technologies are expected to exchange short messages, and therefore, research in the area of short block length communication is sufficient.

Finally, the novel ideas in regards to this method are: 1) the use of SKG in conjunction with PUFs and building a full solution for lightweight authentication; 2) the development of a PHY-based resumption protocol; 3) derivation of a closed form expression for the SKG rate at the finite block-length.

Re-authentication

In Section 3.2.2 it was discussed that using PUF authentication can greatly reduce the computational overhead of a system. Authentication of new keys is required at the start of communication and at each key renegotiation. However, the number of challenges that can be applied to a single PUF is limited. Therefore, to further develop the hybrid crypto-system this work proposes a solution in the form of resumption protocol. The proposed re-authentication approach exploits the use of resumption secrets as used in the 0-RTT authentication mode in the transport layer security (TLS) 1.3 protocol. Instead of performing full authentication before sending data encrypted with a new key, this study proposes a new method which allows Alice (Bob) to authenticate subsequent keys using a lightweight scheme anchored by the initial authentication process. In this physical layer 0-RTT, given that a node identifier state would be required for link-layer purposes, the session cache places little comparative load and thus is the mechanism proposed here for (re-)authentication. Another strong motivation for developing this mechanism is that it is forward secure in the scenario used in this thesis [127]. As discussed earlier, employing the properties of the wireless channel to built a resumption protocol is a novel approach, to the best knowledge of the author of this thesis.

Employed security primitives

In addition, to eliminate the possibility of tampering attacks, this work uses a set of security primitives:

SKG scheme (as described in Section 2.2.1) is denoted by $\mathsf{G} : \mathbb{C} \rightarrow \mathcal{K} \times \mathcal{S}$, accepting as input the fading coefficients (modelled as complex numbers), and generating as outputs

binary vectors $\mathbf{k}$ and $\mathbf{s}_A$ in the key and syndrome spaces, of sizes $|\mathbf{k}|$ and $|\mathbf{s}_A|$, respectively,

$$\mathtt{G}(\mathbf{h}) = (\mathbf{k}, \mathbf{s}_A)\,, \tag{3.16}$$

where $\mathbf{k} \in \mathcal{K}$ denotes the key obtained from $\mathbf{h}$ after privacy amplification and $\mathbf{s}_A$ is Alice's syndrome.

Symmetric encryption algorithm (as described in Section 3.2.1, *e.g.*, AES GCM, denoted by $\mathtt{Es} : \mathcal{K} \times \mathcal{M} \rightarrow \mathcal{C}_{\mathcal{T}}$ where $\mathcal{C}_{\mathcal{T}}$ denotes the ciphertext space with corresponding decryption $\mathtt{Ds} : \mathcal{K} \times \mathcal{C}_{\mathcal{T}} \rightarrow \mathcal{M}$, such that

$$\mathtt{Es}(\mathbf{k}, \mathbf{m}) = \mathbf{c}, \tag{3.17}$$

$$\mathtt{Ds}(\mathbf{k}, \mathbf{c}) = \mathbf{m}, \tag{3.18}$$

for $\mathbf{m} \in \mathcal{M}$, $\mathbf{c} \in \mathcal{C}_{\mathcal{T}}$.

Pair of message authentication code MAC algorithms (as described in 3.2.1), e.g., in HMAC mode, denoted by $\mathtt{Sign} : \mathcal{K}_{\mathtt{MAC}} \times \mathcal{M} \rightarrow \mathcal{T}$, with a corresponding verification algorithm $\mathtt{Ver} : \mathcal{K}_{\mathtt{MAC}} \times \mathcal{M} \times \mathcal{T} \rightarrow (yes, no)$, such that

$$\mathtt{Sign}(\mathbf{k}_{\mathtt{MAC}}, \mathbf{m}) = \mathbf{t}, \tag{3.19}$$

$$\mathtt{Ver}(\mathbf{k}_{\mathtt{MAC}}, \mathbf{m}, \mathbf{t}) = \begin{cases} yes, & \text{if integrity verified} \\ no, & \text{if integrity not verified} \end{cases} \tag{3.20}$$

Security analysis

Finally, to verify the security properties of the proposed protocol, this work concludes with a security analysis using BAN logic and Tamarin prover.

Following from the above the system model for this study can be seen on Fig. 3.8. The model provides a PHY-based multi-factor mutual authentication mechanism between Alice (IoT device) and Bob (resourceful device). As discussed, combining SKG in conjunction with PUFs is a novel approach that is used here to build a secure and lightweight

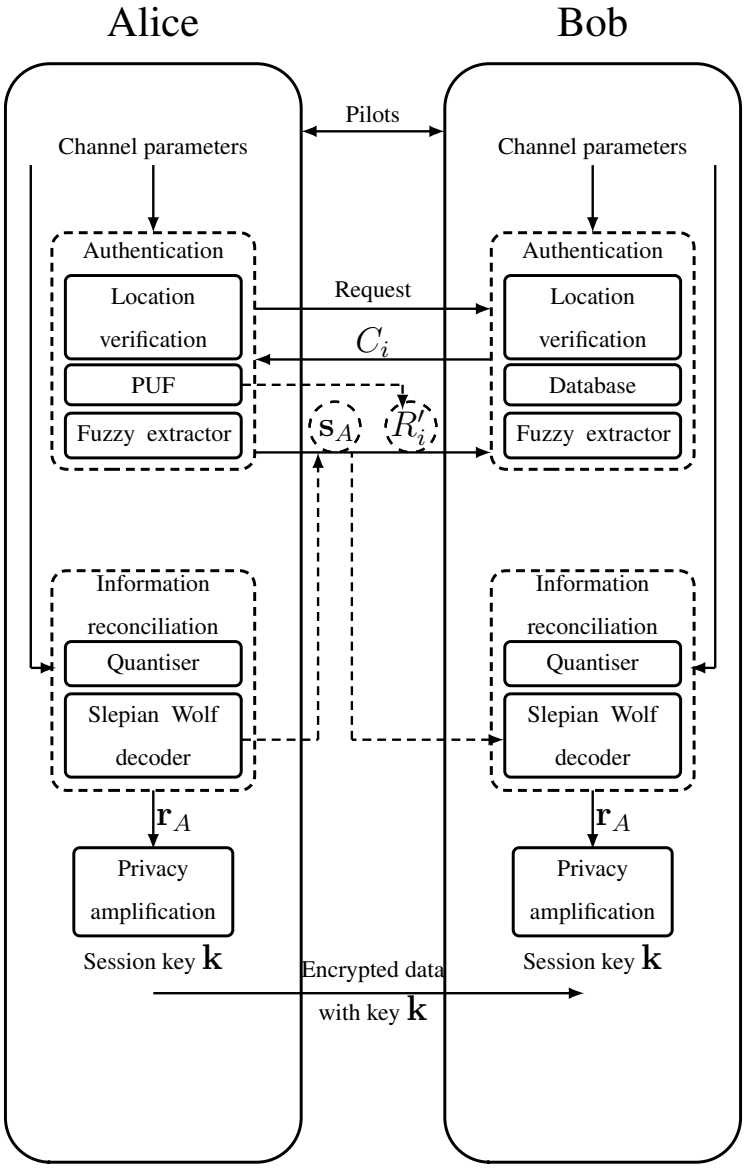

Figure 3.8: System model of the proposed multi-factor authentication scheme, where Alice is a resource constraint IoT device and Bob is a resourceful server.

solution. The figure illustrates that after the link establishment (*i.e.*, pilot exchange) both parties can use the channel parameters as a wireless fingerprint, which allows for proximity detection and SKG. Furthermore, Alice is equipped with a PUF that identifies her hardware fingerprint. The combination of wireless and hardware fingerprints gives the full authentication mechanism.

3.4 Authentication protocol

This section presents a lightweight authentication scheme, consisting of two phases: enrollment phase and authentication phase. First the enrolment, authentication / SKG and resumption schemes will be presented with little justification, this is followed by both an informal and formal security analysis.

3.4.1 Enrollment phase

The enrollment is one-time operation performed between Alice (denoted here as A) and Bob (denoted here as B). This phase is carried on a secure channel (*i.e.*, by using a time-based one time password algorithm [195]). The steps taken during enrollment are summarised in Fig. 3.9 and are performed as follows:

1. In order to establish the link between them, both devices need to exchange pilot signals. During this exchange A and B measure the RSS. Furthermore A downloads (or creates) a map of the premises which contains the location of B.

2. After establishing the connection, Alice sends her ID A with a request for registration `Request`.

3. Upon receiving the request, B first checks if the received ID has already been registered. If B finds the ID within his database the request is rejected. If A has not been registered Bob links the ID A with the computed proximity threshold ζ. Next, B generates two initial PUF challenges C_1, C_2 and an initial one-time alias ID $A_{\mathtt{ID},1}$. These challenges will be used during subsequent authentication and will be updated with each run of the protocol. Next, B generates sets of emergency challenges and one-time alias IDs $\mathcal{C}_{emerg}$ and $\mathcal{A}_{\mathtt{ID},emerg}$, respectively, such that $|\mathcal{C}_{emerg}| = |\mathcal{A}_{\mathtt{ID},emerg}|$. The emergency sets are used only in a case of desynchronisation between the devices and have multiple entries to allow for multiple recoveries. Finally, Bob sends the message $\{C_1, C_2, A_{\mathtt{ID},1}, (\mathcal{C}_{emerg}, \mathcal{A}_{\mathtt{ID},emerg})\}$

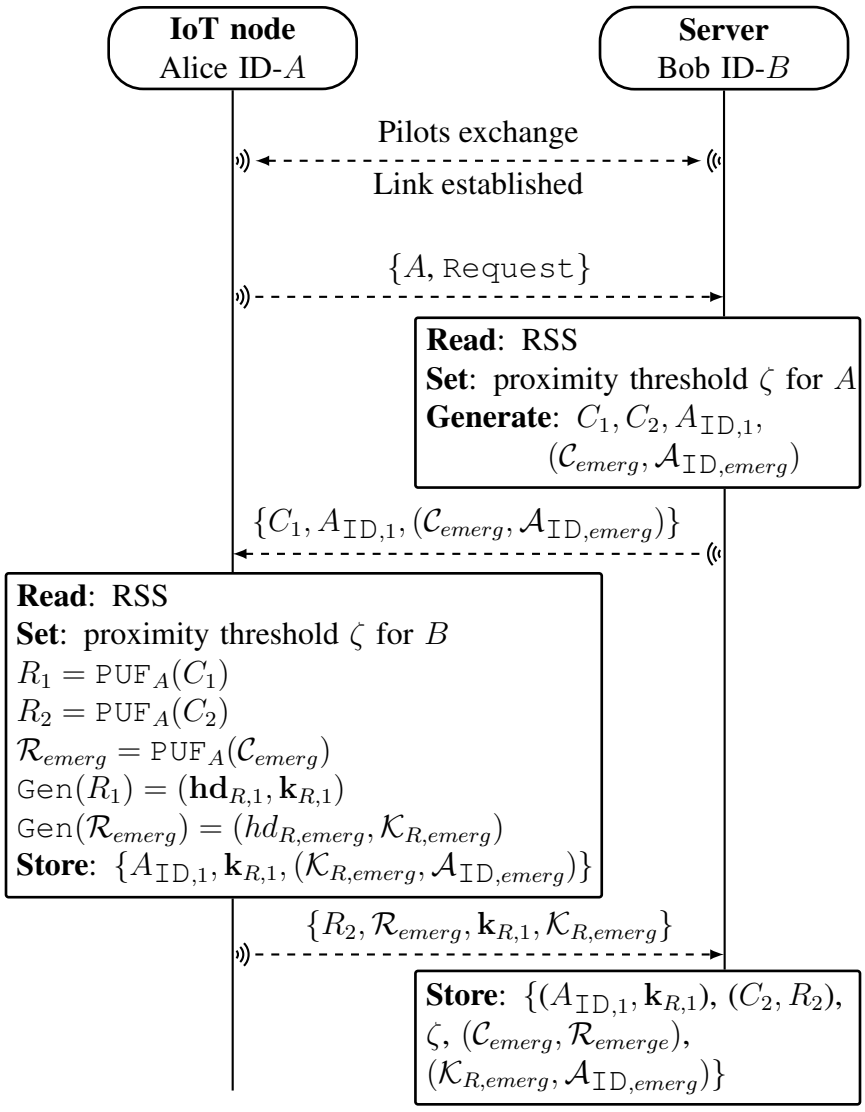

Figure 3.9: Enrollment phase (note using a secure channel)

to Alice. Note that the two emergency sets are linked such that each element has a corresponding one in the other set.

4. After receiving the message, Alice excites her $\texttt{PUF}_A$ with C_1, C_2 and all challenges from the set $\mathcal{C}_{emerg}$, producing responses R_1, R_2 and $\mathcal{R}_{emerg}$, respectively. Next, she uses R_1 and $\mathcal{R}_{emerg}$ as inputs to her fuzzy extractor to generate the pairs $(\mathbf{hd}_{R,1}, \mathbf{k}_{R,1})$ and $(hd_{R,emerg}, \mathcal{K}_{R,emerg})$. Afterwards, Alice stores the parameters: $\{A_{\texttt{ID},1}, \mathbf{k}_{R,1}, (\mathcal{K}_{R,emerg}, \mathcal{A}_{\texttt{ID},emerg})\}$ and sends the following message to Bob

$\{R_2, \mathcal{R}_{emerg}, \mathbf{k}_{R,1}, \mathcal{K}_{R,emerg}\}$.

5. To finalise the registration process B stores the following elements that correspond to the ID A in his database: proximity threshold ζ, initial authentication parameters $(A_{\mathtt{ID},1}, \mathbf{k}_{R,1}), (C_2, R_2)$ and emergency authentication parameters in case of de-synchronisation $(\mathcal{C}_{emerg}, \mathcal{R}_{emerge}), (\mathcal{K}_{R,emerg}, \mathcal{A}_{\mathtt{ID},emerg})$.

3.4.2 Authentication phase

Once the enrollment is finished both devices can use the established parameters for later authentication. The steps taken during authentication are summarised in Fig. 3.10 and are performed as follows:

1. First, in order to estimate the channel and establish connection both devices exchange pilot signals. The following parameters are measured during this phase: i) A and B estimate the CSI. Both parties quantise their CSI measurements into bit string $\mathbf{r}_A$ and $\mathbf{r}_B$, respectively. ii) Alice and Bob measure the RSS of the received signals.

2. Next, Alice performs the enhanced proximity detection mechanism to confirm the location of Bob. If the detection does not succeed she stops the authentication process. If it succeeds, she performs an SKG process obtaining syndrome $\mathbf{s}_A$ and key $\mathbf{k}$. The key will be used later as a session key if the authentication is successful. Then, A sends her request for authentication which contains a one-time alias ID $A_{\mathtt{ID},i}$ and a fresh random nonce N_1.

3. Upon receiving above, Bob first confirms whether Alice is at the expected distance by using the measured RSS and compares against that saved in his database proximity threshold ζ. If the check fails, he rejects the authentication request. If it succeeds he accesses the database and loads the parameters that corresponds to the ID, *i.e.*, CRP (C_2, R_2) and key $\mathbf{k}_{R,1}$. Then he generates a fresh random nonce N_B and breaks $\mathbf{k}_{R,1}$ into two parts as follows: $\mathbf{k}_{R,1} = (\mathbf{k}_{R1,1}, \mathbf{k}_{R1,2})$. Then uses the first

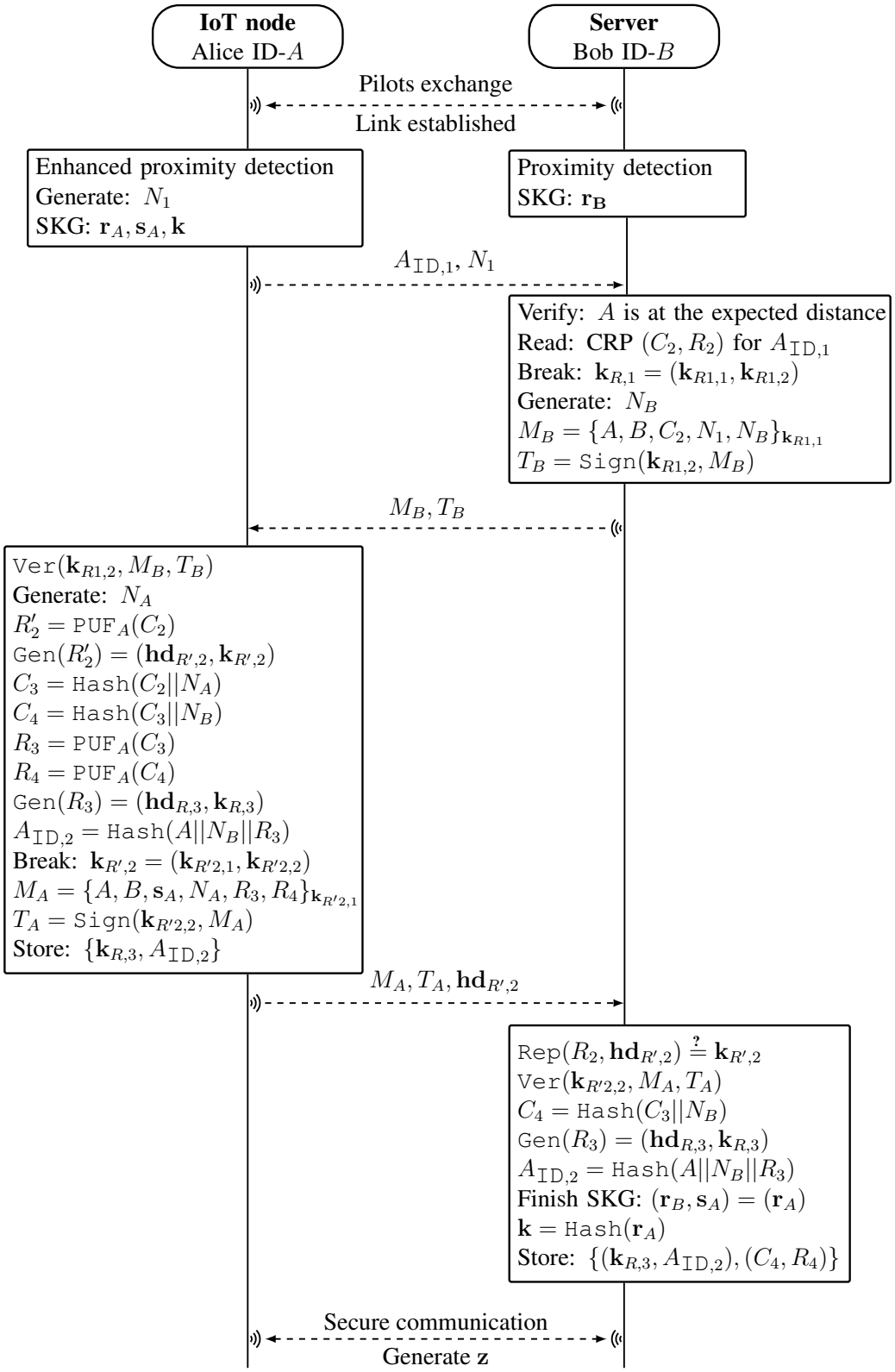

Figure 3.10: Multi-factor authentication protocol

part to encrypt $\{A, B, C_2, N_1, N_B\}$, resulting in cipher text M_B and uses the second part to sign M_B as: $T_B = \texttt{Sign}(\mathbf{k}_{R1,2}, M_B)$. Finally, he sends the cipher text M_B and the signed message T_B to Alice.

4. By using her stored key $\mathbf{k}_{R,1}$, A verifies the authenticity of Bob and the integrity of the message M_B. If one of the verification checks fail Alice rejects the message's claim to authenticity. If the verification succeeds she accepts and excites her PUF with the received challenge: $R_2' = \texttt{PUF}_A(C_2)$. Next, by using the principles of the fuzzy extractor, A uses her noisy response R_2' and produces two bit strings, (one in the helper data space and one the key space) as follows: $\texttt{Gen}(R_2') = (\mathbf{hd}_{R',2}, \mathbf{k}_{R',2})$. Afterwards, she generates a new fresh random nonce N_A and calculates the next two challenges as follows: $C_3 = \texttt{Hash}(C_2||N_A)$ and $C_4 = \texttt{Hash}(C_3||N_B)$. Next, she excites her PUF to produce R_3 and R_4. In order to generate the key that will be used in subsequent running of the authentication protocol A performs the function $\texttt{Gen}(R_3) = (\mathbf{hd}_{R,3}, \mathbf{k}_{R,3})$. Next, she calculates the one-time alias ID for upcoming run of the protocol as: $A_{\texttt{ID},2} = \texttt{Hash}(A||N_B||R_3)$ which due to the randomness of N_B and R_3, cannot be linked to $A_{\texttt{ID},1}$. The pairs (C_4, R_4) and $(\mathbf{k}_{R,3}, A_{\texttt{ID},2})$ will be used in subsequent connection with Bob. Next, Alice breaks her key $\mathbf{k}_{R',2}$ into two parts $\mathbf{k}_{R',2} = (\mathbf{k}_{R'2,1}, \mathbf{k}_{R'2,2})$. Similarly, to the previous step she uses half of the key to encrypt the message $M_A = \{A, B, \mathbf{s}_A, N_A, R_3, R_4\}$. Then, A uses the second half of the key to sign the cipher text as follows: $T_A = \texttt{Sign}(\mathbf{k}_{R'2,2}, M_A)$. Finally, Alice sends M_A, T_A and $\mathbf{hd}_{R',2}$ to Bob and stores the pair $(\mathbf{k}_{R,3}, A_{\texttt{ID},2})$.

5. Upon receiving the preceding message, Bob uses the stored R_2 (from the enrollment phase) and the received helper data $\mathbf{hd}_{R',2}$ in order to verify the condition $\texttt{Rep}(R_2, \mathbf{hd}_{R',2}) \stackrel{?}{=} \mathbf{k}_{R',2}$ by . If the verification fails, B rejects the claim to authenticity. If the claim is accepted, he verifies the integrity of M_A using the signed cipher text T_A. Next, using R_3 and the principles of the fuzzy extractor Bob performs $\texttt{Gen}(R_3) = (\mathbf{hd}_{R,3}, \mathbf{k}_{R,3})$. He calculates $A_{\texttt{ID},2} = \texttt{Hash}(A||N_B||R_3)$. Following that, he stores the pairs $(\mathbf{k}_{R,3}, A_{\texttt{ID},2}), (C_4, R_4)$ which will be used during

the next round of the protocol. Finally, using the received syndrome $\mathbf{s}_A$, B corrects the discrepancies in his observation $\mathbf{r}_B$ to obtain $\mathbf{r}_A$ and calculates the session key using the privacy amplification method described in Section 2.2.1 as follows: $\mathbf{k} = \texttt{Hash}(\mathbf{r}_A)$.

6. After the authentication process finishes A and B enter the secure communication stage with session key $\mathbf{k}$. During this stage, both use the channel properties to generate a shared resumption secret $\mathbf{z}$. Instead of performing full authentication in subsequent sessions, the secret can be used as a parameter to quickly "resume" sessions within 0-RTT.

3.4.3 Resumption protocol

This section presents a novel physical layer resumption protocol that allows Alice to send encrypted data within a 0-RTT. During the secure communication stage of the authentication protocol in Fig. 3.10, B sends to A a look-up identifier $\mathbf{k}_l$. Then, both derive a resumption secret $\mathbf{z}$ that is identified by $\mathbf{k}_l$ (as discussed in Section 3.2.3). Note, $\mathbf{z}$ and session keys have the same length $|\mathbf{k}|$. The usage of a resumption secret for authentication helps avoiding man-in-the-middle attacks in the scenario assumed here. Given the above, the resumption protocol follows the steps:

1. As before, in order to establish the link both devices perform pilot exchange procedure. Alice and Bob obtain channel observations and generate the sequences $\mathbf{r}_A$ and $\mathbf{r}_B$, respectively. Furthermore, both parties measures the RSS.

2. Following the above, Alice performs the enhanced proximity detection mechanism to verify whether Bob is at the expected location. If the verification fails, she aborts the connection. If the verification succeeds she generates a fresh random nonce N_1 and reads the resumption secret $\mathbf{z}$. Next, she obtains $\mathbf{r}^* = \mathbf{z} \oplus \mathbf{r}_A$. Then, using her Slepian Wolf decoder she calculates the syndrome $\mathbf{s}^*$, that corresponds to $\mathbf{r}^*$, and generates the session key as $\mathbf{k}^* = \texttt{Hash}(\mathbf{r}^*)$. She also calculates the one-time alias

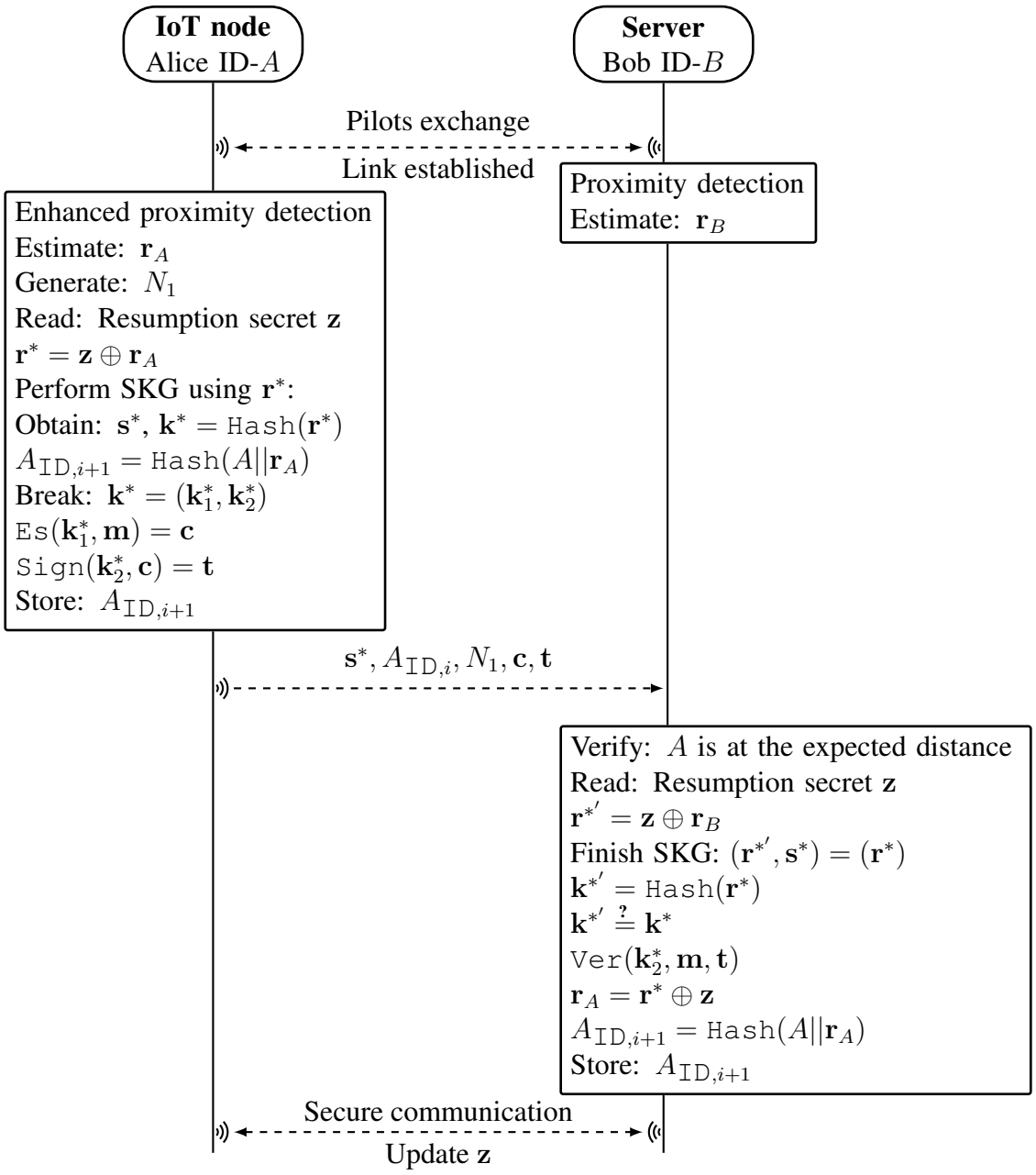

Figure 3.11: Resumption protocol

ID that will be used for subsequent session as: $A_{\mathtt{ID},i+1} = \mathtt{Hash}(A||\mathbf{r}_A)$. Alice breaks her key into two parts $\mathbf{k}^* = (\mathbf{k}_1^*, \mathbf{k}_2^*)$ and uses the first part to encrypt the early 0-RTT data $\mathbf{m}$ as $\mathtt{Es}(\mathbf{k}_1^*, \mathbf{m}) = \mathbf{c}$. The second part she uses to sign the cipher text $\mathtt{Sign}(\mathbf{k}_2^*, \mathbf{c}) = \mathbf{t}$. Finally, she sends $\mathbf{s}^*, A_{\mathtt{ID},i}, N_1, \mathbf{c}, \mathbf{t}$ and stores $A_{\mathtt{ID},i+1}$.

Note that the key $\mathbf{k}^*$ can only be obtained if both the physical layer generated key and the resumption key are valid and this method can be shown to be forward secure [127].

3. Upon receiving that, B first verifies if A is at the expected distance. Then, reads the resumption secret $\mathbf{z}$ and obtains $\mathbf{r}^{*'} = \mathbf{z} \oplus \mathbf{r}_B$. Using that and the received syndrome $\mathbf{s}^*$, Bob finishes the SKG process and obtains $\mathbf{k}^* = \texttt{Hash}(\mathbf{r}^*)$. He uses the condition $\mathbf{k}^{*'} \stackrel{?}{=} \mathbf{k}^*$ to verify the authenticity of Alice and the integrity of the message. If the above succeed he calculates $\mathbf{r}_A = \mathbf{r}^* \oplus \mathbf{z}$ and stores $A_{\texttt{ID},i+1} = \texttt{Hash}(A||\mathbf{r}_A)$.

4. After the resumption process finishes the two devices enter the secure communication stage using $\mathbf{k}^*$ as a session key. During the secure communication stage, they use the channel and session properties to generate a new shared resumption secret that could be used in subsequent resumptions.

3.5 Security analysis

This section provides security analysis of the lightweight authentication protocol illustrated in Fig. 3.10. The security proofs consider a Dolev-Yao [196] type of adversary, who has control over the wireless channel between A and B. Furthermore: 1) the adversary can send any type of messages and queries using its knowledge gained through observation; 2) all functions and operations performed by the legitimate users during the execution of the protocol are public except for $\texttt{PUF}_A(\cdot)$ and the entire enrollment phase; 3) the adversary can launch a DoS attack and block parts of the protocol in order to desynchronise the connection between A and B. For simplicity, the work assumes a rich Rayleigh multipath environment where the adversary is more than a few wavelengths away from each of the legitimate parties. This forms the basis of our hypothesis that the measurements of Alice and Bob are uncorrelated to the adversary's measurements. Finally, the security analysis build upon the fact that the enrollment phase is carried out on

a secure channel. This is a realistic assumption for numerous systems where an initial bootstrapping is required to set up a new device within an existing network [197, 198]. Following that, first an informal, descriptive, analysis will be presented, this will be followed by a formal analysis.

3.5.1 Informal security analysis

Mutual authentication: The proposed protocol uses a set of factors to achieve mutual authentication. It uses enhanced proximity detection as a first factor of authentication. This verifies whether both parties are at the expected position. Next, Alice authenticates Bob by verifying whether the correct key is used for creating M_B and T_B. On the other hand, Bob authenticates Alice by first confirming the validity of the received one-time alias ID $A_{\mathtt{ID},i}$ and second by verifying whether she produced a valid response to C_i. The second condition is confirmed only if Alice uses the correct key to generate the pair M_A, T_A.

Untraceability and Anonymity: During the execution of the authentication protocol, Alice must posses a valid one-time alias ID $A_{\mathtt{ID}}$ for each session. The one-time alias identity cannot be used twice and there is not direct relationship between consequent aliases. Thus, no one except Bob would know what is the origin of the message. Furthermore, in case of de-synchronisation the device can use the set of emergency IDs $\mathcal{A}_{\mathtt{ID},emerg}$. After using an emergency ID it has to be deleted from Alice's and Bob's memory. This approach provides privacy against eavesdroppers and ensures user's anonimity and identity untraceability properties.

Perfect forward secrecy: Assuming an attacker compromises A and obtains all stored secrets, *i.e.*, $(\mathbf{k}_R, A_{\mathtt{ID}})$, he cannot obtain previous keys or one-time alias IDs. First, each $\mathbf{k}_R$ is generated using a CRP and CRPs are randomly generated and independent. Hence, by obtaining $\mathbf{k}_{R,i}$ an adversary cannot learn $\mathbf{k}_{R,i-1}$. Next, one-time alias IDs are generated using a one-way hash function of unique parameters for each session; if an adversary obtains $A_{\mathtt{ID},i}$, he can not inverse the hash function. Furthermore, using the

randomness of the wireless channel ensures that session keys are unique and independent for each session. Therefore, the proposed authentication protocol ensures perfect forward secrecy property.

Protection against replay attack: If an adversary intercepts previous communication between A and B, he can replay the same messages and try to pass the authentication process. In the protocol presented on Fig. 3.10 none of the parameters in the initial request are allowed to be sent twice, hence, if an attacker resends the same message to B the attack will be detected and the request will be rejected. Next, if the adversary tries to re-send M_B to A, he will be detected, since the key used to encrypt M_B is changed with every session. Similarly, if the adversary tries to re-send M_A, he will be detected and the request will be rejected because the key used to encrypt M_A is changed every session. The above proves that the proposed protocol provides resistance against replay attacks.

Protection against impersonation attack: A successful impersonation attack will allow the adversary to be authenticated as a legitimate user. Following from above, an adversary cannot perform a replay attack, which limits his options to perform an impersonation attack. Following from that, in order to impersonate A he must generate 1) a valid one-time alias ID and 2) a valid message M_A 3) be at the expected distance from S. However, due to the unclonability properties of the PUF and the fact that the connection between a device and its PUF is secure, (*i.e.*, system on chip) the adversary cannot generate a valid message M_A, hence cannot impersonate Alice. Next, in order to impersonate Bob the adversary must posses a valid key $\mathbf{k}_{R,1}$ and generate a valid message M_B. To obtain the key an adversary must compromise A (an example of such a scheme vulnerable to this attack can be found in [199]). However, even if A is compromised by performing the enhanced proximity detection Alice will detect the attack. The above proves that our multi-factor authentication protocol provides resistance against impersonation attacks.

Resistance to DoS attack: To ensure security against DoS and de-syncronisation attacks, the authentication protocol uses unlinkable one-time alias IDs and pairs of sets with emergency parameters $(\mathcal{C}_{emerg}, \mathcal{R}_{emerge})$ and $(\mathcal{K}_{R,emerg}, \mathcal{A}_{\mathtt{ID},emerg})$. If an adversary manages to block a message from a legitimate party, such that it does not reach its intended

receiver, the authentication process will stop and the used $A_{\mathtt{ID},i}$ will not be updated. To overcome that A can use one of her emergency IDs from the set $\mathcal{A}_{\mathtt{ID},emerg}$. Bob will then read the corresponding $K_{R,emerg}$ from the set $\mathcal{K}_{R,emerg}$ and use it to encrypt a message containing an emergency challenge C_{emerg} from the set $\mathcal{C}_{emerg}$. Next, both parties can continue the authentication process as usual and setup a new one-time alias ID. In order to prevent replay attacks all used emergency parameters must be deleted from the corresponding set. This approach provides resiliency against DoS to de-synchronisation attacks.

Protection against cloning attacks: A successful cloning attack allows the adversary to use a captured device in order to obtain secrets stored on an another device. In the proposed protocol each device posses a unique pair $(\mathbf{k}_R, A_{\mathtt{ID}})$. Furthermore, all devices have unique PUFs and will produce a unique response to a challenge. Hence, the adversary cannot use secrets derived from one device in order to clone another.

Protection against physical attacks: Successful physical attack could be performed by physical tampering on the IoT device in order to change its behavior. However, by changing its behavior, the PUF will not produce the desired response and therefore B will detect the attack. Therefore, the proposed protocol is resistant against physical attacks.

Session key agreement: It is a common practice in literature to use nonces as part of the session key generation process [153, 199, 200]. However, note that even if N_A and N_B are good shared secrets (the next section gives a formal proof for the secrecy of N_A and N_B) between A and B the low entropy of pseudo-random number generator (PRNG) modules may provoke set of attacks, such as side-channel and prediction attacks [201, 202], and lead to information leakage. Furthermore, it has been shown that true-random number generators (TRNGs) can greatly increase the time complexity in a resource limited systems making the generation time infeasible [203]. Therefore, the role of the nonces in the proposed scheme is limited to only a source of freshness. On the other hand, the randomness already present in the wireless channel allows for secure and lightweight key generation process. The SKG procedure presented in section 2.2.1 can

produce a uniform random key with size as much as [12]:

$$|\mathbf{k}| \leq \mathsf{H}(\mathbf{x}_A) - \mathsf{I}(\mathbf{x}_A; \mathbf{x}_E) - \mathsf{H}(\mathbf{x}_A|\mathbf{x}_B) - r_0, \tag{3.21}$$

where $\mathsf{H}(\mathbf{x}_A)$ represents the entropy of the measurement, $\mathsf{I}(\mathbf{x}_A; \mathbf{x}_E)$ represents the mutual information between Alice's and an eavesdropper's observations, the entropy revealed during information reconciliation is $\mathsf{H}(\mathbf{x}_A|\mathbf{x}_B)$ – which in our scheme is zero due to the encryption of the syndrome $\mathbf{s}_A$, and $r_0 > 0$ is an extra security parameter that ensures uncertainty on the key at an eavesdropper's side. For details and estimation of these parameters in a practical scenario please see [114]. Finally, we note that if the session key somehow gets compromised, the authentication process remains secure as the adversary cannot obtain the PUF response using the session key.

3.5.2 Formal security analysis using BAN logic and Tamarin prover

Secrecy proofs Using BAN Logic

As discussed in Section 3.2.5 secrecy evaluation is a vital step to ensure security of protocols. In this sense, the BAN logic [182] is widely used verification tool. However, several weaknesses of the logic were identified throughout the years. Therefore, for the verification purposes the authentication protocol in Fig. 3.10 an improved and and more reliable version of BAN logic, *i.e.*, MB logic [186] is used in this section. Further details about MB logic are given in Section 3.2.5. The first step of MB logic is to obtain the idealised version of the protocol and to define the initial beliefs. Based on the set of rules and definitions defined in Section 3.2.5, the protocol in Fig. 3.10 is idealised as follows:

1. $A \rightarrow B : A, N_1$
2. $B \rightarrow A : \{N_B \mathfrak{R} N_1\}_{\mathbf{k}_{R,1}}$
3. $A \rightarrow B : \{R_3 | R_4 \mathfrak{R} N_A \mathfrak{R} N_B\}_{\mathbf{k}_{R',2}}$

Table 3.1: Inference rules adopted from Mao and Boyd logic

Notation	**Description**
$\frac{A \models A \overset{\mathbf{k}}{\leftrightarrow} B \wedge A \overset{\mathbf{k}}{\triangleleft} \mathbf{m}}{A \models B \overset{\mathbf{k}}{\mid\sim} \mathbf{m}}$	Authentication rule
$\frac{A \models A \overset{\mathbf{k}}{\leftrightarrow} B \wedge B^{C} \triangleleft \mid\mid \mathbf{m} \wedge A \overset{\mathbf{k}}{\mid\sim} \mathbf{m}}{A \models (A \cup B)^{C} \triangleleft \mid\mid \mathbf{m}}$	Confidentiality rule
$\frac{A \models \#(\mathbf{m}) \wedge A \triangleleft \mathbf{n} \mathcal{R} \mathbf{m}}{A \models \#(\mathbf{n})}$	Fresh rule
$\frac{A \models \{A,B\}^{C} \triangleleft \mid\mid \mathbf{k} \wedge A \models \#(\mathbf{k})}{A \models A \overset{\mathbf{k}}{\leftrightarrow} B}$	Good-key rule
$\frac{A \models \#(\mathbf{n}) \wedge A \models B \overset{\mathbf{k}}{\mid\sim} \mathbf{n}}{A \models B \models A \overset{\mathbf{k}}{\leftrightarrow} B}$	Nonce verification rule
$\frac{A \models B \models X \wedge A \models \text{sup}(B)}{A \models X}$	Super-principal rule
$\frac{A \models X \wedge A \models Y}{A \models (X \wedge Y)}$	Belief axiom 1
$\frac{A \models X \wedge A \models X/Y}{A \models Y}$	Belief axiom 2

where $\mathfrak{R}$ denotes the relation between parameters, as defined in Section 3.2.5. Next, denoting principal as A, B, messages as $\mathbf{m}, \mathbf{k}$ and formulas as X the main properties of MB logic are recalled here for clarity: $A \models X$ denotes A believes X is true; $A \overset{\mathbf{k}}{\triangleleft} \mathbf{m}$ denotes A sees $\mathbf{m}$ using key $\mathbf{k}$, if $\mathbf{m}$ is not encrypted this simplifies to $A \triangleleft \mathbf{m}$; $A \overset{\mathbf{k}}{\mid\sim} \mathbf{m}$ denotes A encrypts $\mathbf{m}$ using key $\mathbf{k}$; $\#(\mathbf{m})$ denotes $\mathbf{m}$ is of type fresh; $A \overset{\mathbf{k}}{\leftrightarrow} B$ denotes $\mathbf{k}$ is a good shared key between A and B; $A \triangleleft \mid\mid \mathbf{m}$ denotes $\mathbf{m}$ is not available to A; sup(B) denotes B is a super-principal, *i.e.*, B is the legitimate source of a specific message. Following that, the inference rules, defined in [186], and used in this section are given in Table 3.1.

Given the above, the initial beliefs are denoted as follows:

A1 $A \models A \overset{\mathbf{k}_{R,1}}{\leftrightarrow} B$ and $B \models A \overset{\mathbf{k}_{R,1}}{\leftrightarrow} B$ – Bob stores CRP (C_1, R_1) with the corresponding ID A in its memory *i.e.* following a correct enrolment phase. The key $\mathbf{k}_{R,1}$ can be generated using a fuzzy extractor only from the respective R_1 and only A can

generate R_1' such that $D(R_1, R_1') \leq t$.

A2 $A \models A \overset{\mathbf{k}_{R',2}}{\leftrightarrow} B$ and $B \models A \overset{\mathbf{k}_{R',2}}{\leftrightarrow} B$ – Bob stores CRP (C_2, R_2) with the corresponding ID A in its memory. The key $\mathbf{k}_{R',2}$ can be generated using a fuzzy extractor only from the respective R_2' such that $D(R_2, R_2') \leq t$.

A3 $B \models A \models A^C \triangleleft ||R_3|R_4\mathfrak{R}N_A\mathfrak{R}N_B$ – the information within message 3 can be accessed only if $\mathbf{k}_{R',2}$ is used as the decryption key.

A4 $B \models \sup(A)$ – A is super-principal w.r.t. R_3, R_4 and N_A.

A5 $B \overset{\mathbf{k}_{R',2}}{\triangleleft} R_3|R_4\mathfrak{R}N_A\mathfrak{R}N_B$ – B can access the information within message 3 of the idealised protocol only if using $\mathbf{k}_{R',2}$ as decryption key.

A6 $A \overset{\mathbf{k}_{R',2}}{\mid\sim} R_3|R_4\mathfrak{R}N_A\mathfrak{R}N_B$ – A uses $\mathbf{k}_{R',2}$ to encrypt message 3 of the idealised protocol.

A7 $A \models B^C \triangleleft ||R_3|R_4\mathfrak{R}N_A\mathfrak{R}N_B$ – A believes that no one can see the information within message 3 unless they posses $\mathbf{k}_{R',2}$.

A8 $A \models \#(N_1)$, $A \models \#(N_A)$, $A \models \#(R_3)$, $A \models \#(R_4)$ – A generates fresh random N_1, N_A, R_3, R_4 with each run of the protocol.

A9 $A \overset{\mathbf{k}_{R,1}}{\triangleleft} N_B\mathfrak{R}N_1$ – A can see the information within message 2 only if using $\mathbf{k}_{R,1}$.

A10 $A \models B \models B^C \triangleleft ||N_B$ and $B \models \#(N_B)$ – N_B is generated by B and is a fresh parameter each time when the protocol is executed.

A11 $A \models \sup(B)$ – B is super-principal w.r.t. N_B.

A12 $B \models A^C \triangleleft ||N_B\mathfrak{R}N_1$ – B believes that no one can see the information within message 2 unless they posses $\mathbf{k}_{R,1}$.

A13 $B \overset{\mathbf{k}_{R,1}}{\mid\sim} N_B\mathfrak{R}N_1$ – B uses $\mathbf{k}_{R,1}$ to encrypt message 2 of the idealised protocol.

$$\frac{A \models A \overset{\mathbf{k}_{R,1}}{\leftrightarrow} B \wedge A \overset{\mathbf{k}_{R,1}}{\triangleleft} N_B}{A \models B \overset{\mathbf{k}_{R,1}}{\mid\sim} N_B}$$

(a)

$$\frac{A \models A \overset{\mathbf{k}_{R',2}}{\leftrightarrow} B \wedge B \overset{\mathbf{k}_{R',2}}{\triangleleft} N_A}{B \models A \overset{\mathbf{k}_{R',2}}{\mid\sim} N_A}$$

(b)

Figure 3.12: Proof of authentication (a) B to A (b) A to B.

The authentication property of the initial run of the protocol can be verified using the authentication rule as shown on Fig. 3.12. The authentication of B to A can be achieved by the fact $A \models B \overset{\mathbf{k}_{R,1}}{\mid\sim} N_B$, *i.e.*, A believes that B sent N_B using $\mathbf{k}_{R,1}$ to encrypt the message. As seen in Fig. 3.12, two facts imply authentication 1) $A \models A \overset{\mathbf{k}_{R,1}}{\leftrightarrow} B$ meaning A believes that $\mathbf{k}_{R,1}$ is a good shared secret between A and B; 2) $A \overset{\mathbf{k}_{R,1}}{\triangleleft} N_B$,*i.e.*, A used $\mathbf{k}_{R,1}$ as a decryption key to see N_B. Following from the fact that the enrollment stage is performed on a secure channel, both of these statements are part of the initial beliefs of the protocol, hence, the authentication of B to A is directly established as shown on Fig. 3.12 a). The authentication of A to B is identical following from Fig. 3.12 b).

Next, follows the proof of secrecy for the parameter R_3 (the proofs for secrecy of N_A and R_4 are identical) which could be used as initial belief for the next run of the protocol. First, by combining the confidentiality rule and two axioms from Table 3.1 one can derive the following rule:

$$\frac{A \models B \overset{\mathbf{k}}{\mid\sim} \mathbf{m} \wedge A \models B \models A \overset{\mathbf{k}}{\leftrightarrow} B \wedge A \models B \models B^C \triangleleft ||\mathbf{m}}{A \models B \models \{A \cup B\}^C \triangleleft ||\mathbf{m}}. \tag{3.22}$$

Given that, the security proofs on Fig. 3.13 (a) and (b) show that both parties A and B agree that R_3 is a good shared secret (identically one can show the property holds for R_4). Given that and using the fuzzy extractor properties [157, 163] it can be concluded

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{B\models\#(N_B)\wedge B\overset{\mathbf{k}_{R',2}}{\triangleleft}R_3\mathfrak{R}N_B}{B\models\#(R_3)}\wedge\dfrac{B\models A\overset{\mathbf{k}_{R',2}}{\leftrightarrow}B\wedge B\overset{\mathbf{k}_{R',2}}{\triangleleft}R_3}{B\models A\overset{\mathbf{k}_{R',2}}{\mid\sim}R_3}}{B\models A\models A\overset{\mathbf{k}_{R',2}}{\leftrightarrow}B}\wedge B\models A\models A^c\triangleleft||R_3\wedge B\models A\overset{\mathbf{k}_{R',2}}{\mid\sim}R_3}{B\models A\models\{B\cup A\}^c\triangleleft||R_3}\wedge B\models\text{sup}(A)}{B\models\{B\cup A\}^c\triangleleft||R_3}\wedge B\models\#(R_3)}{B\models A\overset{R_3}{\leftrightarrow}B}$$

(a)

$$\dfrac{\dfrac{A\models A\overset{\mathbf{k}_{R',2}}{\leftrightarrow}B\wedge A\models B^c\triangleleft||R_3\wedge A\overset{\mathbf{k}_{R',2}}{\mid\sim}R_3}{A\models\{A\cup B\}^c\triangleleft||R_3}\wedge A\models\#(R_3)}{A\models A\overset{R_3}{\leftrightarrow}B}$$

(b)

Figure 3.13: Secrecy proofs: (a) B believes R_3 is a good shared secret between A and B (b) A believes R_3 is a good shared secret between A and B

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A\models\#(N_1)\wedge A\overset{\mathbf{k}_{R,1}}{\triangleleft}N_B\mathfrak{R}N_1}{A\models\#(N_B)}\wedge\dfrac{A\models A\overset{\mathbf{k}_{R,1}}{\leftrightarrow}B\wedge A\overset{\mathbf{k}_{R,1}}{\triangleleft}N_B}{A\models B\overset{\mathbf{k}_{R,1}}{\mid\sim}N_B}}{A\models B\models A\overset{\mathbf{k}_{R,1}}{\leftrightarrow}B}\wedge A\models B\models B^c\triangleleft||N_B\wedge A\models B\overset{\mathbf{k}_{R,1}}{\mid\sim}N_B}{A\models B\models\{A\cup B\}^c\triangleleft||N_B}\wedge A\models\text{sup}(B)}{A\models\{A\cup B\}^c\triangleleft||N_B}\wedge A\models\#(N_B)}{A\models A\overset{N_B}{\leftrightarrow}B}$$

(a)

$$\dfrac{\dfrac{B\models A\overset{\mathbf{k}_{R,1}}{\leftrightarrow}B\wedge B\models A^c\triangleleft||N_B\wedge B\overset{\mathbf{k}_{R,1}}{\mid\sim}N_B}{B\models\{B\cup A\}^c\triangleleft||N_B}\wedge B\models\#(N_B)}{B\models A\overset{N_B}{\leftrightarrow}B}$$

(b)

Figure 3.14: Secrecy proofs: (a) A believes N_B is a good shared secret between A and B (b) B believes N_B is a good shared secret between A and B

```
theory PUF_protocol
begin
builtins: symmetric-encryption, hashing
functions: break1/1, break2/1, mac/2, vermac/3, true/0, fuzzyk/1,
           fuzzyhd/1, rep/2, SKG/2, puf/1 [private]
equations: vermac(x.1, x.2, mac(x.1, x.2)) = true,
           rep(puf(x.1), fuzzyhd(puf(x.2))) = fuzzyk(puf(x.2))
```

Figure 3.15: Example of protocol definition in Tamarin

that $\mathbf{k}_{R,3}$ and $\mathbf{k}_{R,4}$ are good shared keys between A and B. Next, Fig. 3.14 (a) and (b) illustrates that A and B agree that N_B is a good shared secret. As a consequence of the above and by using the properties of universal hash functions [204] we can conclude that $A_{\mathtt{ID},2}$ is also a good shared secret, as it is derived from a hash of $(A||N_B||R_3)$.

Security Verification Using Tamarin-prover

The security properties of the authentication protocol given in Fig. 3.10 were verified using the formal verification tool, Tamarin-prover. Tamarin was used to prove: secrecy of parameters, aliveness, weak agreement, non-injective agreement, injective agreement, untraceablity, and, anonymity. This section gives a model of the authentication protocol using in Tamarin syntax. Next, it provides formal proofs for the security properties of the protocol.

Protocol definition: As discussed in Appendix A a protocol definition within Tamarin begins with calling built-in functions, defining new functions and equations. These are illustrated in Fig. 3.15. The built-in functions used to define the protocol are `symmetric-encryption` and `hashing`. Note, the notation `f/2`, defines function with arity `2`. By simply calling `symmetric-encryption` Tamarin defines the functions `senc/2` and `sdec/2` which are related through the equation `sdec(senc(m,k),k)=m`; calling `hashing` defines the one-way hash function `h/1`. The manually defined functions are as follows: `break1/1`, `break2/1` used by the two parties to break their keys into two parts as discussed in Section 3.4; `mac`, `vermac` and `true` are used to define the MAC mechanism used throughout the protocol, these parameters are related through the

```
1 rule Enrollment:
2 let
3 R_2=puf(~C_2old)
4 in
5 [ Fr(~K_R1), Fr(~A_ID), Fr(~C_2), Fr(~C_2old)]
6 --[Once(),Secure(~C_2),Secure(R_2),Secure(~K_R1),Secure(~A_ID)]->
7 [A0($A, $B, <~K_R1, ~A_ID>), B0($B, $A, <~K_R1, ~A_ID, ~C_2, R_2>)]
```

Figure 3.16: Enrollment phase modelled in Tamarin

equation on line 6 in Fig. 3.15, which can be interpreted as:

$$\texttt{vermac}(\mathbf{k}_{\texttt{MAC}}, \mathbf{m}, \texttt{mac}(\mathbf{k}_{\texttt{MAC}}, \mathbf{m})) = \texttt{true}. \tag{3.23}$$

Next, `fuzzyk/1` and `fuzzyhd/1` define the fuzzy extractor `Gen` function, which for the purpose of the definition here is divided into two functions, one for key generation and one for helper data generation, respectively. The `Rep` FE function is defined as `rep/2`. The SKG procedure is defined as `skg/2`. Next, the PUF at Alice is defined as `puf/1 [private]`, where the addition `[private]`, defines that this function cannot be executed by an adversary. Finally, the equation defined on line 7 can be interpreted as follows:

$$\texttt{Rep}(R'_1, \mathbf{hd}_{R1}) = \mathbf{k}_{R1} \tag{3.24}$$

Following from the above, the first rule that models the enrollment phase is given in Fig. 3.16; the premise (line 5) is used to define fresh random variables that correspond to $\mathbf{k}_{R1}$, $A_{\texttt{ID}}$ and C_2. As the Tamarin functions produce identical outputs if using the same input multiple times two challenges are defined within the premise. This is done to model the measurement noise within the PUF, *i.e.*, `C_2old` is used to produce `R_2` which is stored in the server database and `C_2` will be used later to produce a noisy version of `R_2`. Action facts `Once` and `Secure` will be used later to define that the enrollment phase is performed only once and it is performed on a secure channel. Finally the conclusion (line 7) defines the states of Alice `A0` and Bob `B0` after the enrollment. The states are used to define the knowledge of each party at the beginning and at the end of each rule.

```
rule Init_A:
  [ Fr(~id), Fr(~r_A), A0(A, B, <K_R1, A_ID>)]
--[ Initialise(A, ~id), Role('A')]->
  [ A1(A, B, <K_R1, A_ID, ~r_A>), !KeyA(A, K_R1)]

rule Init_B:
  [ Fr(~id), Fr(~r_B), B0(B, A, <K_R1, A_ID, C_2, R_2>)]
--[ Initialise(B, ~id), Role('B')]->
  [ B1(B, A, <K_R1, A_ID, C_2, R_2, ~r_B>), !KeyB(B, <K_R1, R_2>)]
```

Figure 3.17: Link establishment between Alice and Bob in Tamarin

```
rule Compromise_Alice:
[!KeyA(A, K_R1)]
--[CompromiseA(A)]->
[Out(K_R1)]

rule Compromise_Bob:
[!KeyB(B, <K_R1, R_2>)]
--[CompromiseB(B)]->
[Out(<K_R1, R_2>)]
```

Figure 3.18: Defining compromising action in Tamarin

Next, the rules illustrated in Fig. 3.17 initialise the authentication protocol and define the establishment of the link between the two parties. The premises of the two rules define: 1) a session ID, *i.e.*, `id`; 2) using the properties of the wireless channel Alice and Bob obtain bit sequences $\mathbf{r}_A$ and $\mathbf{r}_B$, respectively; 3) the knowledge of states `A0` and `B0` are imported within the two premises. The action facts `Initialise` and `Role` are used to define authentication properties and to relate actions to their executant. Finally, the conclusion of the two rules define the knowledge of Alice and Bob at this stage through facts `A1` and `B1`, respectively. Furthermore, the two facts `!KeyA` and `!KeyB` will be used to define a compromised agent, *i.e.*, if Alice is compromised the attacker will obtain $\mathbf{k}_{R1}$; if Bob is compromised the attacker will obtain $\mathbf{k}_{R1}$ and R_2. The compromising action is modelled using the rules in Fig. 3.18. As illustrated in the figure, whenever action fact `CompromiseA` appears on the trace of the protocol, $\mathbf{k}_{R1}$ will be send in the network (given by fact `Out`), and will become visible to the adversary. Similarly, whenever action fact `CompromiseB` appears on the trace of the protocol, $\mathbf{k}_{R1}$ and R_2 will be send in the

```
rule Alice_sends1:
let
K_R11=break1(K_R1)
K_R12=break2(K_R1)
m_1=<A_ID,~N_1>
in
[A1(A, B, <K_R1, A_ID, r_A>), Fr(~N_1)]
--[A1_OUT(m_1), Role('A')]->
[Out(m_1), A2(A, B, <~N_1, K_R11, K_R12, r_A>)]

rule Bob_receives_sends1:
let
K_R11=break1(K_R1)
K_R12=break2(K_R1)
m_1=<A_ID,N_1>
m_B=senc{B, C_2, ~N_B, N_1}K_R11
in
[B1(B, A, <K_R1, A_ID, C_2, R_2, r_B>), In(m_1), Fr(~N_B)]
--[SendB(B, m_B), B1_IN(N_1, m_1), B1_OUT(m_B),
Genuine(A), Genuine(B), Role('B'), Involved(B, A, <'init'>)]->
[Out(<m_B, mac(K_R12,m_B)>), B2(B, A, <~N_B, C_2, R_2, K_R11, r_B>)]
```

Figure 3.19: Message exchange between Alice and Bob in Tamarin. Alice sends authentication request to Bob and he replies with a challenge.

network, and will become visible to the adversary. The two action factions are used later to define security properties.

The two rules in Fig. 3.19 define the initial message exchange within the between Alice and Bob, *i.e.*, Alice sends a request to Bob and he replies with a challenge. The premise of `Alice_sends1` rule calls the knowledge of the previous state using fact `A1` and generates a fresh random nonce `N_1`. The lines between `let - in` are used to define local formulas. In the first rule this is used to define that Alice breaks her key into two parts (*i.e.*, `K_R11`, `K_R12`), and her first message that contains $A_{\mathtt{ID}}$ and N_1 is denoted by `m_1`. Next, the action fact `A1_OUT` is used to relate subsequent rules through the, so called source lemmas. It has a corresponding `B1_IN` fact in the `Bob_receives_sends1` rule. As mentioned earlier, `Role` facts are used throughout the protocol definition to identify the agent responsible for the actions. Finally, in the conclusion of the rule Alice updates her state to `A2` and sends `m_1` using the fact `Out`.

Next, the premise of the `Bob_receives_sends1`, calls Bob's knowledge from state

B1; models the receiving of message m_1 using In fact and generates the fresh random nonce N_B. Then, Bob breaks his key into two parts and use the first to encrypt the message m_B, defined within the let - in space. The action facts B1_OUT and Role are used here for identical purposed as the action facts from the previous rule. The purpose of action facts SendB, Genuine and Involved will be defined later, when used as part of security properties. Finally, within the conclusion of the rule, Bob updates his state and sends message m_B to Alice, which contains B, C_2, N_S and N_1 encrypted using key $\mathbf{k}_{R1,1}$. He also sends the signed message denoted here as mac(K_R12,m_B). Both the message and the signed message are sent using the Out fact within the conclusion of the rule.

The last two rules used to define the protocol are given in Fig. 3.20. Within, the premise of the rule Alice_receives_sends2, 1) Alice calls the state A2; 2) receives the message m_B and the signed message; 3) generates a fresh random nonce N_A; 4) generates the syndrome s_A that corresponds to r_A. Note that, the action fact Eq denotes equality, *i.e.*, Alice uses this to verify the signature of the received message. Following from that the rule models he following: Alice runs the received challenge C_2 on her PUF, uses her FE to generate helper data (line 4) and key (line 5); generates the new challenges and corresponding responses R_3 (line 6) and R_4 (line 7); breaks her key into two parts (lines $8-9$); generates the her new one time alias identity (line 10); encrypts the message m_A using the first part of the key (line 11); and generates the session key (line 12). Finally, within the conclusion of the rule, she sends to Bob the message m_A, the signed message denoted as mac(KR_22e, m_A) and the helper data hd_2e. She also updates her state to A3 which now contains the session key. Similarly to the previous rules, some of the action facts are used relate consequent rules (line 15). The purpose of the other action facts will be defined later.

The premise of the last rule of the protocol definition Bob_receives2 calls Bob's state B2 and models the receiving of the message, the signed message and the helper data from Alice (using In fact). Within the let - in space Bob uses his FE, the stored response during the enrollment R_2 and the received helper data to reproduce the key KR_2e (line 23); next he breaks the key into two parts (line $24-25$); uses the first

```
1  rule Alice_receives_sends2:
2  let
3  m_B=senc{B, C_2, N_B, N_1}K_R11
4  hd_2e=fuzzyhd(puf(C_2))
5  KR_2e=fuzzyk(puf(C_2))
6  R_3=puf(h(<C_2,~N_A>))
7  R_4=puf(h(<h(<C_2,~N_A>),N_B>))
8  KR_21e=break1(KR_2e)
9  KR_22e=break2(KR_2e)
10 A_ID_new=h(<A, N_B, R_3>)
11 m_A=senc{A, ~s_A, ~N_A, R_3, R_4}KR_21e
12 sess_key=h(r_A)
13 in
14 [A2(A, B, <N_1, K_R11, K_R12, r_A>), In(<m_B,sig>), Fr(~N_A), Fr(~s_A)]
15 --[ReceiveA(A,m_B), A1_IN(N_B,m_B), SendA(A, m_A), A2_OUT(m_A),
16 Secret(sess_key), Genuine(A), Genuine(B), Commit(A, B, <'init'>),
17 Role('A'), Involved(A, B, <'resp'>), Eq(vermac(K_R12,m_B,sig),true),
18 AuthenticB(B, m_B)]->
19 [Out(<m_A, mac(KR_22e, m_A), hd_2e>), A3(A, sess_key)]
20
21 rule Bob_receives2:
22 let
23 KR_2e=rep(R_2, hd_2e)
24 KR_21e=break1(KR_2e)
25 KR_22e=break2(KR_2e)
26 A_ID_new=h(<A, N_B, R_3>)
27 m_A=senc{A,s_A,N_A, R_3, R_4}KR_21e
28 r_A=SKG(s_A, r_B)
29 sess_key=h(r_A)
30 in
31 [B2(B, A, <N_B, C_2, R_2, K_R11, r_B>),  In(<m_A, sigm_A, hd_2e>)]
32 --[B2_IN(N_A, m_A), ReceiveB(B,m_A), Secret(sess_key), Genuine(B),
33 Genuine(A), Role('B'), Eq(vermac(KR_22e,m_A,sigm_A),true),
34 Commit(B, A, <'resp'>), AuthenticA(A, m_A)]->
35 [B3(B, sess_key)]
```

Figure 3.20: Message exchange between Alice and Bob in Tamarin. Alice receives the challenge from Bob and sends M_A in response.

part to decrypt the message from Alice (line 26) and the second the verify the signature (using action fact `Eq`); then using received syndrome `s_A` and the generated sequence `r_B` he performs the `skg` procedure to obtain the session key (line 29). Finally, within the conclusion of the rule Bob updates his knowledge to `B3` which now includes the session key. This concludes the protocol definition.

Restrictions: Before introducing the security properties, a set of restrictions are de-

```
restriction Equality:
  "All x y #i. Eq(x,y) @i ==> x = y"

restriction Enrollment_is_performed_once:
  "All #i #j. Once( ) @ i & Once( ) @ j ==> #i = #j"

restriction Enrollment_is_secure:
  "not(Ex m #i #j. Secure(m) @ #i & K(m) @ #j)"
```

Figure 3.21: Restrictions within the protocol definition.

```
lemma executable: exists-trace
"Ex A B m #i #j.
    SendA(A,m) @i & ReceiveB(B,m) @j & i < j"
```

Figure 3.22: Lemma used to prove executability property of the protocol definition.

fined in Fig. 3.21. The restriction `Equality` defines that, if an action fact `Eq(x,y)` appears on the trace of the protocol, it implies `x = y`. This action fact is used to verify the signatures of Alice and Bob within the protocol definition. The restriction `Enrollment_is_performed_once` defines that if the action fact `Once( )` appears on the trace at moments `i` and `j`, follows that `i = j`. Usually, in Tamarin, the adversary can execute each rule unlimited number of times, taking different roles each execution. Therefore, by adding this restriction, the rule which contains the fact `Once( )` can be executed only once, which here is the `Enrollment` rule. Finally, the restriction `Enrollment_is_secure` defines that, each parameter within the action fact `Secure` is secret of the adversary. This fact is used within the `Enrollment` rule and it defines that the parameters exchanged during the enrollment phase are secret from the adversary.

Security properties: The following defines the set of security properties used to verify the proposed protocol.

***Executability*:** As discussed in Appendix A security properties are modelled as lemmas. Therefore, to ensure executability property of the defined protocol, this section begins with a sanity check of the model. The lemma used to prove the executability of the defined model is given in Fig. 3.22. The lemma contains the keyword `exists-trace` defining that, it might not hold for all traces, however, it returns a true statement if possible

in just one or few traces. The lemma contains the action facts `SendA` and `ReceiveB`, which are part of the last two rules, illustrated in Fig. 3.20. It can be interpreted as follows: there exists a trace such that Alice sends a message at time instant `i`, Bob receives it at time instant `j`, such that `i` happens before `j`. The proof of the lemma is given in 3.23, note, in the final, printed thesis, this diagram will be printed as a fold out A3 sheet.

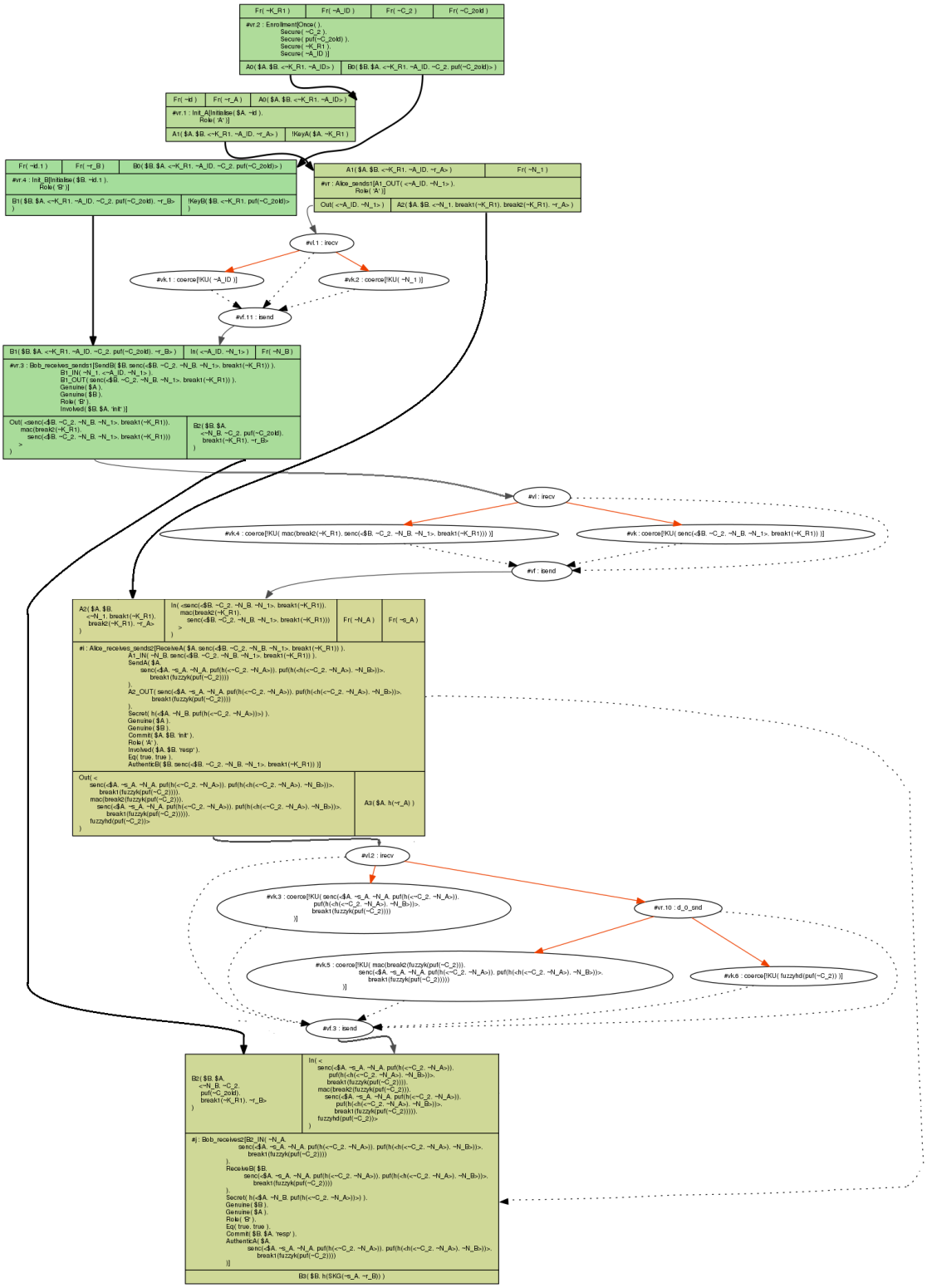

Figure 3.23: Proof for executability of the protocol definition.

```
lemma aliveness:
    "All A B t #i.
    Commit(A,B,t)@i
    ==> (Ex id #j. Initialise(B,id) @ j)
        | (Ex C #r. CompromiseA(C) @ r & Genuine(C) @ i)"
```

Figure 3.24: Lemma used to prove aliveness property of the protocol definition.

```
lemma weak_agreement:
    "All A B t1 #i.
    Commit(A,B,t1) @i
    ==> (Ex t2 #j. Involved(B,A,t2) @j)
        | (Ex C #r. CompromiseB(C) @ r & Genuine(C) @ i)
        | (Ex C #r1. CompromiseA(C) @ r1 & Genuine(C) @ i)"
```

Figure 3.25: Lemma used to prove weak agreement between Alice and Bob within the protocol definition.

***Authentication*:** The authentication specifications are modelled using the hierarchy defined in [205]. The hierarchy of the specifications from the weakest to the strongest is: aliveness, weak agreement, non-injective agreement, injective agreement.

Aliveness

A protocol guarantees to the initiating party A aliveness of another party B if, A completes a run of the protocol apparently with B, then B has previously completed a run of the protocol, but not necessarily with A. The property is modelled as in Fig. 3.24.

Weak agreement

A protocol guarantees to the initiating party A weak agreement with a responding party B if, whenever A completes a run of the protocol apparently with B, therefore, B has previously completed a run of the protocol, apparently with A. However, B may not been acting as a responder. This property is modelled as illustrated in Fig. 3.25. It can be interpreted as follows: For the initiating party A, denoted by action fact `Commit`, there exists a responding party B, denoted by action fact `Involved` or someone who claims to be a genuine party (either Alice or Bob) has been compromised.

Non-injective agreement

A protocol guarantees to A a non-injective agreement with B on a message t if, whenever

```
lemma non_injective_agreement:
    "All A B t #i.
    Commit(A,B,t) @i
    ==> (Ex #j. Involved(B,A,t) @j)
        | (Ex C #r. CompromiseB(C) @ r & Genuine(C) @ i)
        | (Ex C #r. CompromiseA(C) @ r & Genuine(C) @ i)"
```

Figure 3.26: Lemma used to prove non-injective agreement between Alice and Bob within the protocol definition.

```
lemma injective_agreement:
    "All A B t #i.
    Commit(A,B,t) @i
    ==> (Ex #j. Involved(B,A,t) @j
        & j < i
        & not (Ex A2 B2 #i2. Commit(A2,B2,t) @i2
        & not(#i2 = #i)))
            | (Ex C #r. CompromiseB(C)@r & Genuine(C) @i)
            | (Ex C #r. CompromiseA(C)@r & Genuine(C) @i)"
```

Figure 3.27: Lemma used to prove injective agreement between Alice and Bob within the protocol definition.

A completes a run of the protocol apparently with B (who has a defined role), therefore, B has previously completed a run of the protocol, apparently with A and was acting in the defined role and the parties agreed on message t. However, the property does not guarantee full relation between the runs of protocols by A and B, *i.e.*, A may think that he ran the protocol twice, but B was involved only in one run. This property is modelled in Fig. 3.26.

Injective agreement

Finally, a protocol guarantees to A an injective agreement with B on a message t if, whenever A completes a run of the protocol apparently with B (who has a defined role), therefore, B has previously completed a run of the protocol, apparently with A and was acting in the defined role and the parties agreed on message t. Furthermore, there is a unique matching partner for each run of the protocol, *i.e.*, for each `Commit` by a party, there is a corresponding and unique `Involved` by the other party. This addition guarantees there is freshness with each run of the protocol, used to prevent replay attacks. The

```
lemma message_authentication_m_A:
    "All a m #i.
    AuthenticA(a,m) @i
    ==> (Ex #j. SendA(a,m) @j & j<i)
        | (Ex B #r. CompromiseB(B)@r & Genuine(B) @i & r < i)"

lemma message_authentication_m_S:
    "All b m #i.
    AuthenticB(b,m) @i
    ==> (Ex #j. SendB(b,m) @j & j<i)
        | (Ex B #r. CompromiseA(B)@r & Genuine(B) @i & r < i)
        | (Ex B #r. CompromiseB(B)@r & Genuine(B) @i & r < i)"
```

Figure 3.28: Lemmas used to prove message authentication of M_A and M_S, respectively.

injective agreement property is modelled in Fig. 3.27.

Message authentication The two lemmas in Fig. 3.28 are used to prove the authenticity of the signed messages from Alice and Bob, *i.e.*, M_A and M_S, respectively. The first lemma can be interpreted as follows: There exist a time instants `j` and `i` such that Alice sends the message to Bob at `i` and she is still authentic user after that at time `j` (*i.e.*, `j<i`, or Bob has been compromised. This results from PUF unclonability property and that fact that, if Alice is compromised the adversary will obtain only the key used to sign the first M_S and not M_A. However, if Bob is compromised the adversary will first obtain R_2 and then the key used to sign M_A. On the other hand, the authenticity of M_S, proved by `message_authentication_m_S`, requires that none of the parties is compromised.

***Privacy properties*:** The term privacy combines set of properties: while security provides soundness, privacy provides protection against unauthorised access or linking identities [206]. Privacy properties are: anonymity and untraceability. The former ensures that an adversary cannot determine which user is involved within a run of a protocol; the latter ensures that an adversary cannot determine whether the same user is involved into two separate authentication sessions [207]. Privacy is a vital feature for numerous applications, such as healthcare [208], anonymous conferences [209], e-voting [210], e-cash [211].

In Tamarin both properties are modelled using observation equivalence. Observa-

tion equivalence assumes two systems (*i.e.*, two instances of the protocol) and it is used to prove that an adversary cannot distinguish the systems. This can be specified using the operator `diff` by simply adding it to the preferred rules. To build the two systems the following is added to the rule `Alice_receives_sends2` defined in Fig. 3.20: 1) in the premise a new fresh random variable `Random_variable` is generated 2) `Out(diff(A_ID_new, Random_variable)` 3) the rules for compromising an agent (*i.e.*, 3.18) are removed from the protocol definition. Following the above, during the first run of the protocol the left parameter `A_ID_new` will be send out, then during the second run of the protocol the right parameter `Random_variable` will be sent out. If an adversary can distinguish the new one-time alias ID of Alice from a random variable it implies that privacy is not satisfied.

Figure 3.29 shows that, the privacy of Alice is preserved unless one of the parties is compromised. As shown in the figure, regardless of which parameter the adversary observes the final result is equality, *i.e.*, cannot be distinguished. Therefore, the proposed protocol satisfies the privacy properties.

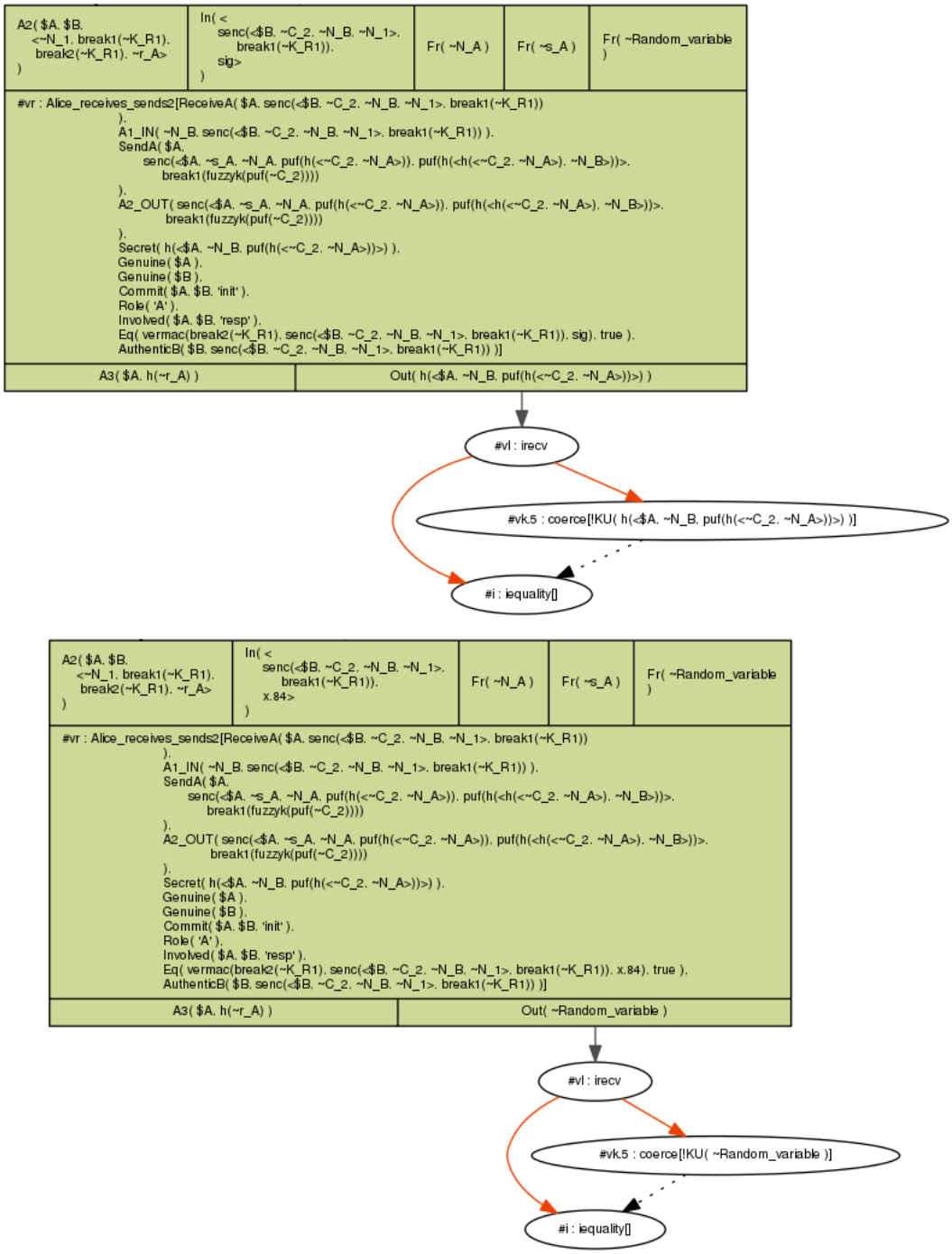

Figure 3.29: Proof for observational equivalence: TOP: Using left parameter; BOTTOM: using the right parameter.

***Secrecy properties*:** Due to the fact that the secrecy of most of the parameters was proved using MB logic, Tamarin is used here to prove perfect forward secrecy of the pro-

```
lemma PFS_session_key_Alice:
    "All x #i.
    Secret(x) @i & Role('A') @i==>
    (not(Ex #j. K(x)@j)) "

lemma PFS_session_key_Bob:
  "All x #i.
  Secret(x) @i & Role('B') @i==>
  (not(Ex #j. K(x)@j)) "
```

Figure 3.30: Lemmas used to prove perfect forward secrecy of the session key, from Alice's and Bob's, perspective.

```
/* All well-formedness checks were successful. */
==============================================================================
summary of summaries:

  aliveness (all-traces): verified (12 steps)
  weak_agreement (all-traces): verified (29 steps)
  noninjective_agreement (all-traces): verified (29 steps)
  injective_agreement (all-traces): verified (65 steps)
  message_authentication_m_S (all-traces): verified (8 steps)
  message_authentication_m_A (all-traces): verified (58 steps)
  PFS_session_key_Alice (all-traces): verified (8 steps)
  PFS_session_key_Bob (all-traces): verified (14 steps)
  types (all-traces): verified (33 steps)
  executable (exists-trace): verified (11 steps)
```

Figure 3.31: Verification of all modelled properties.

tocol in regards to the session key. The lemmas used to prove this property are illustrated in Fig. 3.30. Note that, each run of the protocol both parties generate a new session key that is extracted from the randomness of the wireless channel. Therefore, the perfect forward secrecy of the session key is preserved even if both parties get compromised by an attacker.

Verification of all properties

This section concludes with the verification of all properties. This is illustrated in Fig. 3.31. As it can be seen all of the discussed properties of the protocol have been successfully verified.

Table 3.2: Comparison of existing PUF-based solutions for authentication

Ref.	Privacy preserving	Location factor	Correct noisy PUF responses	Two authentication keys (two parties)	Update CRPs: not all stored in a database	Session key do not rely on PRNG
[154, 169]	No	No	No	No	Yes	No
[216, 217]	No	No	No	No	No	No
[162, 218]	Yes	No	Yes - FE	No	Yes	No
[159, 219]	Yes	No	Yes - reverse FE	No	Yes	No
[153]	Yes	Yes	No	No	Yes	No
This thesis	Yes	Yes	Yes - reverse FE	Yes	Yes	Yes

3.6 Brief discussion

This work proposed a fast and lightweight authentication solution for resource constrained systems. While, existing techniques used in IoT systems, such as the extensible authentication protocol-transport layer security (EAP-TLS), could be used as an authentication mechanism, these are computationally intensive and can lead to significant latency [212, 213]. Measurements performed on current public key operations within EAP-TLS on common devices (such as IoT) give average authentication and key generation times of approximately 160 ms in static environments and this can reach up to 336 ms in high mobility conditions [214].

Therefore, the motivation for using a PUF authentication scheme in conjunction with SKG is to exclude all of the computationally intensive operations required by EAP-TLS, which use modulo arithmetic in large fields. In general, PUF authentication protocols have very low computational overhead and require overall authentication times that can be less than 10 ms [153, 215]. However, the proposed PUF-based schemes in literature do not provide a full solution (as illustrated in Table 3.2). Most of the protocols rely on PUFs as a single security factor and this can expose the system to a variety of threats, especially in an IoT scenario [128]. Therefore, combining two or more independent credentials is essential. As an example, the work within this chapter combines PUFs, location factor

and channel characteristics in order to built a secure multi-factor authentication protocol. Next, it can be seen that other existing schemes cited in Table 3.2 propose to reuse the same authentication key in order to authenticate both parties. However, if the key gets leaked the adversary could impersonate both Alice and Bob, therefore, the scheme proposed within this chapter assumes that each party is authenticated through a unique key. Finally, adding the key generation scheme, proposed in Section 2.2.1 and extended in this thesis, obviates the need of using PRNG in the session key generation process. Furthermore, the SKG implies a negligible penalty in terms of latency, as it requires just a hashing operation and (syndrome) decoding. Hashing mechanisms such as SHA-256 performed on an IoT device require less than 0.3ms [215, 220]. Regarding the decoding, if it assumes the usage of standard LDPC or BCH error correcting mechanisms, even in the worst-case scenario with calculations carried out as software operations, the computation is trivial compared to the hashing and requires less computational overhead [221].

3.7 Summary

This chapter presented a full mutual authentication mechanism which can be used for secure session establishment. The security properties of the proposed solution were validated proving its resistance to numerous types of attacks. Additionally, a novel resumption protocol was developed that can allow for data transmission within 0-RTT. Finally, a closed form expression of the information theoretic bound for SKG within short blocklength was derived, allowing for fair comparison of existing SKG mechanisms.

The novel ideas in regards to the SKG procedure within this chapter are mostly related to expanding its applicability. Therefore, Chapter 4 will explore the SKG in more detail and will optimise the process for the scenarios of a loose and stringent delay constrained systems.

Chapter 4

Optimised key generation for delay-constrained wireless systems

---

This chapter introduces a physical layer security mechanism that jointly optimises physical layer SKG and traditional encrypted data channel (using the key from SKG). The solution gives optimal power and subcarrier allocation.

Part of this chapter was presented as an invited poster presentation at the "Munich Workshop on Coding and Cryptography (MWCC) 2018" [43]. Some of the results given in this chapter were presented at the "The International Wireless Communications and Mobile Computing Conference (IWCMC 2019)" [13]. Further set of results from this chapter were presented at the "IEEE Global Communications Conference (Globecom 2019)" [14]. Finally, after further improvements part of the work presented in this chapter was published in the special issue "Physical Layer Security Solutions for 5G-and-Beyond" within the "Springer EURASIP Journal on Wireless Communications and Networking" [12].

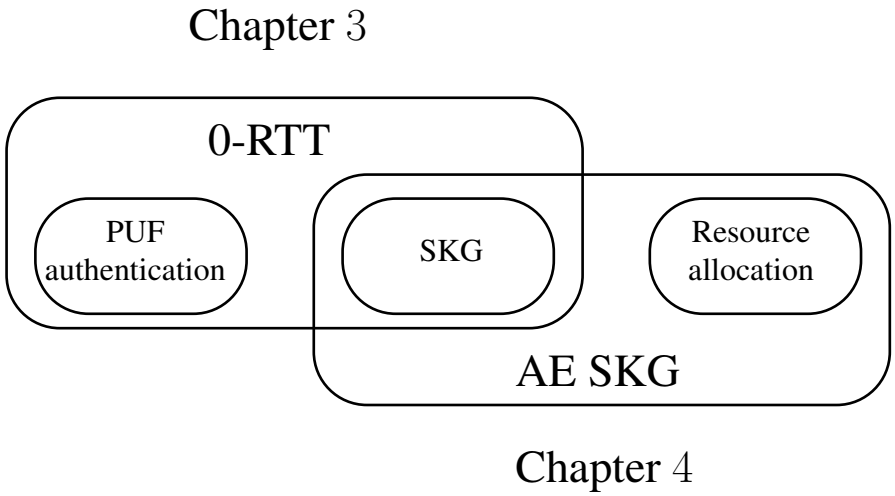

Figure 4.1: Roadmap from Chapter 3 to Chapter 4.

4.1 Introduction

Building on the results presented in Chapter 3, this study focuses on optimising the SKG process. As illustrated in Fig. 4.1, this chapter investigates a fast implementation of an authenticated encryption (AE) SKG and proposes a pipelined (parallel) scheduling method for optimal resource allocation at the PHY. The utilised AE primitive [222–224] jointly optimises data rates and key generation rates. The motivation behind this approach is latency reduction; data could be immediately transmitted whenever they become available, which can be critical in latency sensitive applications such as V2X, URLLC and haptic communication systems [225–227].

The contributions of this chapter are as follows:

1. Proposal of a fast implementation of the AE SKG based on pipelining of key generation and encrypted data transfer.

2. Identifying the optimal resource allocation for the *parallel* approach in a three different scenarios:

    - Under power and security constraints;
    - Under power, security and rate constraints;

- Under power, security, rate and delay constraints;

3. Proposal of a heuristic algorithm of linear complexity that finds the near-optimal subcarrier allocation with negligible loss in terms of efficiency.

4. Numerical evaluation of the achievable sum data rate for short and long term power constraints.

5. Numerical comparison of the efficiency of the proposed *parallel* approach with a *sequential* approach where SKG and data transfer are performed sequentially. This comparison is performed in two delay scenarios:

   - When a relaxed quality of service (QoS) delay constraint is in place;
   - When a stringent QoS delay constraint is in place.

Note that the work presented in this chapter assumes that Alice and Bob have authenticated, *e.g.* by using the method introduced in Chapter 3. Following from the above they can use the SKG solutions presented in this chapter as a rekeying technique. Next, Chapter 5 will discuss the optimal strategies of Alice and Bob in the presence of an active attacker who tries the interrupt the SKG process.

4.2 Respective background

4.2.1 Optimisation methods

A set of optimisation methods are used throughout the present chapter. These include: convex optimisation and combinatorial optimisation (Knapsack problem).

Convex optimisation

Convex optimisation is a field in mathematics that studies minimisation (and maximisation) of convex (and concave) functions. The applicability of convex optimisation ranges from finance and statistics [228] to communication systems and signal processing [229].

Convex optimisation problem is a problem with convex (concave) objective function with convex (concave) feasible set $\mathcal{S}_{cvx}$. In this regard, a function f defined on the set $\mathcal{S}_{cvx}$, such that $f : \mathcal{S}_{cvx} \rightarrow \mathbb{R}$, is convex if:

$$f((1-\lambda)(x_1) + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2), \tag{4.1}$$

and concave if:

$$f((1-\lambda)(x_1) + \lambda x_2) \geq (1-\lambda)f(x_1) + \lambda f(x_2), \tag{4.2}$$

where $\forall x_1, x_2 \in \mathcal{S}_{cvx}$ and $\forall \lambda \in [0, 1]$ [230]. Following from the above, a standard convex optimisation problem can be defined as:

$$\min_{x} \quad f_0(x) \tag{4.3}$$

$$\text{subject to} \quad f_{\text{convex},i}(x) \leq 0 \quad i = 1, \ldots, m, \tag{4.4}$$

$$f_{\text{affine},j}(x) = 0 \quad j = 1, \ldots, n, \tag{4.5}$$

where $x \in \mathbb{R}$ is the parameter of interest (optimisation variable), f_0 is the objective function and $f_{\text{convex}}, f_{\text{affine}}$ represent the set of constraints. Note that functions f_0 and f_{convex}, are convex and satisfy Eq. (4.1), and functions f_{affine} are affine. Overall, Eq. (4.3) defines the problem of finding the value of x that minimises function f_0, such that constraints (4.4), (4.5) are satisfied (similarly, using a $\max$ argument one can optimise concave functions). A well-known method used for solving convex optimisation problems is the Lagrange multipliers method [231]. The Lagrangian function of the problem in Eq. (4.3) is formed as:

$$\mathcal{L}_x = f_0(x) + \sum_{i=1}^{m} \lambda_i f_{\text{convex},i}(x) + \sum_{j=1}^{n} v_j f_{\text{affine},j}(x), \tag{4.6}$$

where λ_i is called Lagrange multiplier and it is associated with the i-th $f_{\text{convex},i}(x) \leq 0$ constraint; and v_j is the Lagrange multiplier associated with the j-th $f_{\text{affine},j}(x) = 0$ constraint. Next, to identify the optimal solution of Eq. (4.3), which is denoted by x^*, the

following set of conditions must be satisfied (also known as Karush-Kuhn-Tucker (KKT) conditions [232, 233]):

1. Stationary condition - the gradient of (4.6) should vanish at x^*, *i.e.*,:

$$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_{\text{convex},i}(x^*) + \sum_{j=1}^{n} v_j^* \nabla f_{\text{affine},j}(x^*) = 0 \tag{4.7}$$

2. Primal feasible condition at x^*:

$$f_{\text{convex},i}(x^*) \leq 0 \quad i = 1, \ldots, m, \tag{4.8}$$

$$f_{\text{affine},j}(x^*) = 0 \quad j = 1, \ldots, n. \tag{4.9}$$

3. Dual feasible condition:

$$\lambda_i^* \geq 0, \quad i = 1, \ldots, m \tag{4.10}$$

4. Complementary slackness:

$$\lambda_i^* f_{\text{convex},i}(x^*) = 0, \quad i = 1, \ldots, m \tag{4.11}$$

The value x^* that satisfies the above conditions minimises $\mathcal{L}_x$ over x. Furthermore, for any convex optimisation problem that has a differentiable objective function and constraints, all points that satisfy the KKT conditions are optimal [230]. Convex optimisation is used as a tool in Chapter 4 to identify the optimal resource allocation policy of Alice and Bob during the SKG process.

0 – 1 Knapsack optimisation

Knapsack problem is a problem defined within the premise of combinatorial optimisation [234]. Assuming a set of items, each with weight and value, by solving the problem one can identify the optimal combination of items such that the sum of the weights is less or equal than a specific threshold and the sum of the values is as high as possible [235].

The 0 – 1 Knapsack problem follows the above description but also limits the copies of each items to either zero or one. Assuming a set of items $i = 1, \ldots, n$, a 0 – 1 Knapsack problem can be formulated as follows:

$$\max_{x_i \in \{0,1\}} \quad \sum_{i=1}^{n} v_i x_i \tag{4.12}$$

$$\text{subject to} \quad \sum_{i=1}^{n} w_i x_i \leq W_{\text{th}} \tag{4.13}$$

$$x_i \in \{0, 1\}, \tag{4.14}$$

where the subscript i denotes the specific item, x_i denotes the number of copies of each item, v_i denotes the value of each item, w_i denotes the weight of each item and W_{th} defines the threshold that defines the maximum weight. The problem can be solved recursively [235] by finding the maximum value of:

$$V[i, y] = \max \{V[i-1, y], V[i-1, y-w_i] + v_i\}, \quad y = 1, \ldots, W, \tag{4.15}$$

where $V[i, y]$ determines the profit of each combination of items, $i = 1, \ldots, n$, and $y = 1, \ldots, W$ is a positive integer. In the present chapter, 0 – 1 Knapsack optimisation is used in the form of a subset-sum Knapsack problem where the weights and values of the item are equal. In the particular problem the Knapsack items refer to a set of available subcarriers and the solution gives optimal subcarrier allocation.

4.2.2 Effective capacity

One of the major requirements of an IoT application is to keep the delay below a certain threshold, according to which the QoS guarantees should be satisfied. In this thesis a flexible delay QoS model was employed using the theory of large deviations (Gärtner-Ellis theorem [236, 237]) that allows defining the metric of the effective capacity on block fading additive white Gaussian noise (BF-AWGN) channels. The theory of large deviations and Gärtner-Ellis theorem are well established and widely employed mechanisms, espe-

cially in the context of effective bandwidth and effective capacity. Therefore, for more details on these concepts the readers are referred to [236–239].

The effective capacity [240] denotes the maximum constant arrival rate that can be served by a given service process, while guaranteeing a required statistical delay provisioning and is closely related to the concept of the effective bandwidth [241, 242].

Theory of effective bandwidth considers that the distribution of the queue length has an exponential tail. In fact, the effective bandwidth of a time-varying source is the minimum amount of bandwidth required to satisfy specified QoS requirments, or it could be represented as the source rate with which the packet loss rate decays exponentially. Therefore, for an arrival process with variable rate one can use the Gartner-Ellis theorem to characterise the probability that this source will behave like a constant rate source of rate α for a time t (note, $A(0,t)$ denotes the number of packets in t time instances), such that, if:

$$\Lambda(\theta) = \lim_{t\to\infty} \frac{1}{t} \ln E[e^{\theta A(0,t)}], \quad \forall\theta \in \mathbb{R}, \tag{4.16}$$

exists and it is everywhere differentible then the effective bandwidth function is [241]:

$$a^*(\theta) = \frac{\Lambda(\theta)}{\theta}. \tag{4.17}$$

The effective bandwidth α^*, defines that the probability of the queue length exceeding a certain threshold, θ, decays exponentially fast as θ increases. The parameter θ indicates the exponential decay rate of the QoS violation probability. A smaller θ corresponds to a slower decay rate, which implies that the system can only provide looser QoS (long delay) and a higher θ, corresponds to a more stringent delay constraint.

The effective capacity is the dual function of the effective bandwidth. While, the effective bandwidth is the minimum constant service rate required by a given arrival process for which θ is fulfilled, the effective capacity is the maximum constant arrival rate that a given service process supports in order to guarantee θ [243–245]. Assuming the sequence $R_i, i = 1, 2, ..., N$ denotes a discrete-time stationary and ergodic stochastic service process and $C = \sum_{i=1}^{N} R_i$ is a sum of the service process. Therefore, if the Gartner-Ellis

limit of C expressed as:

$$\Lambda_C(\theta) = -\lim_{t\to\infty} \frac{1}{\theta t} \ln(E\{e^{-\theta C}\}) \quad \forall \theta \in \mathbb{R}, \tag{4.18}$$

exists and is everywhere differentiable. Then the effective capacity of the service process, denoted by $E_C(\theta)$ is given by:

$$E_C(\theta) = -\frac{\Lambda_C(-\theta)}{\theta} = -\lim_{t\to\infty} \frac{1}{\theta t} \ln(E\{e^{-\theta C}\}). \tag{4.19}$$

Note that, if $R_i, i = 1, 2, ..., N$ are uncorrelated, the effective capacity $E_C(\theta)$ reduces to:

$$E_C(\theta) = -\frac{1}{\theta} \ln(E\{e^{-\theta R_i}\}). \tag{4.20}$$

As it can be seen above, similarly to the effective bandwidth, smaller θ implies looser QoS guarantee and larger θ corresponds to stringent delay requirements. Note that, the theory of effective capacity is based on the assumption of infinite buffer length, and therefore, it could give only probabilistic guarantee on the achievable rates. However, it can provide a good intuition on how a real system will behave given a specific delay constraint.

4.3 Employed methods and system model

Having now presented the necessary background material, the rest of this chapter proposes a new technique for optimising AE SKG based on the combination of several methods. These methods are described below where the novelty that each one brings to the overall solution is identified. Finally, as in Chapter 3, the system model assumes a commonly used adversarial model with an active man-in-the-middle attacker (Eve) and a pair of legitimate users (Alice and Bob). Furthermore a rich Rayleigh multipath environment is assumed, where the legitimate parties communicate over a BF-AWGN channel.

Secret key generation

As in the previous chapter, to ensure privacy after authenticating each other, Alice and Bob obtain a shared secret key following the three-step SKG procedure, described in Section 2.2.1, *i.e.*, i) advantage distillation; ii) information reconciliation; and, iii) privacy amplification. Next, both parties encrypt / decrypt the exchanged information using the shared secret key. The novelty in regards to this method is the combination of the second phase of the SKG process (information reconciliation) with data exchange.

Authenticated encryption SKG

To eliminate the possibility of tampering attacks this work builds on the SKG process to introduce a new AE SKG method. AE can simultaneously guarantee confidentiality and message integrity. In the proposed AE SKG method, side information and encrypted data transfer are pipelined. Throughout this work this is referred the *parallel* transmission method, described in further detail in the next item. This is a novel AE approach entirely based on the physical layer SKG process.

Pipelined transmission

Unlike traditional *sequential* methods, where key generation and data exchange are separated in subsequent frames, in the proposed *parallel* method, the key generation is pipelined with the encrypted data transfer, *i.e.*, side information and data encrypted with the key that corresponds to the side information are transmitted over the same 5G resource block(s) (*i.e.*, in (multiple) frames of 12 OFDM subcarriers, some of the subcarriers are used for data transmission and the others to transmit the side information.) The proposed mechanism achieves optimal power and subcarrier allocation within each frame, in certain scenarios this is a modified water-filling solution. This is an innovative technique that improves the performance of wireless networks – the work is supported by numerical evaluation.

Joint PHY/MAC delay analysis

To analyse the performance of the proposed *parallel* scheduling two methods are used: i) maximisation of the long-term average data rate between Alice and Bob when delay requirements are not specified; ii) maximisation of the effective data rate between Alice and Bob while meeting a probabilistic delay constraint. The latter is based on the theory of *effective capacity* [240] and analyses the scheme's *effective rate*. This allows to identify the optimal resource allocation while satisfying a delay-outage probability constraint.

Optimisation methods

Finally, to optimise the pipelined transmission, this study takes into consideration practical wireless aspects such as the impact of imperfect CSI measurements and formulate several optimisation problems to find the optimal strategy for Alice and Bob. To formulate and solve the optimisation problem set of techniques are employed, such as, combinatorial optimisation (Knapsack), order statistics and convex optimisation. The study does not look into improving these methods, instead it uses them as a tool to identify the optimal solution.

The PHY system model, depicted in Fig. 4.2, assumes two legitimate parties, referred to as Alice and Bob, who wish to establish a symmetric secret key using as a source of shared randomness the wireless fading coefficients. Throughout this work a rich Rayleigh multipath environment is assumed, such that the fading coefficients rapidly decorrelate over short distances [37]. Furthermore, Alice and Bob communicate over a BF-AWGN channel that comprises N orthogonal blocks (*e.g.*, in the frequency domain), which, for simplicity, will be referred to as subcarriers. Without loss of generality the model assumes a unit AWGN variance in all links. The fading coefficients, denoted by $h_j, j = 1, \ldots, N$, are assumed to be independent and identically distributed (i.i.d) zero-mean circularly-symmetric complex Gaussian random variables $h_j \sim \mathcal{CN}(0, \sigma_h^2)$, such that after the pilot exchange Alice and Bob obtain observations $x_{A,j}, x_{B,j}$, respectively, on the j-th subcarrier

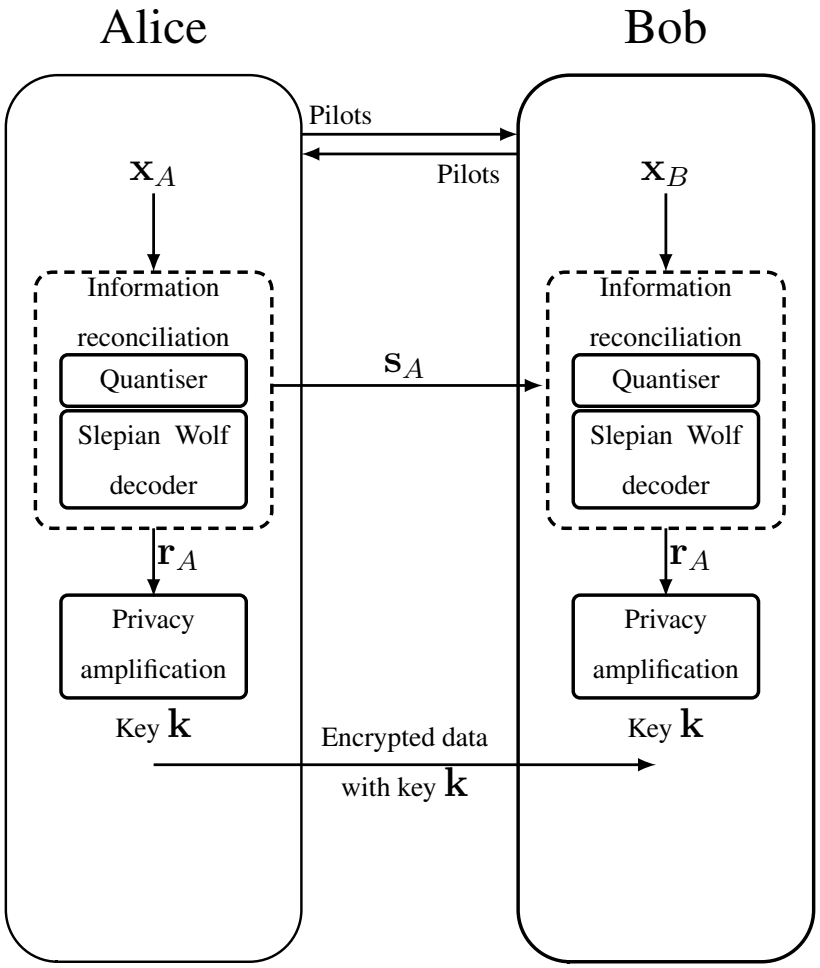

Figure 4.2: Secret key generation between Alice and Bob.

that can be expressed as:

$$x_{A,j} = \sqrt{P} h_j + z_{A,j}, \tag{4.21}$$

$$x_{B,j} = \sqrt{P} h_j + z_{B,j}, \tag{4.22}$$

$j = 1, \ldots, N$, where $z_{A,j}, z_{B,j}$ denote zero-mean, unit variance circularly-symmetric complex AWGN random variables, $(z_{A,j}, z_{B,j}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_2)$, and $\sqrt{P}$ denotes the transmit power during the pilot exchange.

Note, irrespective of whether SKG or data transfer is performed, Alice and Bob need to exchange pilot signals to obtain estimates of their reciprocal CSI. These CSI estimates

can be subsequently used to either conduct the secret key generation or be used to optimally allocate the available power to maximise the data transfer rate using a waterfilling algorithm. In further detail, after obtaining the channel observations, if SKG is to be performed on a specific subcarrier, then it is necessary for Alice (or Bob) to further transmit side information as part of the "information reconciliation" phase; *e.g.*, the syndrome (s_A in Fig. 4.2) of the Slepian-Wolf decoder output if block codes are used. Note that, as discussed in Section 2.2.1, the final step of the SKG process – privacy amplification – in which a common key is extracted at both Alice and Bob is performed locally without any further information exchange. If on the other hand, a given subcarrier is chosen for data transfer, then the estimated CSI will be used to optimise the power allocation. Bearing this in mind, this study assumes an initial authentication and SKG are performed (*e.g.* by using the method presented in Chapter 3) to generate the first keys to be used for the encryption of the first set of data.

Under the system model in Fig. 4.2, the SKG rate on any subcarrier is (note that the noise variances are here normalised to unity for simplicity) [37, 74]:

$$R_{SKG,j} = \log_2\left(1 + \frac{P\sigma^2}{2 + \frac{1}{P\sigma^2}}\right), \tag{4.23}$$

while the corresponding *minimum* necessary reconciliation rate has been show to be $\mathsf{H}(h_{B,j}|h_{A,j})$ [36]. The following sections will focus on optimising the allocation of resources (frequency and power) in multicarrier systems in which keys are generated at the physical layer.

4.4 Authenticated encryption protocols using SKG

To develop robust protocols that can withstand tampering attacks, standard symmetric key block ciphers and message authentication (MAC) schemes can be used in conjunction with SKG. As a sketch of such a protocol, let us assume a system with three parties: Alice and Bob, who wish to exchange messages with confidentiality and integrity, and

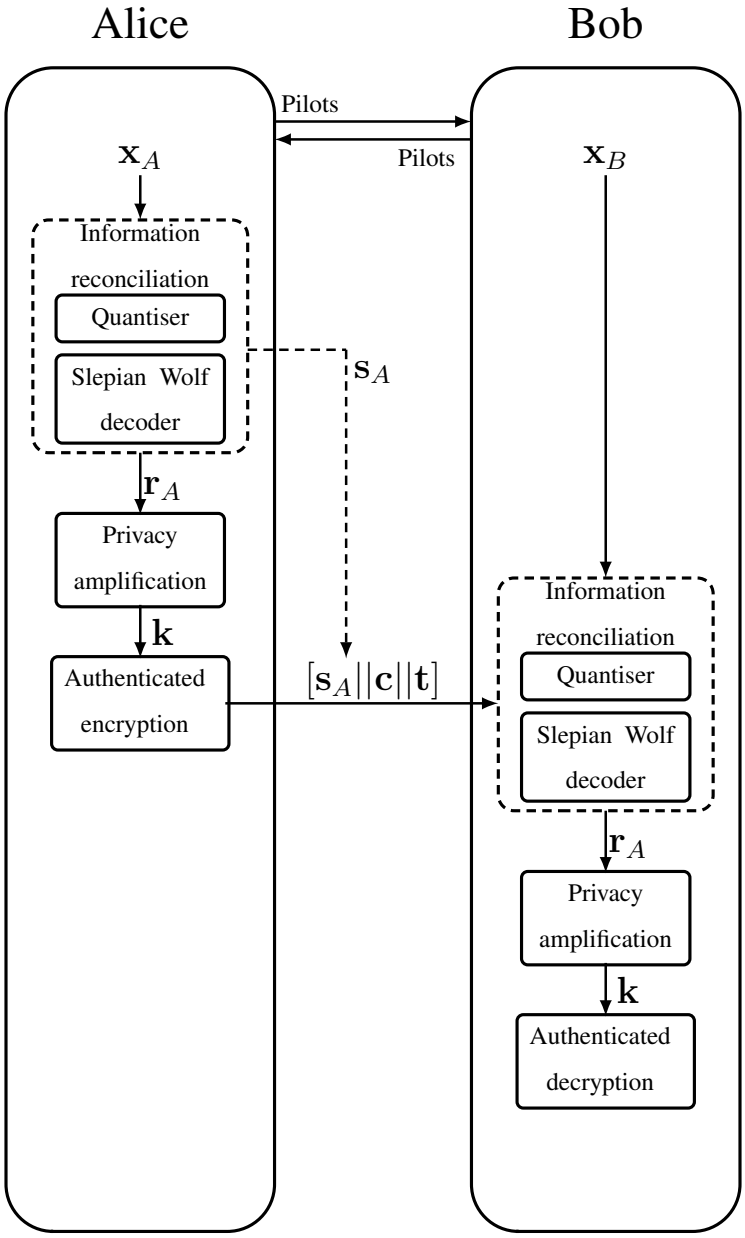

Figure 4.3: Pipelined SKG and encrypted data transfer between Alice and Bob.

Eve, that can act as a passive and active attacker. Alice wishes to transmit over a wireless multipath channel a secret message $\mathbf{m}$ with size $|\mathbf{m}|$ to Bob. The following algorithms are employed: the SKG scheme, a symmetric encryption algorithm denoted by `Es` with corresponding decryption `Ds` and a MAC denoted by `Sign` with a corresponding verification algorithm `Ver`. Using the above, a hybrid crypto-PLS system for AE SKG can be

built as follows:

1. The SKG procedure is launched between Alice and Bob generating a key and a syndrome $\mathtt{G}(\mathbf{h}) = (\mathbf{k}, \mathbf{s}_A)$.

2. Alice breaks her key into two parts $\mathbf{k} = \{\mathbf{k}_e, \mathbf{k}_i\}$ and uses the first to encrypt the message as $\mathbf{c} = \mathtt{Es}(\mathbf{k}_e, \mathbf{m})$. Subsequently, using the second part of the key she signs the ciphertext using the signing algorithm $\mathbf{t} = \mathtt{Sign}(\mathbf{k}_i, \mathbf{c})$ and transmits to Bob the extended ciphertext $[\mathbf{s}_A \| \mathbf{c} \| \mathbf{t}]$, as it is depicted in Fig. 4.3.

3. Bob checks first the integrity of the received ciphertext as follows: from $\mathbf{s}_A$ and his own observation he evaluates $\mathbf{k} = \{\mathbf{k}_e, \mathbf{k}_i\}$ and computes $\mathtt{Ver}(\mathbf{k}_i, \mathbf{c}, \mathbf{t})$. The integrity test will fail if any part of the extended ciphertext was modified, including the syndrome (that is sent as plaintext); for example, if the syndrome was modified during the transmission, then Bob would not have evaluated the correct key and the integrity test would have failed.

4. If the integrity test is successful then Bob decrypts $\mathbf{m}$=$\mathtt{Ds}(\mathbf{k}_e, \mathbf{c})$.

The following, will focus on multicarrier systems in which keys are generated at the physical layer and used in authentication/encryption protocols at upper layers as described above.

4.5 Pipelined SKG and encrypted data transfer

As explained in Section 4.3, if Alice and Bob follow the standard sequential SKG process they can exchange encrypted data only after both of them have distilled the key at the end of the privacy amplification step. This section proposes a method to pipeline the SKG and encrypted data transfer. Alice can unilaterally extract the secret key from her observation and use it to encrypt data transmitted in the same "extended" message that contains the side information (see Fig. 4.3). Subsequently, using the side information, Bob can distill the same key $\mathbf{k}$ and decrypt the received data in one single step.

Section 2.2.1 discussed how Alice and Bob can distill secret keys from estimates of the fading coefficients in their wireless link and Section 4.4 showed how these can be used to develop an AE SKG primitive. At the same time CSI estimates are a prerequisite in order to optimally allocate power across the subcarriers and achieve high data rates[1]. As a result, a question that naturally arises is whether the CSI estimates (obtained at the end of the pilot exchange phase), should be used towards the generation of secret keys or towards the reliable data transfer, and, furthermore, whether the SKG and the data transfer can be inter-woven using the AE SKG principle.

This study is interested in answering this question and shed light into whether following the exchange of pilots Alice should transmit reconciliation information on all subcarriers, so that she and Bob can generate (potentially) a long sequence of key bits, or, alternatively, perform information reconciliation only over a subset of the subcarriers and transmit encrypted data over the rest, exploiting the idea of the AE SKG primitive. Note here that the data can be already encrypted with the key generated at Alice, the sender of the side information, so that the proposed pipelining does not require storing keys for future use. The former approach is denoted as a *sequential* scheme, while the latter is referred to as a *parallel* scheme. The two will be compared in terms of their efficiency with respect to the achievable data rates.

As discussed in Section 4.3, the assumed physical layer system model consists of Alice and Bob who exchange data over a Rayleigh BF-AWGN channel with N orthogonal subcarriers. Without loss of generality the variance of the AWGN in all links is assumed to be unity. During channel probing, constant pilots are sent across all subcarriers [37, 74] with power P. Using the observations (4.21), Alice estimates the channel coefficients as

$$\hat{h}_j = h_j + \tilde{h}_j, \tag{4.24}$$

for $j = 1, \ldots, N$ where $\tilde{h}_j$ denotes an estimation error that can be assumed to be Gaus-

[1] As an example, despite the extra overhead in URLLC systems, advanced CSI estimation techniques are employed in order to be able to satisfy the strict reliability requirements [246, 247].

sian, $\tilde{h}_j \sim \mathcal{CN}(0, \sigma_e^2)$ [248]. Under this model, the following rate is achievable on the j-th subcarrier from Alice to Bob when the transmit power during data transmission is p_j [248]:

$$R_j = \log_2\left(1 + \frac{g_j p_j}{\sigma_e^2 P + 1}\right) = \log_2(1 + \hat{g}_j p_j), \tag{4.25}$$

where $g_j = |h_j|^2$ and $\hat{g}_i = \frac{g_j}{\sigma_{j,e}^2 P + 1}$, to denote the estimated channel gains. The metric, R_j, measures the average rate on the subcarriers devoted to data transmission assuming Shannon's capacity is achievable. This measure is the maximal achievable rate for a specified maximal error probability while assuming infinite code-word length [249, 250]. This is a very optimistic assumption since no code-word can be designed to ensure that. An alternative measures are the outage capacity and the rate for finite block lengths. The former is defined as the achievable rate of transmission under a given error probability [251,252], and the latter defines the achievable rates taking into account factors such as code-word size, probability of error and channel dispersion [253, 254]. Overall, outage capacity and achievable rate for finite block lengths gives a more accurate solution, and therefore, must be considered in practical implementations. However, for the purpose of the present thesis Shannon's capacity is used as an indicative metric. Although this is an optimistic measure for the achievable rates, it reflects the overall performance based on different system parameters. As a result, the channel capacity is defined as $C = \sum_{j=1}^{N} R_j$ under the short term power constraint:

$$\sum_{j=1}^{N} p_j \leq NP, \;\; p_j \geq 0, \; \forall j \in \{1, \ldots, N\}, \tag{4.26}$$

is achieved with the well known waterfilling power allocation policy $p_j = \left[\frac{1}{\lambda} - \frac{1}{\hat{g}_j}\right]^+$, where the water-level λ is estimated from the constraint (4.26). In the following, the estimated channel gains $\hat{g}_j$ are – without loss of generality – assumed ordered in descending order, so that:

$$\hat{g}_1 \geq \hat{g}_2 \geq \ldots \geq \hat{g}_N. \tag{4.27}$$

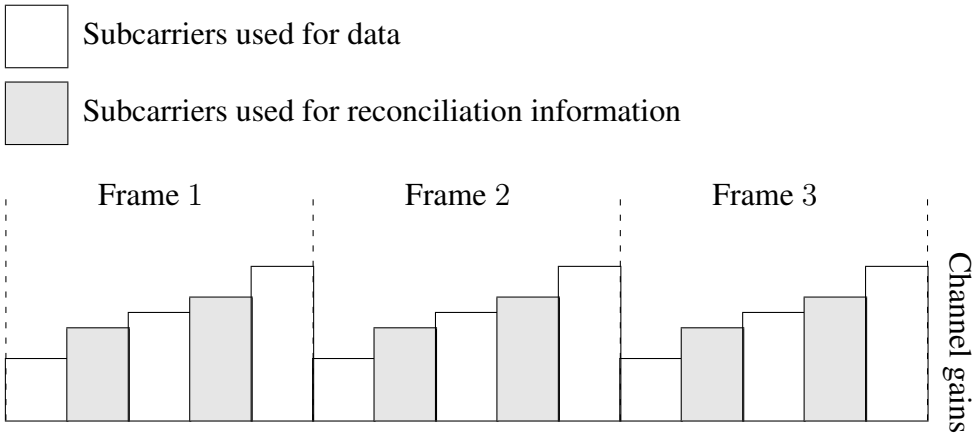

Figure 4.4: Parallel approach

As mentioned in Section 2.2.1, the advantage distillation phase of the SKG process consists of the two-way exchange of pilot signals during the coherence time of the channel. On the other hand, the CSI estimation phase can be used to estimate the reciprocal channel gains in order to optimise data transmission using the waterfilling algorithm. In the former case, the shared parameter is used for generating symmetric keys, in the latter for deriving the optimal power allocation. In the parallel approach the idea is to interweave the two procedures and investigate whether a joint encrypted data transfer and key generation scheme as in the AE SKG in Section 4.4 could bear any advantages with respect to the system efficiency. While in the sequential approach the CSI across all subcarriers will be treated as a source of shared randomness between Alice and Bob, in the parallel approach it plays a dual role.

4.5.1 Parallel approach

In the parallel approach, after the channel estimation phase, the legitimate users decide on which subcarrier to send the reconciliation information (*e.g.*, the syndromes) and on which data (*i.e.*, the SKG process here is not performed on all of the subcarriers). This approach is illustrated in Fig. 4.4. The total capacity has now to be distributed between data and reconciliation information bearing subcarriers. As a result, the overall set of orthogonal subcarriers comprises two subsets; a subset $\mathcal{D}$ that is used for encrypted data

transmission with cardinality $|\mathcal{D}| = D$ and a subset $\breve{\mathcal{D}}$ with cardinality $|\breve{\mathcal{D}}| = N - D$ used for reconciliation such that, $\mathcal{D} \cup \breve{\mathcal{D}} = \{1, \ldots, N\}$. Over $\mathcal{D}$ the achievable sum data transfer rate, denoted by C_D is given by

$$C_D = \sum_{j \in \mathcal{D}} \log_2(1 + \hat{g}_j p_j), \tag{4.28}$$

while on the subset $\breve{\mathcal{D}}$, Alice and Bob exchange reconciliation information. As stated in Section 4.3 the fading coefficients are assumed to be zero-mean circularly-symmetric complex Gaussian random variables.

Following from the above, the SKG rate can be expressed as [37, 74]:

$$C_{SKG} = \sum_{j \in \breve{\mathcal{D}}} R_{SKG,j} = \sum_{j \in \breve{\mathcal{D}}} \log_2 \left(1 + \frac{P_j \sigma^2}{2 + \frac{1}{P_j \sigma^2}} \right). \tag{4.29}$$

Finally, the efficiency of the proposed parallel method – measured as the ratio of the long-term data rate versus the average capacity – is evaluated as:

$$\eta_{\text{parallel}} = \frac{\mathbb{E}\left[C_D\right]}{\mathbb{E}[C]}. \tag{4.30}$$

This efficiency quantifies the expected back-off in terms of data rates when part of the resources (power and frequency) are used to enable the generation of secret keys at the physical layer.

4.5.2 Sequential approach

In the sequential approach encrypted data transfer and secret key generation are two separate events; first, the secret keys are generated over the whole set of subcarriers, leading to a sum SKG rate given as

$$C_{SKG} = N R_{SKG} = N \log_2 \left(1 + \frac{P \sigma^2}{2 + \frac{1}{P \sigma^2}} \right). \tag{4.31}$$

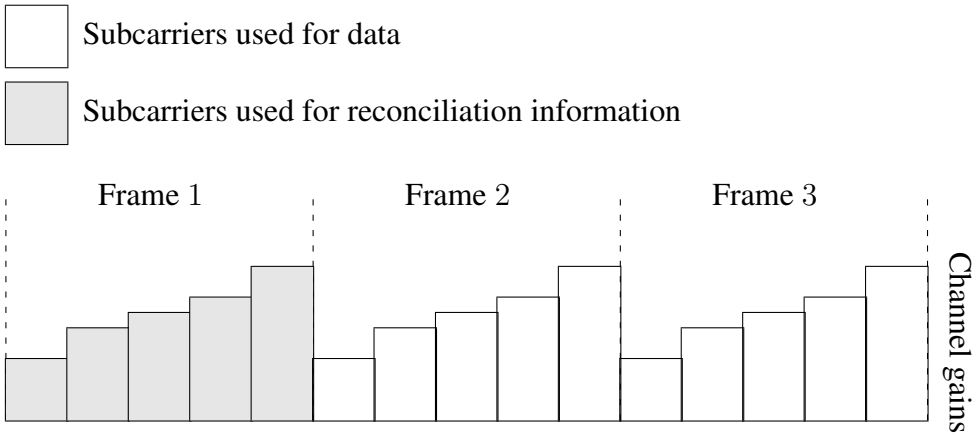

Figure 4.5: Sequential approach

This is illustrated in Fig. 4.5. To estimate the efficiency of the scheme, the necessary resources for the exchange of the reconciliation information need to be identified. An estimate of the number of transmission frames that will be required for the transmission of the syndromes can be obtained as the expected value of the reconciliation rate (*i.e.*, it's long-term value). For example, in Fig. 4.5 this is only frame 1, however, depending on the scenario more frames might be needed to establish a long enough key. Therefore, the the efficiency of the sequential method is then calculated as:

$$\eta_{\text{sequential}} = \frac{L}{L+M}, \tag{4.32}$$

where M denotes the average number of frames needed for reconciliation and L denotes the average number of frames with encrypted data before a new key establishment needs to take place. Both of this parameters will be precisely defined later in this chapter.

4.6 Optimal power and subcarrier allocation

This section will identify the optimal power and subcarrier allocation for Alice and Bob under three different scenarios. The study begins with a simplified version of the optimisation problem. This initial version assumes power and security constraints, where the reconciliation rate is roughly approximated to the SKG rate [13]. After solving this sim-

plified problem, a more realistic scenario was investigated where the rate of transmitting reconciliation information is explicitly accounted and differentiated from the SKG rate. Finally, this study was extended by taking into account delay requirements. In detail, it investigates the optimal resource allocation for Alice and Bob, when their communication has to satisfy specific delay constraints.

4.6.1 Optimal allocation under security and power constraints

The study within this section evaluates only the performance of the parallel approach. Once having a good understanding of the problem, the parallel approach will be compared to the sequential in the next sections.

As discussed in Section 4.5.1 the achievable sum rates for data transfer and SKG are denoted by C_D and C_{SKG}, respectively, where data is transferred over the subset of subcarriers $\mathcal{D}$ and keys are generated over the subset $\breve{\mathcal{D}}$. Depending on the exact choices of the cryptographic suites to be employed, it is possible to reuse the same key for the encryption of multiple blocks of data, *e.g.*, as in the AES Galois/Counter Mode, that is being considered for employment in the security protocols for URLLC systems [21]. In practical systems, a single key of length 128 to 256 bits can be used to encrypt up to gigabytes of data. As a result, a realistic assumption is that for a particular application it is possible to identify the ratio of key to data bits, which in the following we will denote by β. As noted, in practical systems, this ratio might reach values in the order of 10^{-5}, i.e., a single key of length of 256 bits could be used to encrypt large amount of data. On the other hand, some applications which require higher level of security will apply a higher value for β, *i.e.*, secret keys will be updated more often. In fact, the case $\beta = 1$ would correspond to a one-time-pad, *i.e.*, the generated keys could be simply x-ored with the data to achieve perfect secrecy without the need of any cryptographic suites. Therefore, the following security constraint should be met

$$C_{SKG} \geq \beta C_D, \quad 0 < \beta \leq 1, \tag{4.33}$$

where, depending on the application, the necessary minimum value of β can be identified.

Following from the above, the initial problem discussed in this chapter finds the optimal trade-off between data transmission and SKG in terms of subcarrier allocation and power allocation, under the short-term power constraint (4.26). To this end, the following optimisation problem is formulated:

$$\max_{p_i, i\in\mathcal{D}} \sum_{i\in\mathcal{D}} \log_2(1+\hat{g}_i p_i), \quad \text{s.t. (4.26) and (4.33).} \tag{4.34}$$

To solve the problem, the optimisation methods described in Section 4.2.1 are used. In fact i) the optimal power allocation is calculated through convex optimisation, based on the Lagrangian approach; ii) the optimal subcarrier allocation is identified numerically through MATLAB simulation using a dynamic programming approach, in the form of Knapsack optimisation. Further details regarding the formulation of these problems are given later in this chapter. First, the following two Lemmas are stated.

Lemma 4.1. *In order to maximise C_D the strongest D subcarriers – in terms of SNR – are used for data transmission and the rest, $N-D$, for SKG.*

Proof. Assume that $\mathcal{D}^*$ is the subset of subcarriers indices which maximises C_D and $\mathcal{D}_{ord} = \{1, 2, \ldots, D\}$ the subset of the first D ordered subcarrier indices. Then, after fixing a subcarrier power level $p_d > 0$, $\forall\ d \in \mathcal{D}^*$ with $d \notin \mathcal{D}_{ord}$ it follows that a better index exists, *i.e.*, $\exists\ d' \in \mathcal{D}_{ord}$ with $d' \notin \mathcal{D}^*$, s.t.,

$$\log_2(1+\hat{g}_d p_d) < \log_2(1+\hat{g}_{d'} p_d). \tag{4.35}$$

As a consequence of Bellman's principle [255] the optimal sum rate in (4.34) has to consist of optimal subcarrier rates, (4.35) contradicts this fact and hence, $\mathcal{D}^* = \mathcal{D}_{ord} = \{1, 2, \ldots, D\}$. ■

Consequently, the following fact is used in later derivations

$$\mathcal{D}^* = \{1, 2, \ldots, D\}. \tag{4.36}$$

The optimality of this approach, *i.e.,* choosing the worst subcarriers for SKG and best for data transmission is confirmed through MATLAB simulation using the dynamic programming approach, described in Sec. 4.2.1 in the form of 0-1 Knapsack optimisation. The problem is formulated as in Eq. (4.12) and (4.13) where $v_i = w_i$ denote the achievable rate of the i-th subcarrier and W_i is chosen such that constraint (4.33) is satisfied. The problem is solved using the iterative approach given in Eq. 4.15. Numerical results on the achievable versus different values of β and different subcarrier allocation are presented later in this chapter.

The next question is: how the available power for SKG should be used? We first assume that the overall power expended for SKG can be expressed as

$$P_s = (N - D)p_s, \tag{4.37}$$

where p_s denotes the average SKG power and P_s the overall SKG power. Given (4.37), and by taking the first and the second derivative of $R_{SKG}(p_s)$, it is straightforward to see it is a monotonic function in p_s and it is convex if $p_s < \frac{1}{\sqrt{2}\sigma^2}$ and concave if $p_s > \frac{1}{\sqrt{2}\sigma^2}$.

Lemma 4.2. *If $p_s > \frac{1}{\sqrt{2}\sigma^2}$ the set $\breve{\mathcal{D}}$ comprises the weakest $N - D$ subcarriers – in terms of SNR – and the power allocation is equal on all of them, so that:*

$$C_{SKG} = (N - D)\log_2\left(1 + \frac{p_s\sigma^2}{2 + \frac{1}{p_s\sigma^2}}\right). \tag{4.38}$$

Consequently the overall power allocation vector takes the form:

$$\mathbf{p} = \{p_1, p_2, \ldots, p_D, p_s, p_s, \ldots, p_s\}, \tag{4.39}$$

where the number of elements equal to p_s is $N - D$.

If $p_s < \frac{1}{\sqrt{2}\sigma^2}$ the set $\breve{\mathcal{D}}$ consists of a single subcarrier on which the full power available

for SKG is allocated, so that:

$$C_{SKG} = \log_2\left(1 + \frac{P_s\sigma^2}{2 + \frac{1}{P_s\sigma^2}}\right). \tag{4.40}$$

Proof. Note that when multiple subcarriers are to be used the power should be equally distributed to them. To prove this the definition of a concave function is applied together with Jensen's inequality [256], [257]:

$$R_{SKG}\left(\sum_{i=1}^{N-D} \delta_i x_i\right) > \sum_{i=1}^{N-D} \delta_i R_{SKG}(x_i). \tag{4.41}$$

Substituting $\delta_i = 1/(N-D)$ and $x_i = P_s/b_i$ with $\sum_{i=1}^{N-D} \delta_i = 1$, results in:

$$\begin{aligned} &R_{SKG}\left(\sum_{i=1}^{N-D} \frac{P_s}{(N-D)b_i}\right) > \sum_{i=1}^{N-D} \frac{1}{N-D} R_{SKG}\left(\frac{P_s}{b_i}\right) \Leftrightarrow \\ &(N-D)R_{SKG}\left(\frac{1}{N-D}\sum_{i=1}^{N-D} \frac{P_s}{b_i}\right) > \sum_{i=1}^{N-D} R_{SKG}\left(\frac{P_s}{b_i}\right) \end{aligned} \tag{4.42}$$

From the RHS of (4.42) it can be seen that the power allocation on each subcarrier is P_s/b_i, therefore the the following power constraint $\sum_{i=1}^{N-D} P_s/b_i \leq P_s$ is added. The fact, R_{SKG} is monotonically increasing function with P_s, implies:

$$\begin{aligned} &(N-D)R_{SKG}\left(\frac{P_s}{N-D}\right) \geq (N-D)R_{SKG}\left(\frac{1}{N-D}\sum_{i=1}^{N-D} \frac{P_s}{b_i}\right) \Leftrightarrow \\ &(N-D)R_{SKG}\left(\frac{P_s}{N-D}\right) > \sum_{i=1}^{N-D} R_{SKG}\left(\frac{P_s}{b_i}\right). \end{aligned} \tag{4.43}$$

Equation (4.43) proves that in order to maximise the sum rate R_{SKG} the legitimate users have to distribute their power equally when multiple subcarriers are used.

Next, it is proved that all the subcarriers have to be used. Recalling the definition of a concave function for a single δ on the interval $[0, b]$ results in:

$$R_{SKG}((1-\delta)0 + \delta b) > (1-\delta)R_{SKG}(0) + \delta R_{SKG}(b), \tag{4.44}$$

with $\delta = a/b$ for $f(0) > 0,$ and $0 < a < b$, implies:

$$\begin{aligned} &R_{SKG}\left(\left(1-\frac{a}{b}\right)0+\frac{a}{b}b\right) > \left(1-\frac{a}{b}\right)R_{SKG}(0)+\frac{a}{b}R_{SKG}(b) \Leftrightarrow \\ &R_{SKG}(a) > R_{SKG}(0)\frac{b-a}{b}+\frac{a}{b}R_{SKG}(b) \geq \frac{a}{b}R_{SKG}(b) \Leftrightarrow \\ &\frac{R_{SKG}(a)}{a} > \frac{R_{SKG}(b)}{b}. \end{aligned} \tag{4.45}$$

Setting $a = x/v$ and $b = x/u$ gives:

$$uf\left(\frac{x}{u}\right) < vf\left(\frac{x}{v}\right) \tag{4.46}$$

for $0 < u < v$ and $x > 0$. Given that when $p_s > \frac{1}{\sqrt{2\sigma^2}}$, *i.e.*, when $R_{SKG}(p_s)$ is concave, results in:

$$\begin{aligned} R_{SKG}\left((N-D)p_s\right) &< 2R_{SKG}\left(\frac{N-D}{2}p_s\right) < \dots \\ &\dots < (N-D-1)R_{SKG}\left(\frac{N-D}{N-D-1}p_s\right) < (N-D)R_{SKG}(p_s), \end{aligned} \tag{4.47}$$

which shows that all available subcarriers have to be used.

On the other hand, when $p_s < \frac{1}{\sqrt{2\sigma^2}}$, *i.e.*, $R_{SKG}(p_s)$ is convex and by the definition of a convex function results in:

$$R_{SKG}((1-\delta)0+\delta b) < (1-\delta)R_{SKG}(0)+\delta R_{SKG}(b), \tag{4.48}$$

which is equivalent to:

$$\begin{aligned} R_{SKG}\left((N-D)p_s\right) &> 2R_{SKG}\left(\frac{N-D}{2}p_s\right) > \dots \\ &\dots > (N-D-1)R_{SKG}\left(\frac{N-D}{N-D-1}p_s\right) > (N-D)R_{SKG}(p_s), \end{aligned} \tag{4.49}$$

which shows that in this case it is optimal to use a single subcarrier for SKG. Lemma 4.2 follows. ■

As a result of Lemma 4.2, the following two cases, Case 1 for R_{SKG} concave and Case 2 for R_{SKG} convex are explored.

Case 1: $p_s > \frac{1}{\sqrt{2}\sigma^2}$

Theorem 4.1. *When $p_s > \frac{1}{\sqrt{2}\sigma^2}$ the optimal power allocation for data transmission and SKG on each subcarrier are:*

$$p_i^* = \left[\frac{1-\beta\mu}{\lambda \ln(2)} - \frac{1}{\hat{g}_i}\right]^+, i = 1, \ldots, D \tag{4.50}$$

$$p_s^* = \left[\frac{Q \pm \sqrt{Q^2 - F^2}}{\frac{4\sigma^2}{\sqrt{8}}F}\right]^+, \tag{4.51}$$

where:

$$[x]^+ \triangleq \max(x, 0), \tag{4.52}$$

$$Q = 2\sigma^4\mu N - 2\sigma^4\mu D - 3\sigma^2\lambda \ln(2), \tag{4.53}$$

$$F = \sqrt{8}\lambda\sigma^2 \ln(2). \tag{4.54}$$

For the feasibility of (4.50) and (4.51) we have:

$$\mu < \frac{1}{\beta}, \tag{4.55}$$

$$\mu \geq \frac{\lambda \ln(2)(3+\sqrt{8})}{2\sigma^2(N-D)}, \tag{4.56}$$

$$\lambda > 0. \tag{4.57}$$

Proof. Following the results drawn in Lemma 4.2, if $p_s > \frac{1}{\sqrt{2}\sigma^2}$ (*i.e.*, the function R_{SKG} is concave and the size of the set $|\breve{\mathcal{D}}| = (N - D)$) the optimisation problem (4.34) can be re-written as:

$$\max_{p_i} \sum_{i=1}^{D} \log_2(1 + \hat{g}_i p_i) \tag{4.58}$$

s.t. (4.26) and

$$\beta\left(\sum_{i=1}^{D}\log_2(1+\hat{g}_i p_i)\right) = (N-D)\log_2\left(1+\frac{p_s\sigma^2}{2+\frac{1}{p_s\sigma^2}}\right). \tag{4.59}$$

The optimisation problem (4.58) is solved using the Lagrangian method described in Sec. 4.2.1. Note that the Karush-Kuhn-Tucker (KKT) conditions are satisfied, only if Eq. (4.55)-(4.57) are satisfied. Next, to obtain the optimal power allocation the Lagrangian is formulated, as:

$$\begin{aligned}\mathcal{L}_p = &\sum_{i=1}^{D}\log_2(1+\hat{g}_i p_i) - \lambda\left(\sum_{i=1}^{N} p_i + NP\right) - \\ &\mu\left(\beta\left(\sum_{i=1}^{D}\log_2(1+\hat{g}_i p_i)\right) - (N-D)\log_2\left(1+\frac{\sigma^2 p_s}{2+\frac{1}{\sigma^2 p_s}}\right)\right)\end{aligned} \tag{4.60}$$

which after a few algebraic manipulations takes the form:

$$\begin{aligned}\mathcal{L}_p = &\left(\sum_{i=1}^{D}\log_2(1+\hat{g}_i p_i)\right)(1-\mu\beta) - \lambda\left(\sum_{i=1}^{N} p_i + NP\right) \\ &+ \mu\left((N-D)\log_2\left(1+\frac{\sigma^2 p_s}{2+\frac{1}{\sigma^2 p_s}}\right)\right),\end{aligned} \tag{4.61}$$

where λ and μ are the dual Lagrange multipliers, that correspond to (4.26) and (4.59), respectively. Solving the problem based on the stationary condition given in (4.7), *i.e.*, setting the gradient of Eq. (4.61) to zero, and solving for p gives the optimal power allocation. In fact, the problem (4.58) has concave objective and constraint functions, and as a result the optimal power allocation is given in (4.50) and (4.51) is the unique solution. ■

Case 2: $p_s < \frac{1}{\sqrt{2}\sigma^2}$

Theorem 4.2. *When $p_s < \frac{1}{\sqrt{2}\sigma^2}$ a single random subcarrier with index $j > D$ can be used for SKG and without loss of generality we can set $j = D+1$. The optimal power*

allocation for data transmission and SKG is then expressed as:

$$p_i^* = \left[\frac{1-\beta\mu}{\lambda\ln(2)} - \frac{1}{\hat{g}_i}\right]^+, i = 1, \ldots, D, \tag{4.62}$$

$$P_s^* = \left[\frac{Q' \pm \sqrt{Q'^2 - F'^2}}{\frac{4\sigma^2}{\sqrt{8}}F'}\right]^+, \tag{4.63}$$

where:

$$Q' = 2\mu\sigma^4 - 3\sigma^2\lambda\ln(2) \tag{4.64}$$

$$F' = \sqrt{8}\lambda\sigma^2\ln(2). \tag{4.65}$$

For the feasibility of (4.62) and (4.63) we have:

$$\mu < \frac{1}{\beta}, \tag{4.66}$$

$$\mu \geq \frac{\lambda\ln(2)(3+\sqrt{8})}{2\sigma^2}, \tag{4.67}$$

$$\lambda > 0. \tag{4.68}$$

Proof. Similarly to Theorem 4.1 the results drawn in Lemma 4.2 are used. If $p_s < \frac{1}{\sqrt{2}\sigma^2}$ (*i.e.*, the function R_{SKG} is convex and the size of the set $|\breve{\mathcal{D}}| = 1$), therefore, Eq. (4.33) takes the form:

$$\beta\left(\sum_{i=1}^{D}\log_2(1+\hat{g}_i p_i)\right) = \log_2\left(1 + \frac{P_s\sigma^2}{2+\frac{1}{P_s\sigma^2}}\right). \tag{4.69}$$

Following the approach in Theorem 4.1 and by using the KKT conditions it is straightforward to see that the optimal power allocation for each subcarrier is presented in (4.62) and (4.63). ■

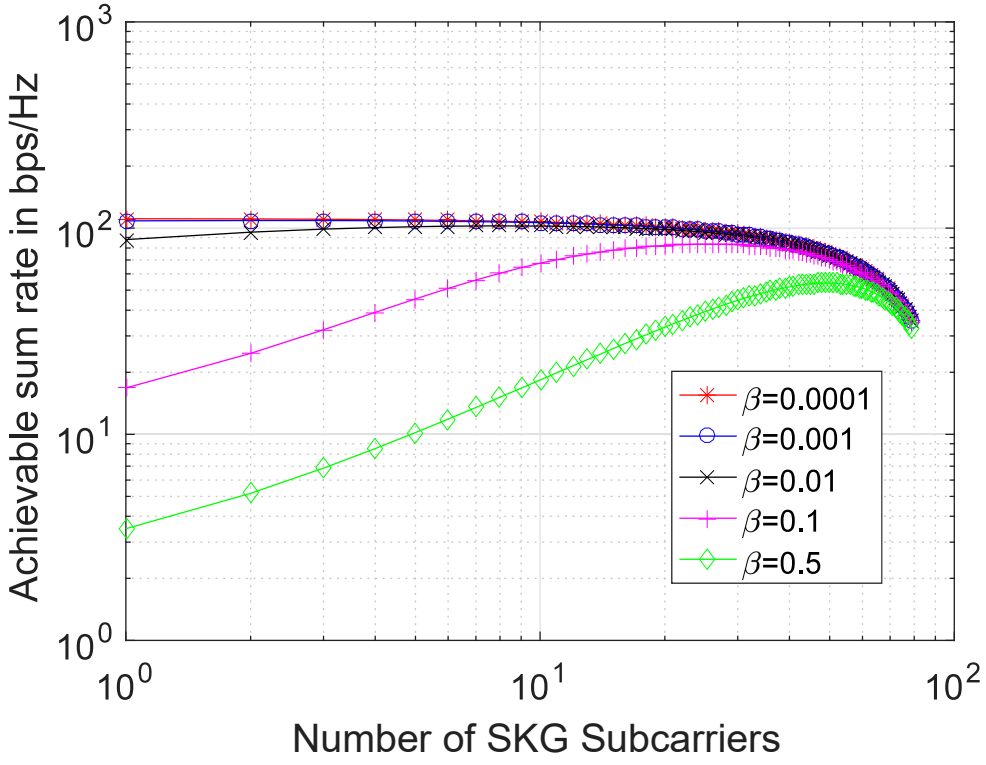

Figure 4.6: Case 1: Achievable sum rate C_D averaged over $10,000$ simulations for different values of β, defined in (4.33), number of subcarriers used for data transmission and for SKG. Parameters: $(p_1 + \cdots + p_D)/D = 5, \sigma^2 = 1, N = 100$.

Numerical evaluation

By applying (4.28) and using the modified water-filling solution of the maximisation problem given in (4.34) the achievable data rates can be simulated (See Fig. 4.6 and Fig. 4.7).

In Fig. 4.6 the dependence of the achievable sum rate on β and D easily can be seen. While varying β the achievable C_D changes and due to its concavity the unique maximum, achieved with the optimal D and power allocation, can always be identified. For small values of β it can be seen that the fewer SKG subcarriers that are used the greater sum rate can be achieved. When β increases the optimal subcarrier allocation changes, and a larger number of subcarriers are needed to meet the security constraint.

As expected, in Case 2, varying β directly affects the achievable sum rate. In agreement with intuition, from Fig. 4.7 we see that the smaller the β, the greater rate we

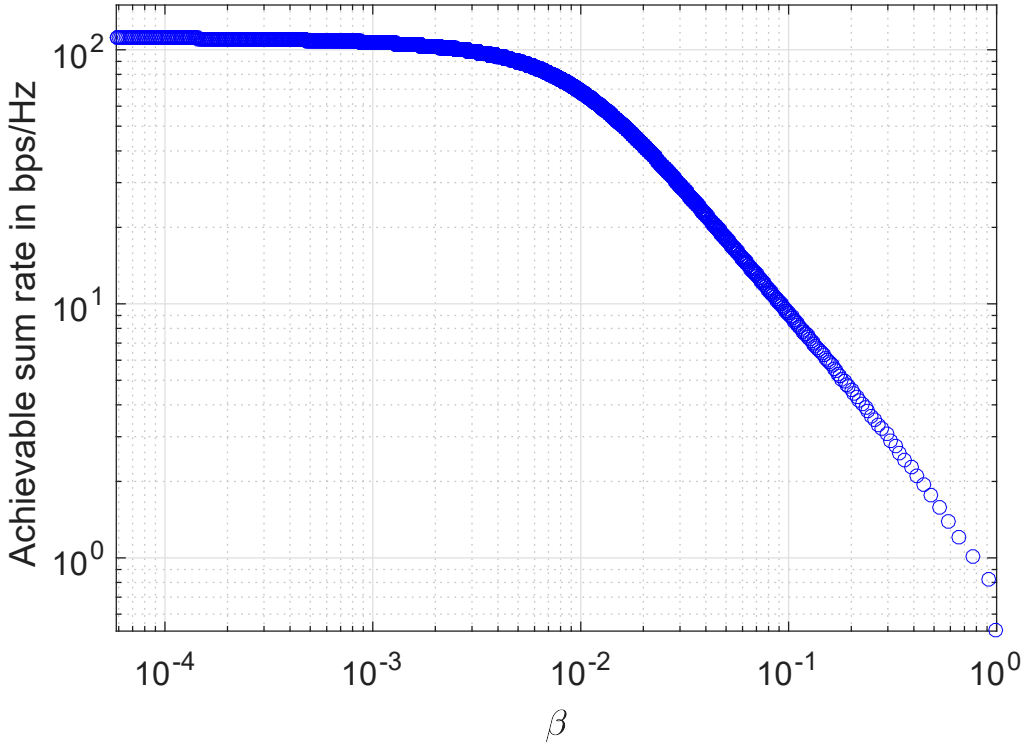

Figure 4.7: Case 2: Achievable sum rate C_D averaged over $10,000$ simulations for different values of β, defined in (4.33). Parameters: $N = 100, D = 99, (p_1 + \cdots + p_D)/D = 5, \sigma^2 = 1$.

achieve.

Long-term Issues with the Short-term Policy

Characterising the distribution of the fading coefficients is important to understand the channel properties. Having a short-term power constraint allows to optimally allocate the subcarriers and the available power. However, the optimal short-term solution suggests that the weakest subcarriers should be used for SKG. If this policy is applied in the long-term, it is obvious that it will have an impact on the actual statistical properties and the distribution of the coefficients used for SKG. This effect is investigated in the present section.

As discussed in Section 4.3, the fading coefficients are assumed to be zero-mean circu-

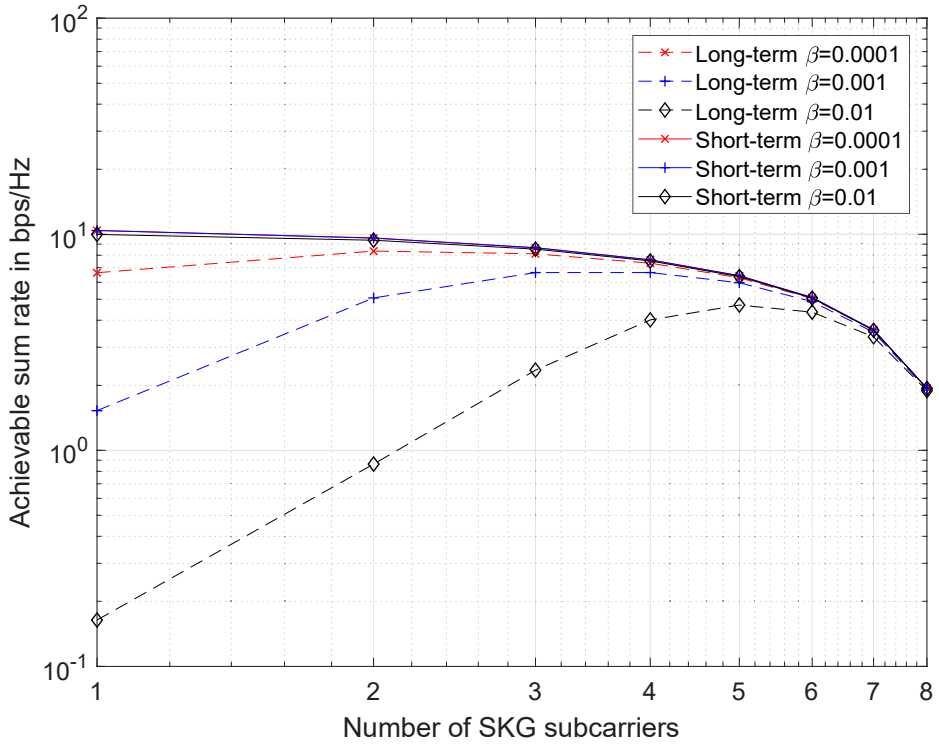

Figure 4.8: Case 1: Achievable sum rate C_D for different values of β, defined in (4.33), number of subcarriers used for data transmission and for SKG. Parameters: $(p_1 + \cdots + p_D)/D = 5, \sigma^2 = 1, N = 10$.

larly symmetric complex Gaussian random variables. It has been shown that the weakest $N - D$ subcarriers should be used for SKG, where D depends on the system parameters as well as the exact fading realisations. Therefore, the distribution of the channel gains of the SKG subcarriers for $j = \{D + 1, \ldots, N\}$ can then be expressed as [258]:

$$p(g_j) = \frac{N!}{\sigma^2(N-j)!(j-1)!}\left(1 - e^{-\frac{g_j}{\sigma^2}}\right)^{N-j}\left(e^{-\frac{g_j}{\sigma^2}}\right)^{i} \tag{4.70}$$

where $\sigma^2 = 4\sigma_h^4$ is the variance of channel gains. As a result, choosing the weakest $N-D$ subcarriers for SKG in the long-term will impact the variance of each of them, which is now given by:

$$\sigma_j^2 = \sigma^2 \sum_{q=j}^{N} \frac{1}{q^2}, \quad j \in \{1, \ldots, N\}. \tag{4.71}$$

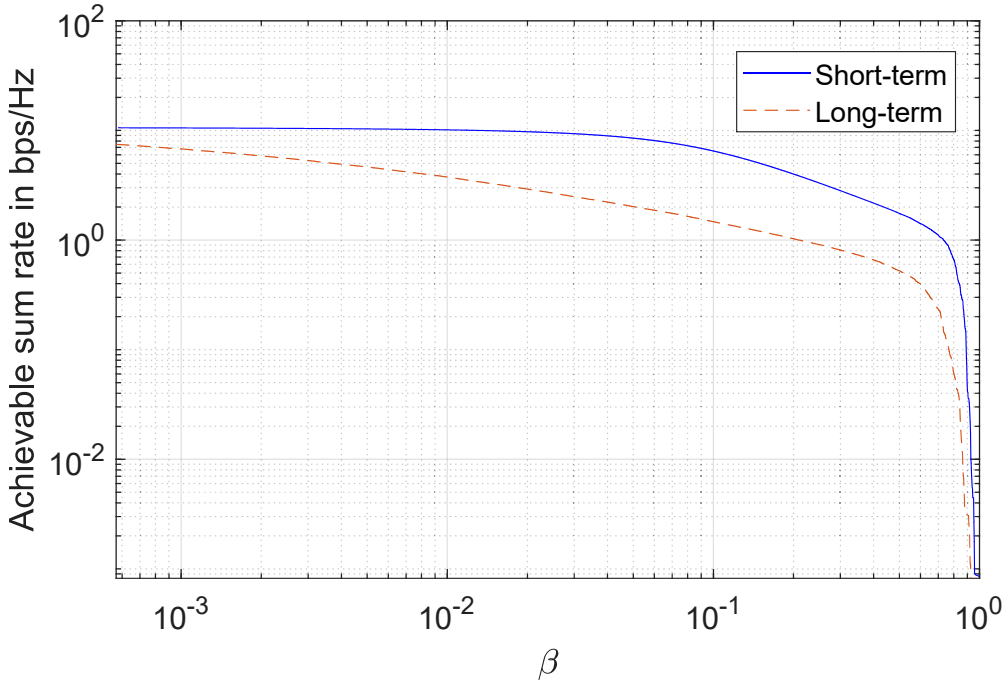

Figure 4.9: Case 2: Achievable sum rate C_D for different values of β, defined in (4.33). Parameters: $N = 10, D = 9, (p_1 + \cdots + p_D)/D = 5, \sigma^2 = 1$.

As an example, the variance of the weakest of N subcarriers is scaled with a factor N^{-2}. This effect has high impact on the SKG rate which is now given as:

$$C_{SKG} = \sum_{j \in \breve{\mathcal{D}}} R_{SKG,j} = \sum_{j \in \breve{\mathcal{D}}} \log_2 \left(1 + \frac{P\sigma_j^2}{2 + \frac{1}{P\sigma_j^2}} \right). \tag{4.72}$$

where σ_j is given in Eq. 4.71. This effect is specifically accounted in the next figures.

Fig. 4.8 and 4.9 compare the rates that are achievable for Case 1 and 2, respectively, in the short-term and in the long-term with the proposed short-term policy. It can be seen that for small values of β the incurred penalty for both cases is small. However, when β increases a higher reduction of the sum rates is observed.

To summarise this work investigated the possibility of jointly performing data transfer and SKG in a Rayleigh BF-AWGN environment. This initial simplified approach

investigated the maximisation of the data transfer rate under two constraints: a security constraint, and, a short-term overall power constrain, where the information reconciliation rate was roughly approximated to the SKG rate. The analysis demonstrated that in this scenario the strongest subcarriers – from an SNR point of view – should be allocated to data transfer and the weakest to SKG. Accordingly, the optimal power allocation for the data transfer has been shown to be expressed in the waterfilling form while the power allocation for the SKG subcarriers depends on the overall available power and might not be unique. Furthermore, the impact of utilising the optimal short-term policy in the long-term was investigated. The use of order statistics revealed that systematically choosing the weakest subcarriers for SKG can result, in the worst case, in a scaling inversely proportional to the square of N, the number of subcarriers, for the SKG variance. However, for small values of N and β the incurred penalty is small.

Based on the results presented within this section, the next sections will account for the exact rate of transmitting reconciliation information. Furthermore, it will be confirmed whether the policy of using the strongest subcarriers for data transmission is still optimal when the full optimisation problem is considered, including the communication cost for reconciliation.

4.6.2 Optimal allocation under security, power and rate constraints

As discussed in Section 4.5.2 in the parallel approach, the legitimate users decide on which subcarrier to send the reconciliation information and on which data – the total capacity has now to be distributed between data and reconciliation information bearing subcarriers. In the previous section the reconciliation rate was roughly approximated to the SKG rate. Therefore, this section accounts for the exact rate on the subset $|\mathcal{D}| = D$ that is used for encrypted data transmission and subset $|\breve{\mathcal{D}}| = N - D$ used for reconciliation. Over $\mathcal{D}$ the achievable sum data transfer rate, denoted by C_D is given by

$$C_D = \sum_{j \in \mathcal{D}} \log_2(1 + \hat{g}_j p_j), \tag{4.73}$$

while on the subset $\breve{\mathcal{D}}$, Alice and Bob exchange reconciliation information at rate

$$C_R = \sum_{j \in \breve{\mathcal{D}}} \log_2(1 + \hat{g}_j p_j). \tag{4.74}$$

As discussed in Section 4.3, the minimum rate necessary for reconciliation has been theoretically derived in [36]. Here, alternatively, a practical design approach is employed in which the rate of the encoder used is explicitly taken into account. Noting that in a rate $\frac{k}{n}$ block encoder the side information is $n-k$ bits long, *i.e.*, the rate of syndrome to output key bits after privacy amplification is $\frac{n-k}{k}$. Therefore, a parameter $\kappa = \frac{n-k}{k}$ – the inverse of the encoder rate, is defined here. The parameter reflects the ratio of the reconciliation rate to the SKG rate, for example, for a rate $\frac{k}{n} = \frac{1}{2}$ encoder, $\kappa = 1$, for $\frac{k}{n} = \frac{1}{3}$, $\kappa = 2$, while for $\frac{k}{n} = \frac{1}{4}$, $\kappa = 3$. Note, in practice κ needs to be chosen depending on the scenario and the channel characteristics. Based on this discussion, the minimum requirement for the reconciliation rate is captured through the following expression:

$$C_R \;\; \geq \;\; \kappa C_{SKG}, \tag{4.75}$$

where the achievable SKG rate C_{SKG} is given in Eq. 4.72.

Accounting for the reconciliation rate and security constraints in (4.75) and (4.33) the following maximisation problem is formulated:

$$\max_{p_j, j \in \mathcal{D}} \quad \sum_{j \in \mathcal{D}} R_j \tag{4.76}$$

s.t.

$$\sum_{j=1}^{N} p_j \leq NP \tag{4.77}$$

$$C_R \geq \kappa C_{SKG}, \tag{4.78}$$

$$C_{SKG} \geq \beta C_D, \quad 0 < \beta \leq 1, \tag{4.79}$$

$$\sum_{j \in \mathcal{D}} R_j + \sum_{j \in \breve{\mathcal{D}}} R_j \leq C. \tag{4.80}$$

Constraint (4.79) can be integrated with (4.78) to the combined constraint

$$\sum_{j\in\mathcal{D}} R_j \quad \leq \quad \frac{\sum\limits_{j\in\breve{\mathcal{D}}} R_j}{\kappa\beta}. \tag{4.81}$$

The optimisation problem at hand is a mixed-integer convex optimisation problem with unknown entities being both the sets $\mathcal{D}, \breve{\mathcal{D}}$, as well as the power allocation policy $p_j, j \in \{1, \ldots, N\}$. These problems are typically NP hard and addressed with the use of branch and bound algorithms and heuristics.

Based on the result from the previous section, this work proposes a simple heuristic to make the problem more tractable by reducing the number of free variables. In the proposed approach, it is assumed that the constraint (4.80) is satisfied with equality. The only power allocation that allows this is the water-filling approach that uniquely determines the power allocation p_j and also requires that the constraint (4.77) is also satisfied with equality. Thus, following that approach, the power allocation vector can be uniquely determined and can be used to combine the remanining constraints (4.80) and (4.81) into a single one as:

$$\sum_{j\in\mathcal{D}} R_j \leq \frac{C}{\kappa\beta + 1}. \tag{4.82}$$

The new optimisation problem can be re-written as

$$\max_{x_j\in\{0,1\}} \sum_{j=1}^{N} R_j x_j \tag{4.83}$$

$$\text{s.t. } \sum_{j=1}^{N} R_j x_j \leq \frac{C}{\kappa\beta + 1}. \tag{4.84}$$

The problem in (4.83)-(4.84) is a subset-sum problem from the family of $0-1$ Knapsack problems, that is known to be NP hard [259]. However, these type of problems are solvable optimally using dynamic programming techniques in pseudo-polynomial time [235, 259]. Furthermore, it is known that greedy heuristic approaches are bounded away from the optimal solution by half [260].

Algorithm 1 Heuristic Greedy Algorithm for (4.83)-(4.84)

procedure HEURISTIC(start, end, R_j)
  $j \leftarrow 1, R_0 \leftarrow 0, R_{N+1} \leftarrow 0$
  while $j \leq N-1$ and $\sum_{j=1}^{N} R_j x_j \leq \frac{C}{1+\kappa\beta}$ **do**
    $\sum_{j=1}^{N} R_j x_j \leftarrow \sum_{j=1}^{N} R_{j-1} x_{j-1} + R_j x_j$
    if $\sum_{j=1}^{N} R_j x_j \leq \frac{C}{1+\kappa\beta}$ **then**
      $x_j \leftarrow 1; j \leftarrow j+1$
    else do $x_j \leftarrow 0; j \leftarrow j+1$
    end if
  end while
end procedure

This work proposes a simple greedy heuristic algorithm of *linear complexity,* as follows.[2] The data subcarriers are selected starting from the best – in terms of SNR – until (4.84) is not satisfied. Once this situation occurs the last subcarrier added to set $\mathcal{D}$ is removed and the next one is added. This continues either to the last index N or until (4.84) is satisfied with equality. The algorithm is described in *Algorithm 1*.

The performance of *Algorithm 1* and the optimal solution achieved by solving the Knapsack problem 4.83 were evaluated. Furthermore, the parallel method (described in Section 4.5.1) was compared the sequential approach (described in Section 4.5.2) in terms of efficiency. For ease of reading the efficiencies of both methods are re-called here:

$$\eta_{\text{parallel}} = \frac{\mathbb{E}\left[\sum\limits_{j\in\mathcal{D}} R_j\right]}{\mathbb{E}[C]}, \tag{4.85}$$

$$\eta_{\text{sequential}} = \frac{L}{L+M}, \tag{4.86}$$

where the average number of frames needed for reconciliation is computed as:

$$M = \left\lceil \frac{\kappa C_{SKG}}{\mathbb{E}[C_R]} \right\rceil, \tag{4.87}$$

[2]Without loss of generality, the algorithm assumes that the channel gains are ordered in decreasing order as in (4.27), and, consequently, the rates R_j are also ordered in descending order. The ordering is a $\mathcal{O}(N \log N)$ operation and required in common power allocation schemes such as the waterfilling, and, therefore does not come at any additional cost.

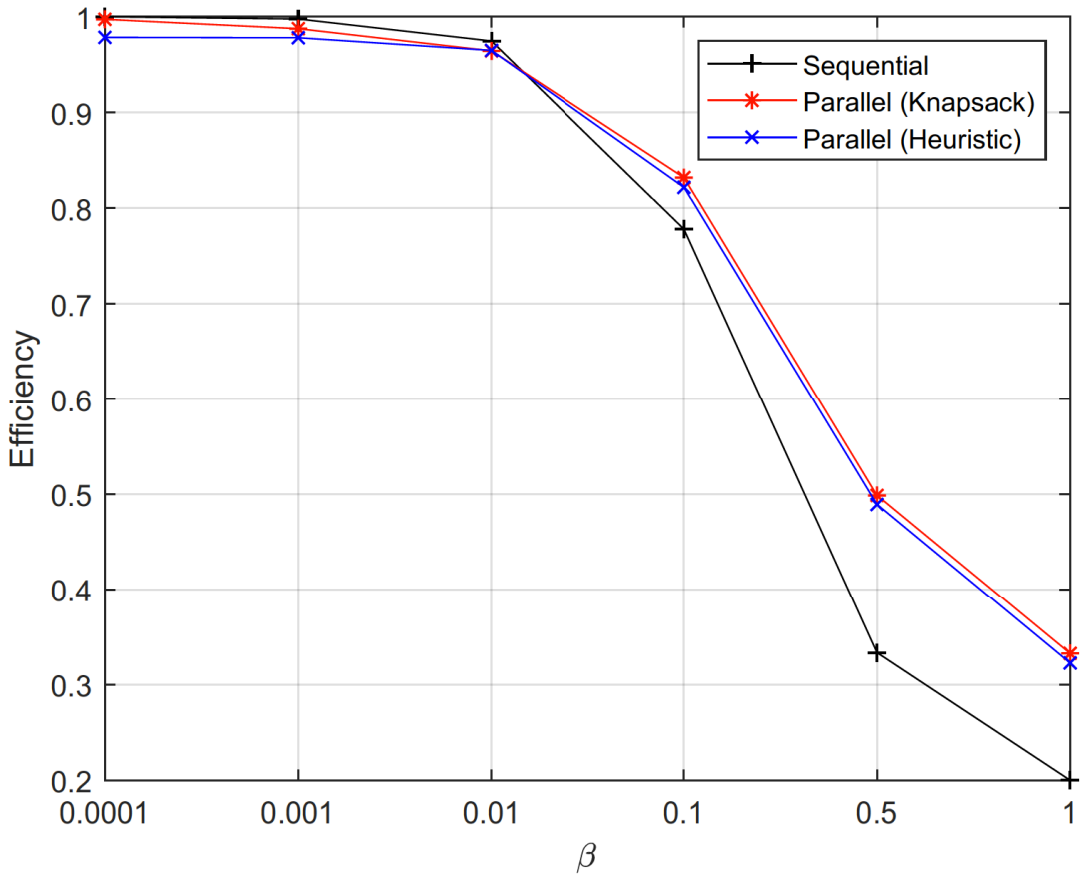

Figure 4.10: Efficiency comparison for $N = 12$, the transmit SNR=10 dB and $\kappa = 2$.

where $\lceil x \rceil$ denotes the smallest integer that is larger than x; the average number of the frames that can be sent while respecting the secrecy constraint is:

$$L = \left\lfloor \frac{C_{SKG}}{\beta \mathbb{E}[C]} \right\rfloor, \tag{4.88}$$

where $\lfloor x \rfloor$ denotes the largest interger that is smaller than x.

Numerical evaluation

This section provides numerical evaluations of the efficiency that can be achieved with the presented methods (*i.e.*, sequential and parallel) for different values of the main parameters. With respect to the parallel approach, the section provides numerical results of the optimal dynamic programming solution of the subset-sum $0-1$ knapsack problem,

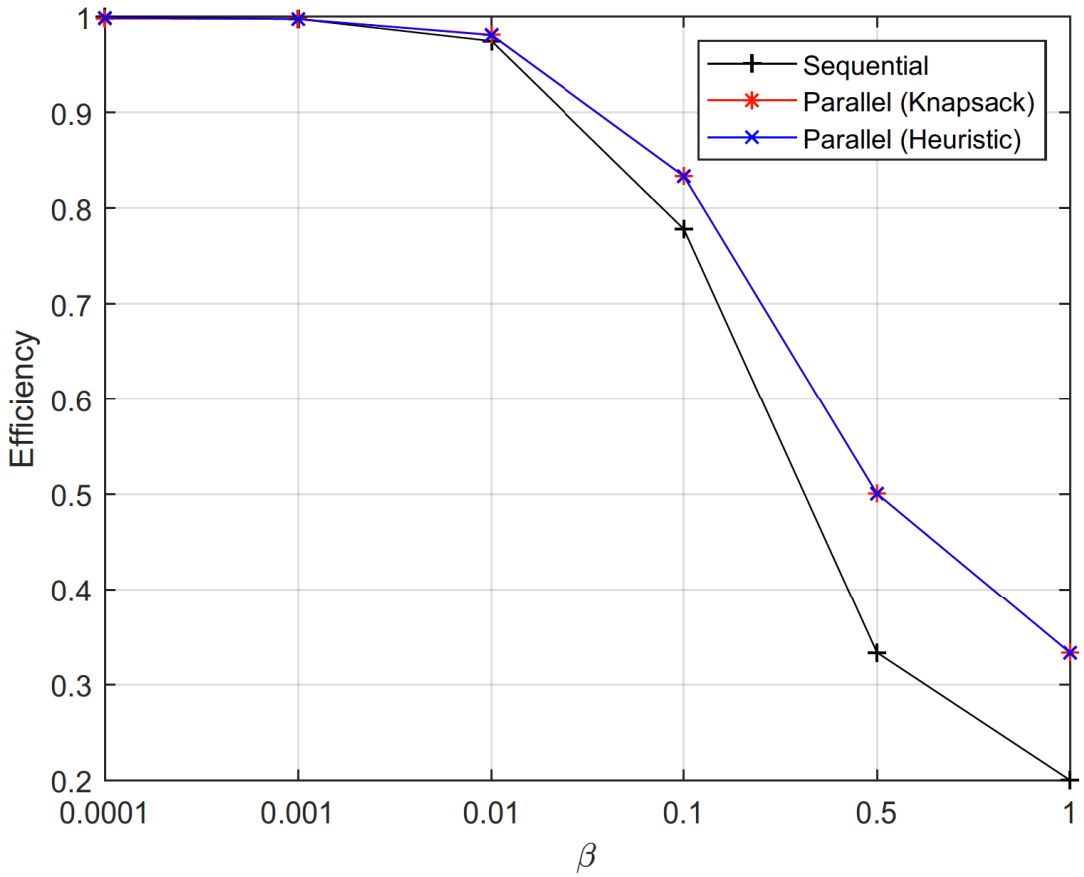

Figure 4.11: Efficiency comparison for $N = 64$, the transmit SNR=10 dB and $\kappa = 2$.

as well as of the greedy heuristic approach presented in *Algorithm 1*. Furthermore, two methods are compared through their efficiencies, *i.e.* η_{parallel} and $\eta_{\text{sequential}}$ given in (4.85) and (4.86), respectively.

Figures 4.10 and 4.11 show the efficiency of the methods for $N = 12$, and $N = 64$, respectively, while $\kappa = 2$ and $P = 10$. Note that the proposed heuristic algorithm has a near-optimal performance (almost indistinguishable from the red curves achieved with dynamic programming). Due to this fact (which was tested across all scenarios that follow) only the heuristic approach is shown in subsequent figures for clarity in the graphs. It can be seen, in Fig. 4.10, that when there are a small number of subcarriers (N=12, typical for NB-IoT) and small β the efficiency of both the parallel η_{parallel} and the sequential $\eta_{\text{sequential}}$ approaches are very close to unity, a trend that holds for increasing N. While the efficiency of the sequential and parallel methods coincide almost until around

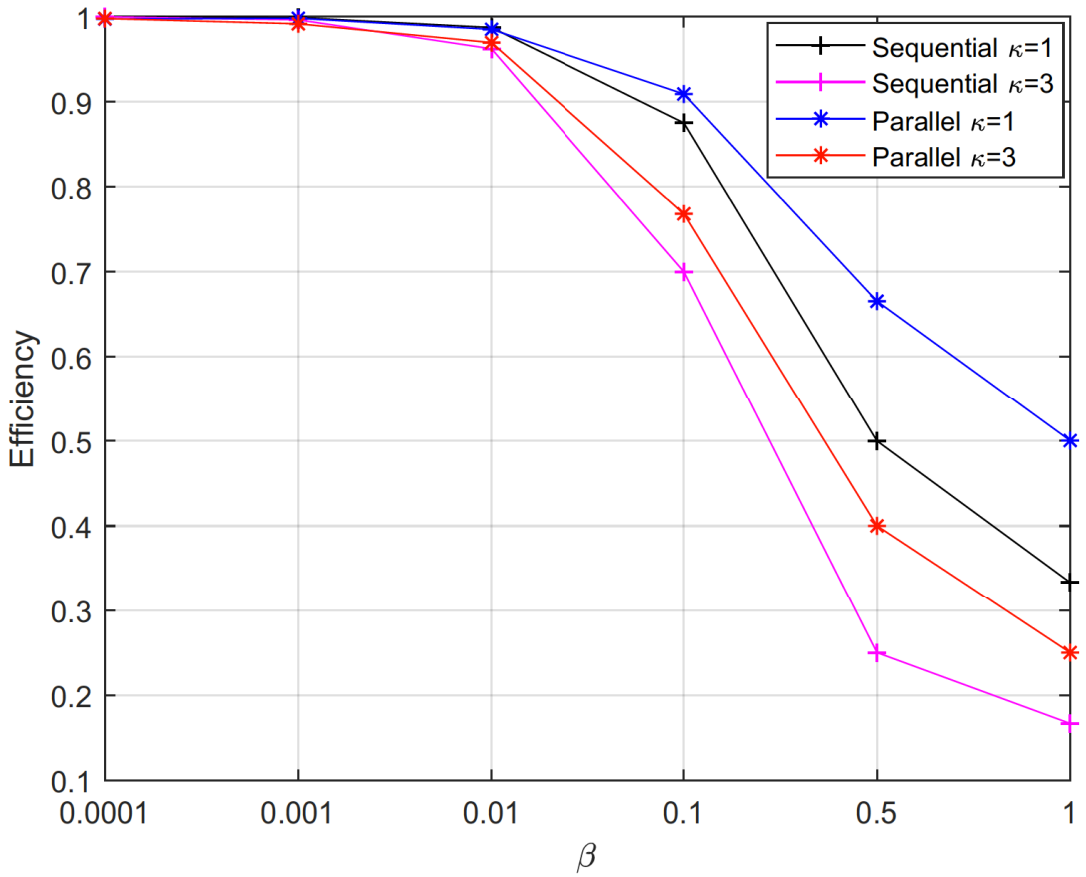

Figure 4.12: Efficiency vs κ, for $N = 24$, SNR=10 dB.

$\beta = 0.01$ with increasing β, due to the fact that more frames are needed for reconciliation in the sequential approach (*i.e.*, M increases), the parallel method proves more efficient than the sequential.

This trend can also be seen for higher number of subcarriers, *i.e.*, $N = 64$ in Fig. 4.11. Furthermore, for $N = 64$ the crossing point of the curves moves to the left and the efficiency of the two methods coincide until around $\beta = 0.001$. This trend was found to be consistent across many values of N, only two of which are shown here for compactness of presentation.

Next, in Fig. 4.12 the efficiency of the parallel η_{parallel} and the sequential $\eta_{\text{sequential}}$ methods are shown for two different values of $\kappa \in \{2, 3\}$ where SNR $= 10$ dB and $N = 24$. It is straightforward to see that they both follow similar trends and when κ increases the efficiency decreases. On the other hand, regardless of the value of κ they

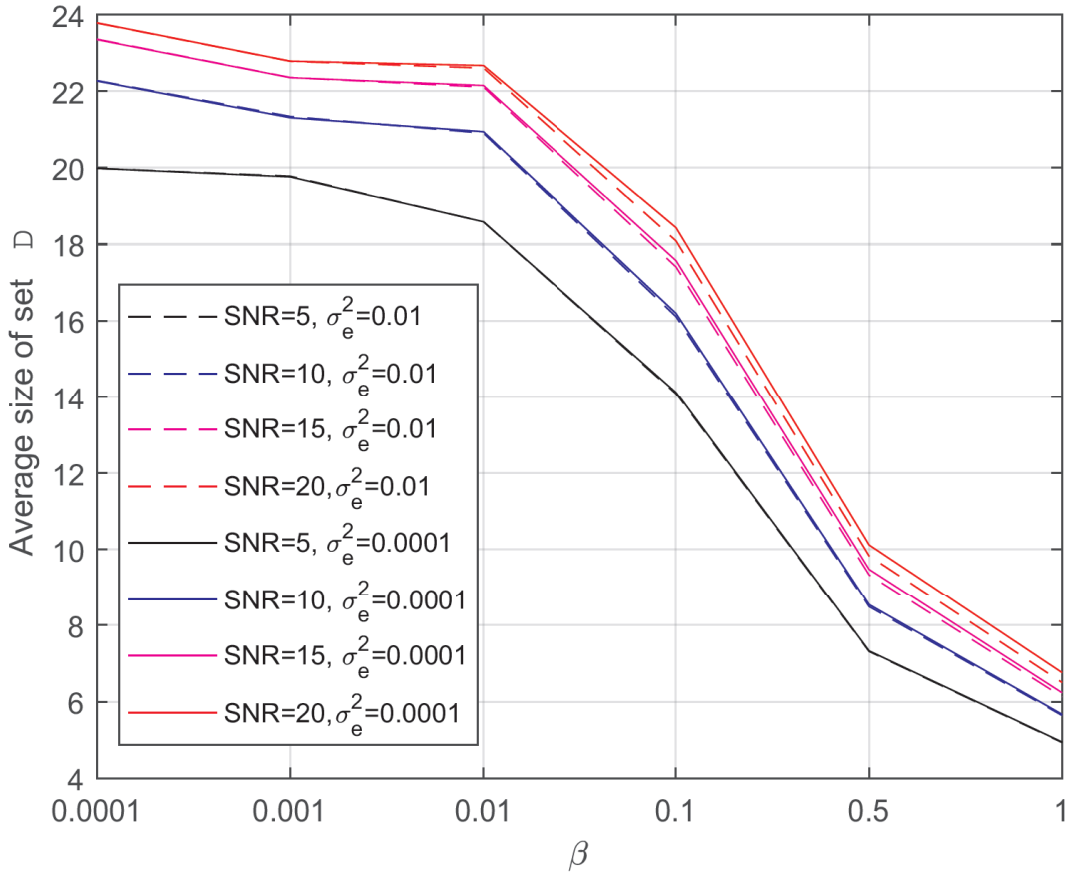

Figure 4.13: Size of set $\mathcal{D}$ for different SNR levels and σ_e^2 when $N = 24$.

both perform identically until around $\beta = 0.001$.

Finally, Fig. 4.13 and 4.14, focus on the parallel method. The figures illustrate the average size of set $\mathcal{D}$, the data channels, for different values of σ_e^2 and SNR levels, Fig. 4.13, and κ, Fig. 4.14, when $N = 24$. As expected, it can be seen that when the SNR increases the size of the set increases, too. This is due to the fact that more power is used on any single subcarrier and consequently a higher reconcilliation rate can be sustained. Regarding the estimation error σ_e^2 of the CSI, it only slightly affects the performance at high SNR levels. Hence more subcarriers have to be used for reconciliation, and fewer for data.

The SNR level in the Fig. 4.14 is set to 10 dB. The figure shows that when increasing κ the size of set $\mathcal{D}$ decreases. This result can be easily predicted from inequality (4.78), meaning, when κ increases more reconciliation data has to be sent, hence fewer subcarri-

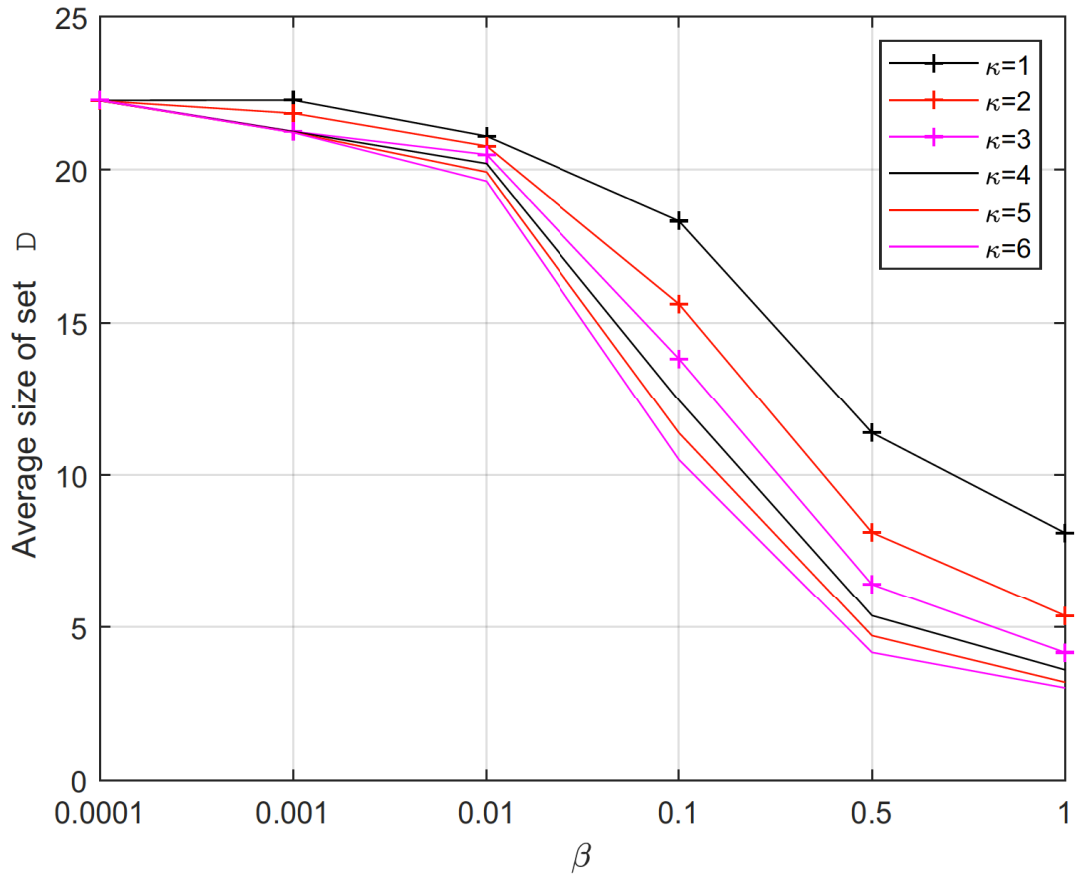

Figure 4.14: Size of set $\mathcal{D}$ for different values of κ when $N = 24$.

ers can be used for data. In both figures when β increases the size of set $\mathcal{D}$ decreases; this effect is a consequence of constraint (4.84) as the data rate is decreasing with β.

This work studied the maximisation of the data transfer rate under power, security and rate constraints, captured through the following system parameters: a factor β, representing the minimum ratio of the SKG to the data rate, and, a factor κ representing the maximum ratio of the SKG rate over the reconciliation rate. The proposed parallel method, in which SKG and data transfer are inter-weaved, was shown to perform equally well or better than a sequential approach in which the two operations were separated. Furthermore, a significant result is that although the optimal subcarrier scheduling is a $0-1$ knapsack problem, with a potentially high bound on complexity, it can be solved in linear time using a simple heuristic algorithm with virtually no loss in performance.

The next section extends this study by taking into account delay requirements.

4.6.3 Optimal allocation under security, power, rate and delay constraints

The previous section investigated the optimal power and subcarrier allocations strategy of Alice and Bob in order to maximise their long-term average data rate and proposed a greedy heuristic algorithm of linear complexity. Here, this work is extended and takes into account delay requirements. In detail, it investigates the optimal resource allocation for Alice and Bob, when their communication has to satisfy specific delay constraints. To this end, the theory of *effective capacity* [240] is used. It gives a limit for the maximum arrival rate under delay-bounds with a specified violation probability.

This section studies the *effective data rate* for the proposed pipelined SKG and encrypted data transfer scheme; the effective rate is a data-link layer metric that captures the impact of statistical delay QoS constraints on the transmission rates. As background, refer to Section 4.2.2. As showed in [261], the probability of a steady-state queue length process $Q(t)$ exceeding a certain queue-overflow threshold x converges to a random variable $Q(\infty)$ as:

$$\lim_{x\to\infty} \frac{\ln(\mathsf{Pr}[Q(\infty) > x])}{x} = -\theta, \tag{4.89}$$

where θ indicates the asymptotic exponential decay-rate of the overflow probability. For a large threshold x, (4.89) can be represented as $\mathsf{Pr}[Q(\infty) > x] \approx e^{-\theta x}$. Furthermore, the delay-outage probability can be approximated by [240]:

$$\mathsf{Pr}^{\text{out}}_{\text{delay}} = \mathsf{Pr}[\text{Delay} > D_{\max}] \approx \mathsf{Pr}[Q(\infty) > 0] e^{-\theta\zeta D_{\max}}, \tag{4.90}$$

where $D_{\max}$ is the maximum tolerable delay, $\mathsf{Pr}[Q(\infty) > 0]$ is the probability of a non-empty buffer, which can be estimated from the ratio of the constant arrival rate to the averaged service rate, ζ is the upper bound for the constant arrival rate when the statistical delay metrics are satisfied.

Using the delay exponent (θ) and the probability of non-empty buffer, the effective

capacity, that denotes the maximum arrival rate, can be formulated as [240]:

$$E_C(\theta) = -\lim_{t\to\infty} \frac{1}{\theta} \ln \mathbb{E}[e^{-\theta S[t]}](\text{bits/s}), \tag{4.91}$$

where $S[t] = \sum_{i=1}^{t} s[i]$ denotes the time-accumulated service process, and $s[i], i = 1, 2, ...$ denotes the discrete-time stationary and ergodic stochastic service process. Therefore, the delay exponent θ indicates how strict the delay requirements are, *i.e.*, $\theta \to 0$ corresponds to looser delay requirements, while $\theta \to \infty$ implies exceptionally stringent delay constraints. Assuming a Rayleigh block fading system, with frame duration T_f and total bandwidth B, we have $s[i] = T_f B\tilde{R}_i$, with $\tilde{R}_i$ representing the instantaneous service rate achieved during the duration of the i^{th} frame. In the context of the investigated data and reconciliation information transfer, $\tilde{R}_i$, is given by:

$$\tilde{R}_i = \frac{1}{F} \sum_{i\in\mathcal{D}} \log_2(1 + p_i \hat{g}_i), \tag{4.92}$$

where F is the equivalent frame duration, *i.e.*, the total number of subcarriers used for data transmission, so that for the parallel approach we have $F = |D|$ while for the sequential approach $F = N(L + M)L^{-1}$.

Under this formulation and assuming that Gärtner-Ellis theorem [236,237] is satisfied, the *effective data rate*[3] $E_C(\theta)$ is given as:

$$E_{C,\mathcal{D}}(\theta) = -\frac{1}{\theta T_f B} \ln\left(\mathbb{E}\left[e^{-\theta T_f B\tilde{R}_i}\right]\right). \tag{4.93}$$

By setting $\alpha = \frac{\theta T_f B}{\ln(2)}$ and inserting (4.92) into (4.93):

$$E_{C,\mathcal{D}}(\theta) = -\frac{1}{\ln(2)\alpha} \ln\Big(\mathbb{E}\Big[e^{-\ln(2)\alpha F^{-1}\sum_{i\in\mathcal{D}}\log_2(1+p_i\hat{g}_i)}\Big]\Big),$$

[3]Since part of the transmission rate is used for reconciliation information, and part for data transmission the terms "*effective syndrome rate*" and "*effective data rate*" are introduced instead of the term "effective capacity", for rigour. Note that the information data and reconciliation information are accumulated in separate independent buffers within the transmitter.

$$E_{C,\mathcal{D}}(\theta) = -\frac{1}{\alpha}\log_2\left(\mathbb{E}\left[\prod_{i\in\mathcal{D}}\left(1+p_i\hat{g}_i\right)^{-\alpha F^{-1}}\right]\right). \tag{4.94}$$

Assuming i.i.d. channel gains, by using the distributive property of the mathematical expectation, (4.94) becomes [262]:

$$E_{C,\mathcal{D}}(\theta) = -\frac{1}{\alpha}\log_2\left(\prod_{i\in\mathcal{D}}\mathbb{E}\left[\left(1+p_i\hat{g}_i\right)^{-\alpha F^{-1}}\right]\right). \tag{4.95}$$

After further manipulations and using the log-product rule:

$$E_{C,D}(\theta) = -\frac{1}{\alpha}\sum_{i\in\mathcal{D}}\log_2\left(\mathbb{E}\left[\left(1+p_i\hat{g}_i\right)^{-\alpha F^{-1}}\right]\right). \tag{4.96}$$

Similarly, the *effective syndrome rate* can be written as:

$$E_{C,\breve{D}}(\theta) = -\frac{1}{\alpha}\sum_{i\in\breve{\mathcal{D}}}\log_2\left(\mathbb{E}\left[\left(1+p_i\hat{g}_i\right)^{-\alpha \breve{F}^{-1}}\right]\right), \tag{4.97}$$

where the size of $\breve{F}$ here is $|N-D|$.

Using that, the maximisation problem given in (4.76) is now reformulated by adding a delay constraint. The reformulated problem can be expressed as follows:

$$\max_{p_j, j\in\mathcal{D}} \quad E_{C,\mathcal{D}}(\theta), \tag{4.98}$$

s.t.

$$\sum_{j=1}^{N} p_j \leq NP, \tag{4.99}$$

$$\sum_{j\in\mathcal{D}} R_j \leq \frac{\sum\limits_{j\in\breve{\mathcal{D}}} R_j}{\kappa\beta}, \tag{4.100}$$

$$E_{C,\mathcal{D}}(\theta) + E_{C,\breve{\mathcal{D}}}(\theta) \leq E_C^{\text{opt}}(\theta), \tag{4.101}$$

where $E_C^{\text{opt}}(\theta)$ represents the maximum achievable effective capacity for both key and

data transmission for a given value of θ over N subcarriers:

$$E_C^{\text{opt}}(\theta) = \max_{p_i, i=1,2,\ldots N} \left\{ -\frac{1}{\alpha} \log_2 \left(\mathbb{E}\left[\prod_{i=1}^{N} (1 + p_i \hat{g}_i)^{-\alpha N^{-1}} \right] \right) \right\}. \tag{4.102}$$

The proposed approach, assumes that the constraint (4.101) is satisfied with equality. Given that, the optimisation problem in (4.98) can be evaluated as two sub-optimisation problems: i) finding the optimal long term power allocation from (4.99) and (4.102); ii) finding the optimal subcarrier allocation that satisfies (4.100). The first problem that gives the optimal power allocation can be solved using convex optimisation tools. Next, as in Section 4.6.2 two methods are used to solve the subcarrier allocation problem, *i.e.*, by formulating a subset-sum 0 – 1 knapsack optimisation problem or through a variation of *Algorithm 1*. The efficiency of both methods will be again compared numerically to the sequential method.

Now, following the same steps from (4.94) to (4.96) and using the fact that maximising $E_C(\theta)$ is equivalent to minimising $-E_C(\theta)$ (this is due to $\log(\cdot)$ being a monotonically increasing concave function for any $\theta > 0$) the following minimisation problem is formulated:

$$\min_{p_i, i=1,2,\ldots N} \sum_{i=1}^{N} \left(\mathbb{E}\left[(1 + p_i \hat{g}_i)^{-\alpha N^{-1}} \right] \right), \tag{4.103}$$

$$\text{s.t. (4.99)}.$$

where $F = N$ in this case as the full set of subcarriers is concerned. Next the the Lagrangian function $\mathcal{L}$ is formed as:

$$\mathcal{L} = \left(\mathbb{E}\left[(1 + p_i \hat{g}_i)^{-\alpha N^{-1}} \right] \right) + \lambda \left(\sum_{i=1}^{N} p_i - NP \right). \tag{4.104}$$

By differentiating (4.104) w.r.t. p_i and setting the derivative equal to zero [230] results

into:

$$\frac{\partial \mathcal{L}}{\partial p_i} = \lambda - \frac{\alpha \hat{g}_i}{N} \left(\hat{g}_i p_i + 1\right)^{-\frac{\alpha}{N}-1} = 0. \tag{4.105}$$

Solving (4.105) gives the optimal power allocation policy:

$$p_i^* = \frac{1}{g_0^{\frac{N}{\alpha+N}} \hat{g}_i^{\frac{\alpha}{\alpha+N}}} - \frac{1}{\hat{g}_i}, \tag{4.106}$$

where $g_0 = \frac{N\lambda}{\alpha}$ is the cutoff value which can be found from the power constraint. The expression of $E_C^{\text{opt}}(\theta)$ can be found by inserting p_i^* in $E_C(\theta)$:

$$E_C^{\text{opt}}(\theta) = -\frac{1}{\alpha} \sum_{i=1}^{N} \log_2 \left(\mathbb{E} \left[\left(\frac{\hat{g}_i}{g_0} \right)^{-\frac{\alpha}{\alpha+N}} \right] \right) \tag{4.107}$$

When $\theta \to 0$ the optimal power allocation is equivalent to water-filling and when $\theta \to \infty$ the optimal power allocation transforms to total channel inversion.

Now, fixing the power allocation as in (4.106) the optimal subcarrier allocation that satisfies (4.100) easily can be identified. As in Section 4.6.2, to do that first a subset-sum 0 – 1 knapsack optimisation problem is formulated, which is solved using the standard dynamic programming approach. Furthermore, the performance of the heuristic algorithm presented in *Algorithm 1* is also evaluated.

Numerical evaluation

Inspired by the good performance of *Algorithm 1*, in the case where long-term average rate is the metric of interest, here, the investigation continues with a variation of *Algorithm 1*, with the following differences: at lines 3 and 5 instead of (4.82) the constraint (4.100) is used and the power allocation is fixed as in (4.106). The performance of the system is again compared with a sequential method and the metric of interest here is the *effective data rate*. The comparison is performed by taking into account the following parameters: signal to noise ration (SNR); number of subcarriers N; ratio of the reconciliation rate to

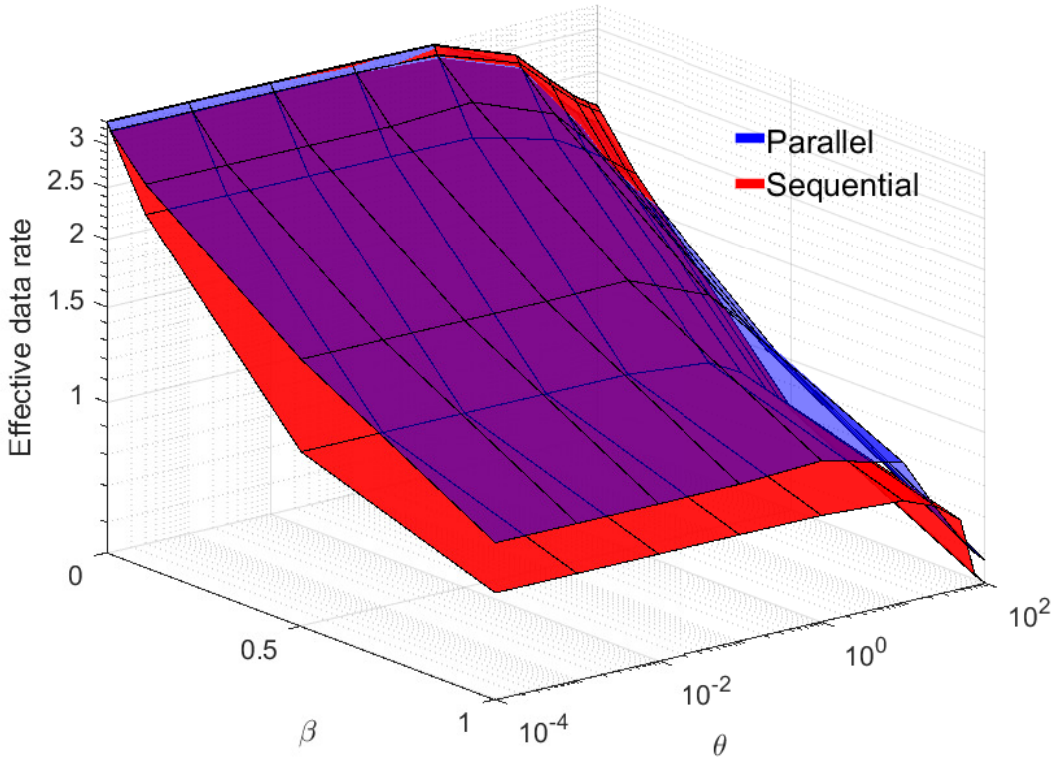

Figure 4.15: Effective data rate achieved by the parallel heuristic approach and the sequential approach when, SNR$=10$ dB and $\kappa = 2$ and $N = 12$.

the SKG rate κ; delay exponent θ; and, the ratio of key bits to data bits β.

Figures 4.15 - 4.18 illustrate three-dimensional plots showing the dependence of the achievable *effective data rate* $E_{C,D}(\theta)$ on β and θ. Figures 4.15 and 4.17 compare the parallel heuristic approach and the sequential approach for high SNR levels, whereas Fig. 4.16 and 4.18 compare their performance for low SNR level. In Fig. 4.15 and 4.16 $N = 12$ while in Fig. 4.17 and 4.18 the total number of subcarriers is $N = 64$. All graphs compare the performance of the heuristic parallel approach and the sequential approach for $\kappa = 2$.

As discussed earlier in this section, when the delay exponent θ increases, the optimal power allocation transforms from water-filling to total channel inversion. Consequently, the rate achieved on all subcarriers converges to the same value, hence when having a

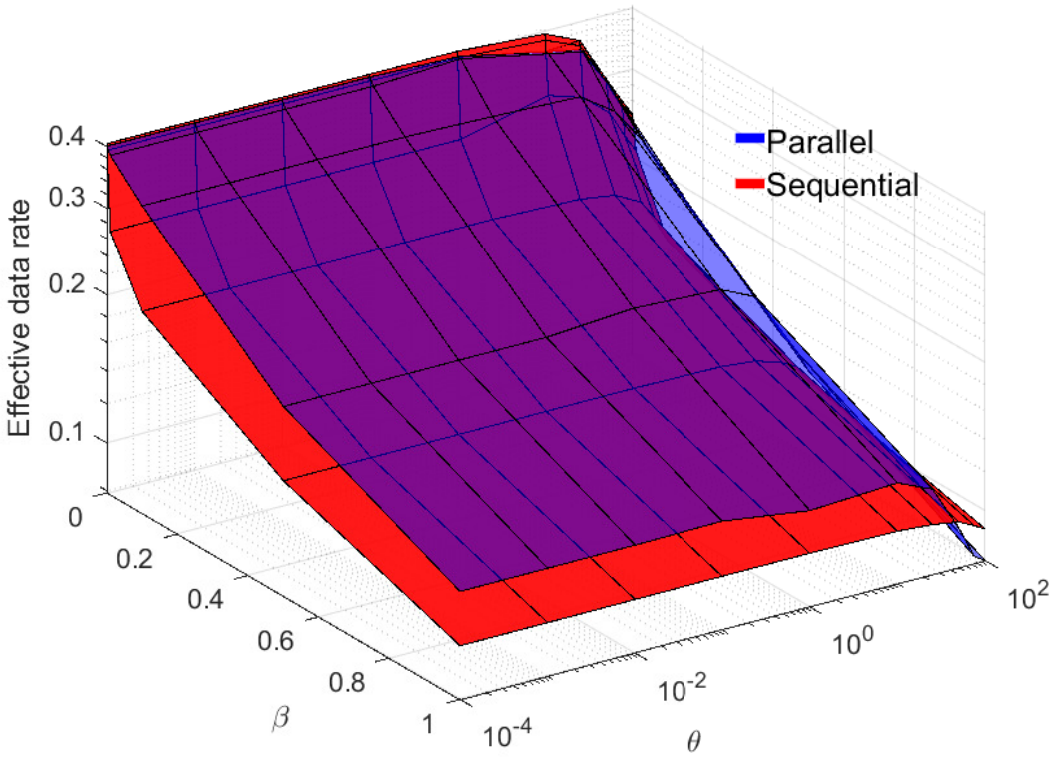

Figure 4.16: Effective data rate achieved by the parallel heuristic approach and the sequential approach when, SNR$=0.2$ dB and $\kappa = 2$ and $N = 12$.

small number of subcarriers (such as $N = 12$ in Fig. 4.15) and small values of β then using a single subcarrier for reconciliation data will use more capacity than needed and most of the rate on this subcarrier is wasted. Devoting a whole subcarrier for sending the reconciliation data for the case of $N = 12$ and $\beta = 0.0001$ is almost equivalent of losing $1/12$ of the achievable rate. This effect can be seen in both Fig. 4.15 and 4.16 where $N = 12$. When the SNR is high (See Fig. 4.15, as discussed, this effect is mostly noticeable for large values of θ and small values of β[4], whereas for small values of β and θ both algorithms perform nearly identically. A similar trend can be seen at the low SNR regime in Fig. 4.16. However, at a low SNR the sequential approach has a lower effective data rate. This happens because at high SNR levels each reconciliation frame will contain

[4] *i.e* that the ratio of reconciliation information to data is small as seen from Eq. (4.100))

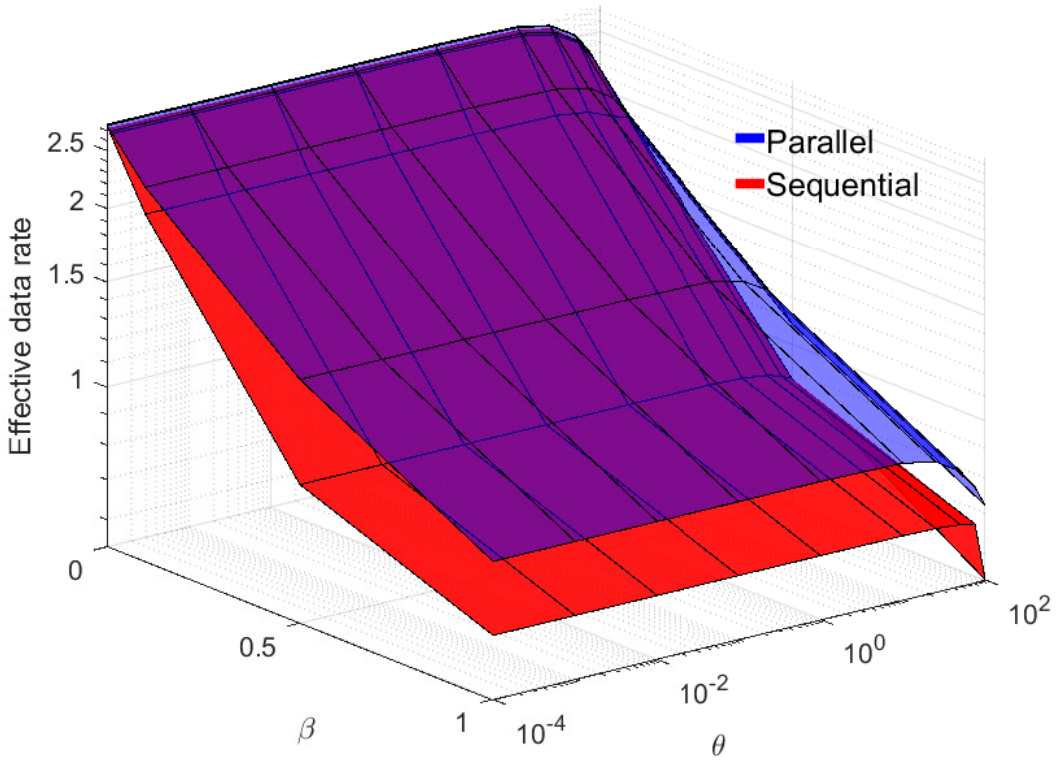

Figure 4.17: Effective data rate achieved by the parallel heuristic approach and the sequential approach when, SNR$=10$ dB and $\kappa = 2$ and $N = 64$.

more information and hence more data frames will follow. Therefore, at the low SNR regime, the reconciliation information received will decrease, hence less data can be sent afterwards. This does not affect the parallel approach. However, in both scenarios with high and low SNR, when β increases, regardless of the value of θ, the parallel approach always achieves higher *effective data rate* $E_{C,D}(\theta)$.

In the next case, when the total number of subcarriers is $N = 64$, illustrated in Fig. 4.17 and 4.18, it can be seen that the penalty of devoting a high part of the achievable effective capacity $E_C^{\text{opt}}(\theta)$ to reconciliation disappears and the heuristic parallel approach always achieves higher or identical *effective data rate* $E_{C,D}(\theta)$ compared to the sequential approach. This trend repeats for high and low SNR levels as given in Fig. 4.17 and 4.18, respectively.

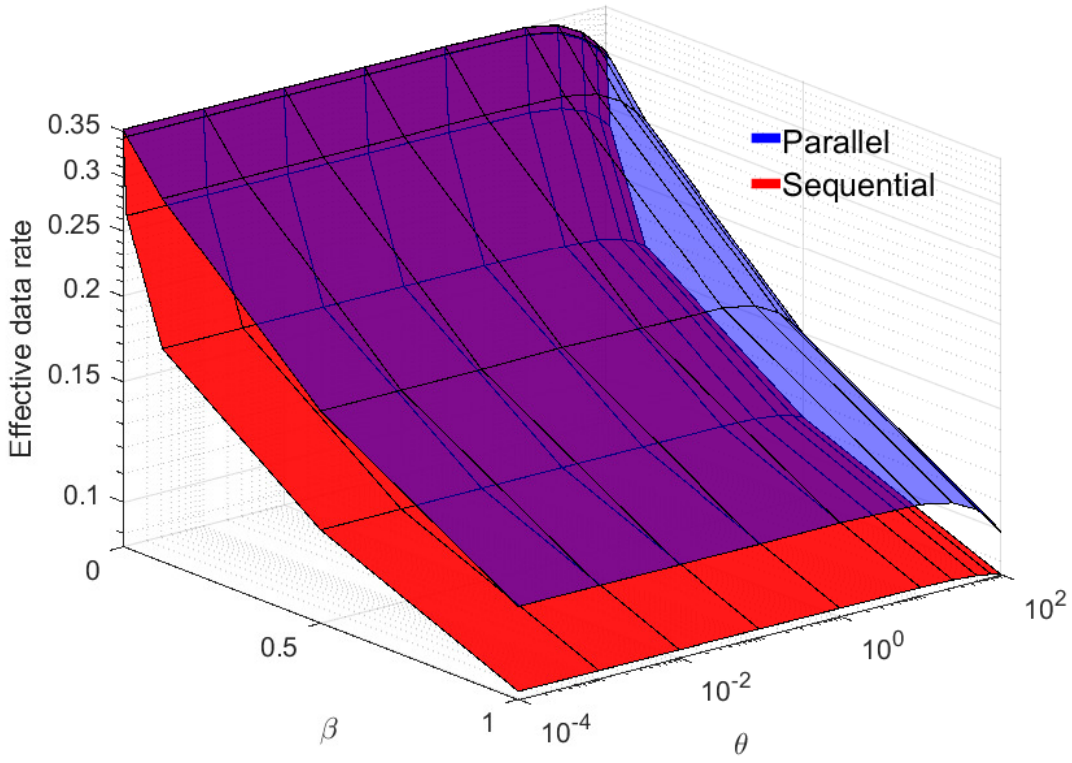

Figure 4.18: Effective data rate achieved by the parallel heuristic approach and the sequential approach when, SNR$=0.2$ dB and $\kappa = 2$ and $N = 64$.

Now, taking to take a closer look some specific cases from the 3d plots are transformed to two-dimensional graphs. Figures 4.19 – 4.22 show the achieved *effective data rate* $E_{C,D}(\theta)$ given in (4.96), for different values of N and θ while the SNR=5 dB and $\kappa = 2$. Fig. 4.19 gives the achieved effective rate on set $\mathcal{D}$ for $N = 12$ and $\theta = 0.0001$ (relaxed delay constraint). Similarly to the case of long term average value of C_D it can be seen that for small values of β the sequential approach achieves slightly higher effective data rate. As before, the increase of β results in more reconciliation frames M required in the sequential case. This effect is not seen in the parallel case and for high values of β it performs better.

Fig. 4.20 illustrates the case when $N = 12$ and $\theta = 100$ (very stringent delay constraint). It can be seen that for small values of β the sequential approach performs better

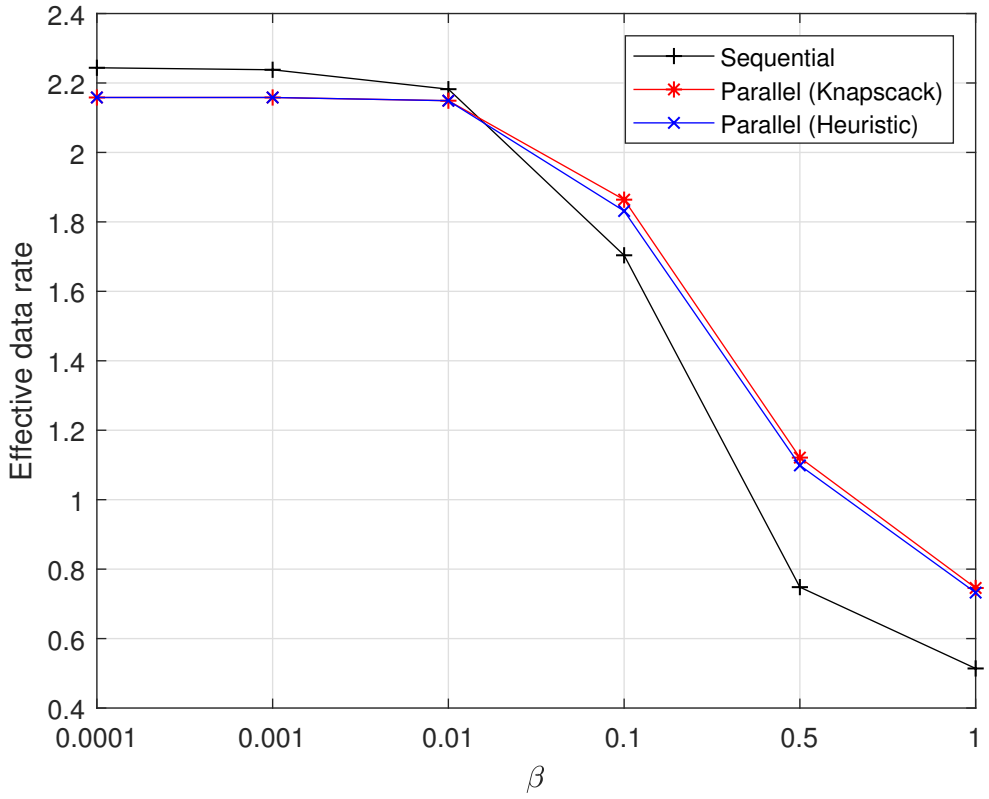

Figure 4.19: Effective data rate achieved by parallel and sequential approaches when $N = 12$, SNR$= 5$dB, $\kappa = 2$ and $\theta = 0.0001$.

than the parallel. As discussed, the efficiency loss is caused by the fact that the devoted part of the total achievable effective capacity $E_C^{\text{opt}}(\theta)$ to reconciliation (syndrome communication) is more than what is required. However, a higher β leads to an increase in the reconciliation information that needs to be sent, and the rate of the subcarriers in set $\breve{\mathcal{D}}$ will be fully or almost fully utilised and the parallel approach shows better performance for these values.

Next, Fig. 4.21 and 4.22 show the performance of the two algorithms for higher value of $N = 64$. It is easy to see that regardless of the value of θ and β both algorithms perform identical or the parallel is better. In the previous case of $N = 12$ increasing θ might reduce the effectiveness of the parallel approach, however when $N = 64$ increasing θ does not incur such a penalty and the parallel is either identical to the sequential or outperforms it.

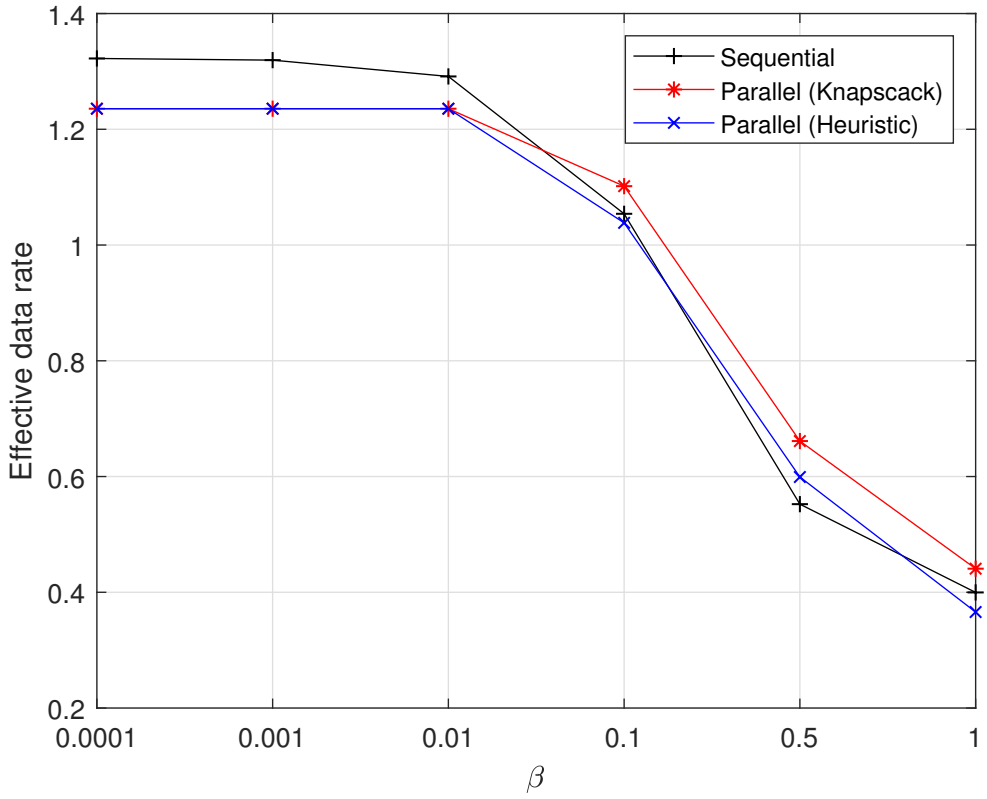

Figure 4.20: Effective data rate achieved by parallel and sequential approaches when $N = 12$, SNR$= 5$dB, $\kappa = 2$ and $\theta = 100$.

This can be seen in both cases, *i.e.*, $\theta = 0.0001$, in Fig. 4.21, and $\theta = 100$, in Fig. 4.22.

Another interesting fact from Figures 4.19 – 4.22 is that looking at the parallel approach, it can easily be seen that in all cases the heuristic approach almost always performs as well as the optimal knapsack solution. The case of small values of θ is similar to the one when long term average rate was used and choosing the best subcarriers for data transmission works as well as the optimal Knapsack solution. Interestingly, *Algorithm 1* works well for high values of θ, too. This can be explained by the fact that when θ increases the rate on all of the subcarriers becomes similar and switching the subcarriers in set $\mathcal{D}$ does not incur high penalty.

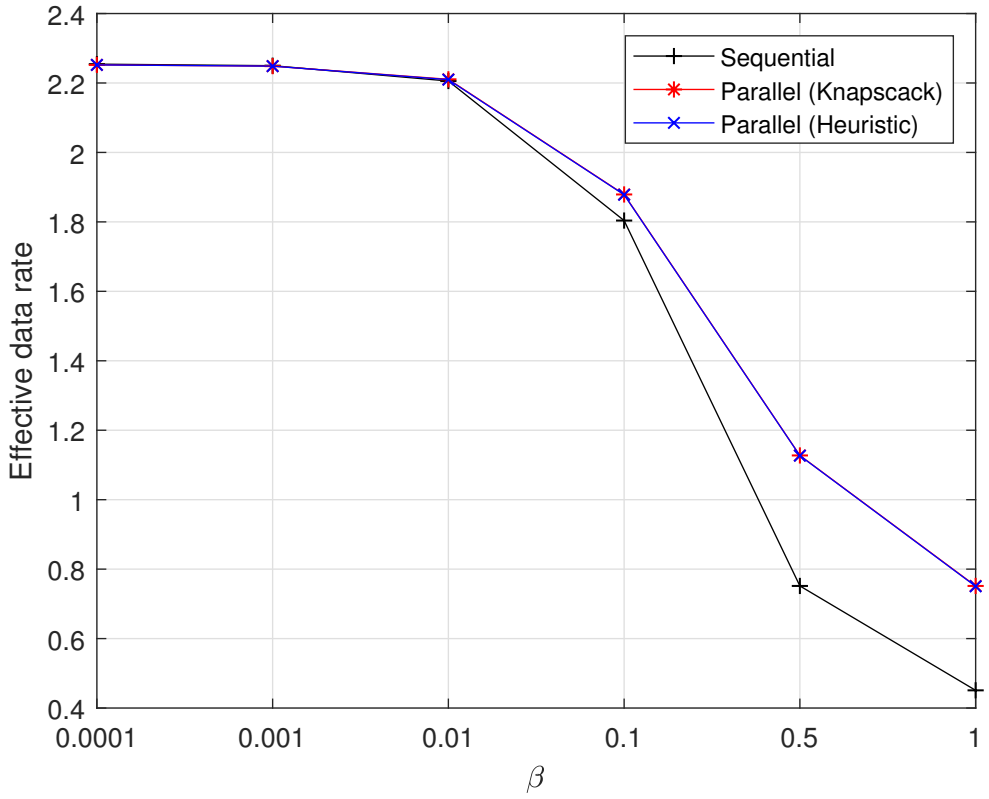

Figure 4.21: Effective data rate achieved by parallel and sequential approaches when $N = 64$, SNR$= 5$dB, $\kappa = 2$ and $\theta = 0.0001$.

4.7 Summary

This work explored the possibility of pipelining encrypted data transfer and SKG in a Rayleigh BF-AWGN environment. It investigated the maximisation of the data transfer rate in parallel to performing SKG. The work took into account imperfect CSI measurements and the effect of order statistics on the channel variance. Three scenarios were differentiated in this study: i) the optimal data transfer rate was found under power and security constraints – represented by the system parameters β, which represents the minimum ratio of SKG rate to data rate; ii) the optimal data transfer rate was found under power, security and rate constraint – represented by the system parameters κ, which represents the maximum ratio of SKG rate to reconciliation rate; iii) by adding a delay con-

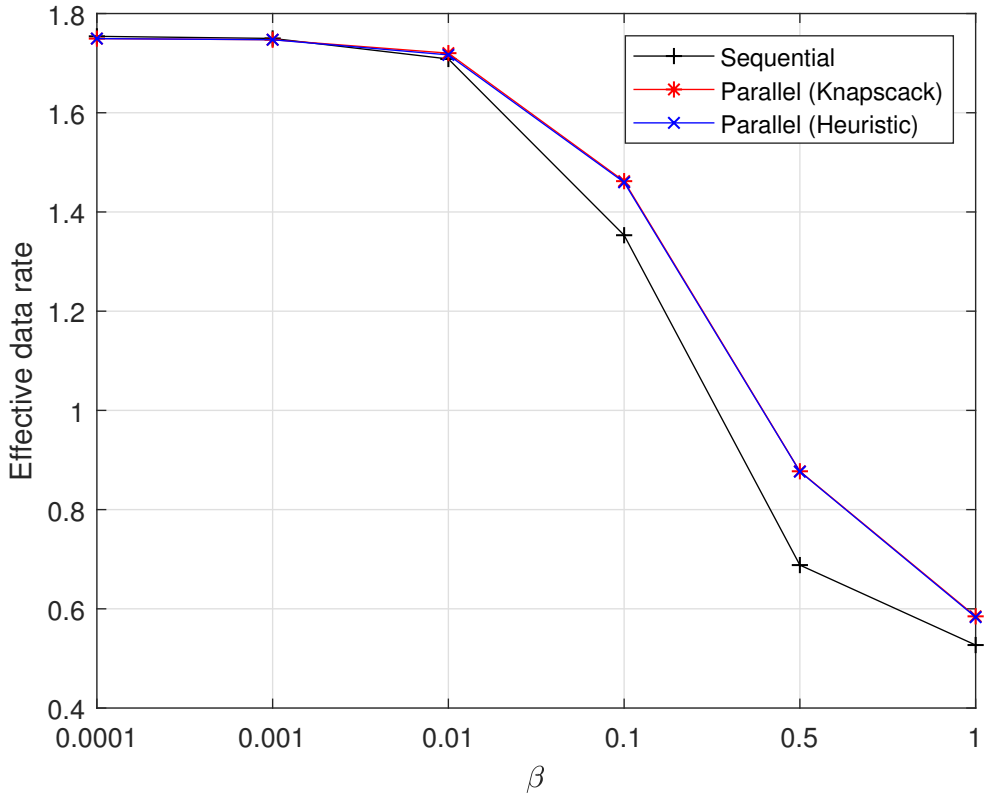

Figure 4.22: Effective data rate achieved by parallel and sequential approaches when $N = 64$, SNR$= 5$dB, $\kappa = 2$ and $\theta = 100$.

straint, represented by parameter θ, the optimal *effective data rate* was found. The latter shows the first study introducing the theory of effective capacity with SKG.

The study was finalised by numerical comparisons of the efficiency of the proposed parallel method, in which SKG and data transfer are inter-weaved and a sequential method where the two operations are done separately. The results of the two scenarios showed that in most of the cases the performance of both methods, parallel and sequential, is either equal or the parallel performs better. As the possible advantage of using the sequential is small and only applies in particular scenarios, the parallel scheme is recommended as a universal mechanism for general protocol design, when latency is an issue. Furthermore, a significant result is that although the optimal subcarrier scheduling is an NP hard $0-1$ knapsack problem, it can be solved in linear time using a simple heuristic algorithm with

virtually no loss in performance.

Overall this chapter presented a novel SKG approach that can greatly reduce the system's latency through pipelining as the SKG information is sent in parallel with the data. This allows the data to be decrypted immediately without a further round of SKG transmission. Next, Chapter 5 will discuss what is the optimal strategy of Alice and Bob when an active attacker tries to interrupt their SKG process.

Chapter 5

Man-in-the-middle and denial of service attacks in PLS systems

---

This chapter addresses the problem of active jamming attacks. The chapter presents countermeasures in the form of power allocation strategies. Game theoretic analysis is used to confirm the optimal solutions. The results in this chapter were presented at the "IEEE Global Communications Conference (Globecom 2019)" [15].

5.1 Introduction

The SKG process from shared randomness (*e.g.*, from the wireless channel fading realisations), was well investigated in the previous chapters: Chapter 3 showed how the SKG can be used in authentication protocols and session key agreement schemes; Chapter 4 proposed an improved SKG solution where data and key rates are jointly optimised. The roadmap from Chapter 4 to this chapter is illustrated in Fig. 5.1. This chapter will investigate the SKG process accounting for jamming attacks in PLS systems. It will first investigate the impact of injection and reactive jamming attacks during the "advantage distillation" phase in SKG (when observations of the shared randomness are obtained at

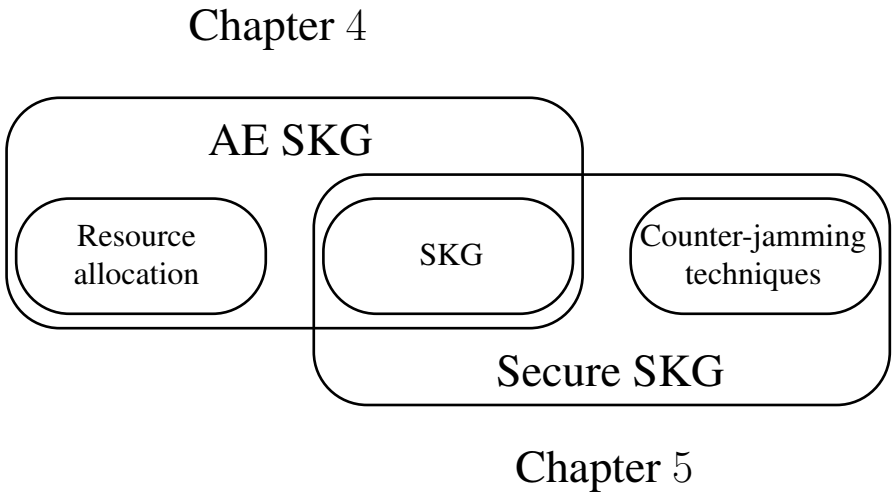

Figure 5.1: Roadmap from Chapter 4 to Chapter 5.

the legitimate parties). As an example, an active attacker can act as a man-in-the-middle (MiM) by injecting pilot signals and/or can mount denial of service attacks (DoS) in the form of jamming. The optimal strategies of jammer and legitimate users are identified in a BF-AWGN channel, using a game theoretic formulation.

The contributions of this chapter are as follows:

1. Proposal of pilot randomisation scheme as a countermeasure to injection attacks. The scheme reduces injection attacks injection MiM attacks to less harmful jamming attacks.

2. Investigation of the optimal strategy of a reactive jammer who has a fixed sensing threshold.

3. Identifying the Stackelberg equilibrium that gives the optimal power allocation for Alice and Bob in the present of a reactive jammer with fixed sensing threshold.

4. Investigation of the optimal strategy for an intelligent reactive jammer who can strategically choose their sensing threshold.

5. Identifying the Stackelberg equilibrium that gives the optimal power allocation for Alice and Bob in the present of a reactive jammer when they strategically choose

their sensing threshold.

5.2 Respective background

SKG schemes have been shown to be vulnerable to DoS attacks in the form of jamming and to man in the middle attacks implemented as injection attacks. Therefore, this section introduces the types of jamming attacks and known counter-jamming techniques. The section concludes with a brief introduction to game theory which is used in this chapter as a tool to identify the optimal strategies in presence of a jammer.

5.2.1 Jamming attacks

Various types of jamming techniques in wireless communications are presented in [263], these include:

Noise jamming attack: Theoretically proven as the most harmful attack when the adversary does not have knowledge of the legitimate users' channel. The goal of the attack is to increase the noise and interrupt the communication.

Equalisation jamming attack: Instead of jamming the entire signal the goal of this attack is to affect only part of the transmission. An example is a pilot jamming attack which is concentrated during the pilot exchange. This attack aims to disrupt the equalisation process.

Injection jamming attack: The adversary injects signals during the advantage distillation phase between two legitimate users. The signals target each of the legitimate users and are constructed such that both parties have the same channel observation. A successful attack will allow the adversary to derive a substantial part of the generated key.

The efficiency of some of these jamming approaches are numerically evaluated in [264]. This chapter concentrates on equalisation jamming during the pilot exchange and

shows how the injection jamming attack can be transformed into a less harmful equalisation jamming attack.

On the other hand, jammers can be divided based upon their strategy as: active and reactive jammers. Active jamming is simplistic approach where the adversary aims to increase the noise / interference level of the channel, regardless if there is ongoing communication. Reactive jamming is a stealthy jamming approach where the jammer first senses the spectrum and jams only if detects an ongoing transmission. Reactive jamming attacks are considered as the most harmful [265–269] as the optimise the jamming dynamically to maximise the impact. This chapter proposes a set of countermeasures to both active and reactive jammers

5.2.2 Countermeasures

In [265] - [267] the authors propose reactive jammer detection algorithms assuming the jammer needs a small period of time (reaction time) in order to detect the transmission and switch from listening to jamming. As an example, [265] proposes a jamming detector based on threshold. The detector uses the unjammed bits, transmitted during the jammer's switching period, to estimate a threshold and compares following bits to that threshold in order to detect jamming attacks. A similar approach is taken in [267, 270, 271], where reactive jamming attacks are detected based on pre-stored non-jammed samples. As an example, in [271] the variance of pilot signals is used to categorised them as jammed or non-jammed. The authors of [269] estimate the probabilities of correct and false detection of the reactive jammers using a game-theoretic analysis. As discussed above, numerous techniques have been introduced for the detection of reactive jammer. Therefore, the work presented in this chapter identifies the optimal strategy of the legitimate users in a presence of a reactive jammer.

An interesting approach to cope with a reactive jammers that employs uniform power allocation is given in [272]. During the pilot exchange phase between two legitimate users each of them estimates signal-to-interference-plus-noise (SINR) ratio on each subcarrier.

Next, both parties determine the received jamming power on each subcarrier and use only the subcarriers with high SINR ratio for communication. Another perspective of the jammer's role in a communication system is given in [273]. The work assumes a scenario of a single transmitter that has knowledge of the channel, a pair of receivers and a pair of jammers. While the transmitter is trying to maximise the overall rate, each jammer aims to help a single user by jamming the other receiver. The work identifies the optimal power allocations for jammers and transmitter which, in fact, are both modified water-filling algorithms. Section 5.5.1 investigates the scenario where channel observations are not yet present at the receiver or transmitter (*i.e.*, during the SKG process) and identifies the optimal power allocation in such a scenario.

As jamming attacks represent a critical vulnerability for wireless SKG systems, in [73,74,102,274], the employment of energy harvesting (EH) was investigates as a counter-jamming approach. In particular, in the case of EH receivers, the existence of a critical transmission power for the legitimate nodes enabled the complete neutralisation of the jammer. This thesis has not considered this type of approach, but it would interesting future work and the author is currently working on topic.

5.2.3 Game-theoretic analysis of active attacks

Game theory is a study of mathematical models which provides a set of tools for analysing interactive decision problems. It is usually used to model the strategic interaction between two or more players in situations where a player's choice have direct impact on others [275, 276]. The type of games within the theory can be divided as cooperative games and non-cooperative games. In a cooperative game players form coalitions in order to strengthen their position in a game. In a non-cooperative games each player is independent and chooses their strategy in order to increase their own benefit. The present thesis focus on non-cooperative games. More specifically, Section 5.5.1 uses the theory to evaluate the optimal strategy of a pair of legitimate users in the presence of a jammer.

Game theory has proven a reliable analysis tool for numerous problems in communi-

cations [15,74,277–279]. A non-cooperative game is framed by 1) identifying the players and their possible actions; 2) defining the payoffs as a function of the actions; 3) evaluating equilibria and possible outcomes. The mathematical definition of a non-cooperative game is described by three elements, $\mathcal{G} = (\{L, J\}, \{\mathcal{A}_{\mathrm{L}}, \mathcal{A}_{\mathrm{J}}(p)\}, \{u_L, u_J\}))$. First the players are identified, *i.e.*, player L representing the legitimate users, who are considered to act as a single player, and player J representing the jammer. Secondly, the action set of player L and player J are determined as $\mathcal{A}_L$ and $\mathcal{A}_J$, respectively. Finally the utility function (payoff) of each player is defined as u_L, for player L, and u_J for player J. A fundamental assumption of the theory is that the players are rational, *i.e.*, each player chooses the action that optimises his utility knowing that the other players will act likewise. The games defined in this chapter belong to the group of zero-sum games. This is a type of game where the gain of one player equals the loss of the other player, *i.e.*, $u_L = -u_J$, hence while maximising their utility function each player minimises the profit of the others.

In a game-theoretic analysis the saddle point of a game is defined by its Nash or Stackelberg equilibria. A Nash equilibrium is the profile of strategies where all player choose simultaneously their best responses to the equilibrium strategies of the other players [280, 281], such that:

$$u_i(a_i^*, a_{-i}^*) \geq u_i(b_i, a_{-i}^*), \tag{5.1}$$

where $a_i^* \in \mathrm{BR}_i$ and $a_{-i}^* \in \mathrm{BR}_{-i}$ define the fixed points of best responses for players i and all other players in the game, respectively, $b_i \in \mathcal{A}_i \cap \mathrm{BR}_i^C$ denotes all other actions that player i can take. On the other hand, Stackelberg games can be distinguished from Nash games by the fact that in Stackelberg games the players act in specific order in time, *i.e.*, not simultaneously as in Nash equilibrium games. Therefore, in Stackelberg games there is a leader who chooses his strategy first and followers who choose their best response based upon the leader's action. Section 5.5.1 of the present chapter evaluates the Stackelberg equilibrium in order to identify the optimal strategy of legitimate users in presence of reactive jammer.

5.3 Employed methods and system model

The previous section summarised the necessary background material, the rest of this chapter explores novel work on the jamming attacks on the SKG process and proposes a set of countermeasures. As discussed in Chapter 4, building semantically secure AE protocols using the SKG procedure is straightforward, as long as the channel probing phase of the scheme is robust against active attacks [108], [282]. Therefore, an important next step is to study MiM and DoS attacks during the channel excitation phase of the SKG protocol, commonly referred to as "advantage distillation". The investigations are based on the methods described below.

Secret key generation

As discussed in the previous chapters, Alice and Bob can obtain a shared secret key following the three-step SKG procedure, (described in Section 2.2.1), *i.e.*, i) advantage distillation; ii) information reconciliation; and, iii) privacy amplification. This chapter provides countermeasures to existing attacks during the process. This is discussed in detail in the next items.

Jamming attacks and countermeasures

Firstly, MiM attacks, referred to as "injection" attacks, are investigated: an active adversary tries to control part of the generated secret key by spoofing the channel estimation phase of the SKG scheme. Existing works have considered jamming attacks and formulate these in game-theoretic form [283], [284]. However, they have not considered the close relationship between injection and jammming. The work in this chapter proposes an approach of the MiM attack that assumes that the adversary has one additional antenna with respect to the legitimate users. This is a generous assumption with respect to the adversary's capabilities and reveals a critical vulnerability of SKG, that needs to be addressed. As a countermeasure, this study proposes a concrete pilot randomisation scheme

using QPSK modulated random pilots. This work proves that the source of shared randomness remains Gaussian and that the adversary can no longer mount the MiM attack. An interesting conclusion of the analysis is that the MiM injection attack is reduced to a jamming attack when pilot randomisation is employed.

Next, motivated by the above result, DoS in the form of reactive jamming is studied for BF-AWGN channels – used as an abstraction for orthogonal frequency division multiplexing (OFDM) modulation systems. The attacker's optimal strategies are derived. In the present contribution it is assumed that the legitimate users blindly adopt a uniform power allocation policy, the level of which is optimally identified. The study demonstrates that a reactive jammer can have a far more serious impact on the SKG process compared to a simple active jammer.

Game theory

The optimal strategies within this chapter are identified through game-theoretic analysis (for more details on game theory please see Section 5.2.3). All of the scenarios, described in the previous item, are formulated as zero-sum games where legitimate parties and adversary are presented as players with opposite goals. The study within this chapter does not introduce any novelty in regards to game theory, instead it uses it as a tool to determine the best actions of all players (*i.e.*, legitimate users and jammer).

The system model, depicted in Fig. 5.2, assumes two legitimate parties, referred to as Alice and Bob, and a active adversary, referred to as Mallory. This work assumes a Rayleigh multipath environment, where the legitimate parties communicate over a BF-AWGN channel, that comprises N subcarriers.

5.4 MiM in SKG Systems: Injection Attacks

MiM in the form of injection attacks constitutes one of the most critical limitations in SKG systems [285–287]. While jamming attacks aim at interrupting a legitimate communication, injection attacks may reveal sensitive information, such as secret key bits. Recently,

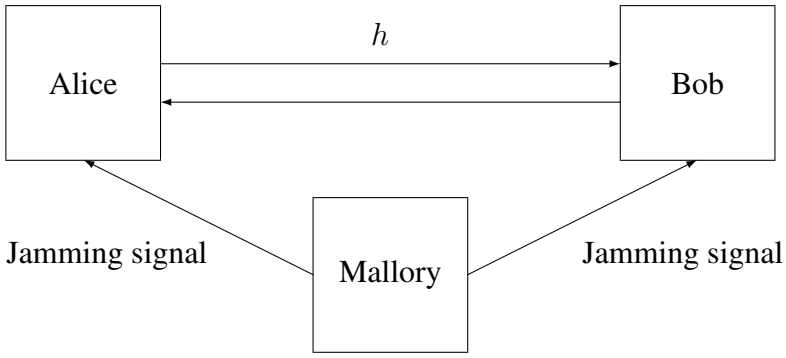

Figure 5.2: Alice and Bob are communicating over a Rayleigh fading channel with realisation h. A MiM, Mallory, act as an active adversary, trying to interrupt their communication.

various possible approaches for injection attacks have been published: in [285], the attacker controlled the movement of objects in an indoor wireless network, thus generating predictable changes in the RSS, (*e.g.*, by obstructing, or not, a line-of-sight). In [286], whenever similar channel envelope measurements in the links to the legitimate nodes were observed, the MiM spoofed the SKG process by injecting a strong signal. The following will prove that – even when full CSI is used to extract the keys – it suffices that the adversary has one additional antenna with respect to the legitimate users to be able to mount an injection MiM attack.

To capture the main components of injection attacks in SKG systems, the system model depicted in Fig. 5.3 is employed. It comprises three nodes: a legitimate transmitter, its intended receiver, and a MiM, referred to as Alice, Bob and Mallory, respectively. Alice and Bob are assumed to have a single antenna each for simplicity, while Mallory has two transmit antennas.[1] The fading channel realisation in the link Alice-Bob is denoted by the complex circularly symmetric Gaussian random variable $\mathbf{h} \sim \mathcal{CN}(0, \sigma^2)$. To obtain estimates of $\mathbf{h}$, Alice and Bob exchange pilot signals $\mathbf{y}$ with $\mathsf{E}[|\mathbf{y}|^2] \leq P$. Furthermore, following [286], it is assumed that Mallory has perfect knowledge of the channel

[1] It is straightforward to see that the scenario can easily be generalised to a multi-antenna setting in which Mallory has one more antenna than Alice and Bob.

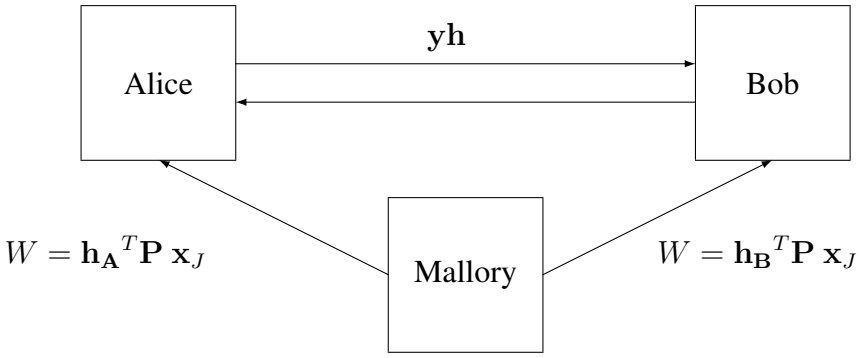

Figure 5.3: Alice and Bob have single transmit and receive antennas and exchange pilot signals $\mathbf{x}$ over a Rayleigh fading channel with realisation $\mathbf{h}$. A MiM, Mallory, with multiple transmit antennas can inject a suitably pre-coded signal $\mathbf{P}\mathbf{x}_J$, such that the received signal at both Alice and Bob coincide $\mathbf{w} = \mathbf{h_A}^T\mathbf{P} = \mathbf{h_B}^T\mathbf{P}$.

vectors in the multiple input single output (MISO) links Mallory-Alice and Mallory-Bob. The channel coefficients of Mallory-Alice and Mallory-Bob are assumed to be independent and identically distributed (i.i.d.), *i.e.*, $\mathbf{h_A} = [h_{A1}, h_{A2}]^T, \mathbf{h_B} = [h_{B1}, h_{B2}]^T$ with $(h_{A1}, h_{A2}, h_{B1}, h_{B2}) \sim \mathcal{CN}(\mathbf{0}, \sigma_J^2/2\ \mathbf{I}_4)$; this assumption is realistic since Mallory can estimate the channel vectors while Alice and Bob exchange pilot signals, as long as the channel's coherence time is respected (a plausible scenario in slow fading, low mobility environments). The latter is a best case scenario for Mallory.

To mount the attack, Mallory transmits a signal $\mathbf{x}_J$, suitably precoded as $\mathbf{P}\mathbf{x}_J$. The precoding matrix $\mathbf{P} = [P_1, P_2]^T$ is chosen such that the same signal is "injected" at both Alice and Bob, *i.e.*,

$$\begin{aligned} \mathbf{h_A}^T\mathbf{P}\mathbf{x}_J = \mathbf{h_B}^T\mathbf{P}\mathbf{x}_J \Rightarrow \\ P_1 = \frac{h_{B2} - h_{A2}}{h_{A1} - h_{B1}} P_2, \end{aligned} \tag{5.2}$$

where, due to the i.i.d. assumption and to the continuous distribution of the channels, $h_{A1} \neq h_{B1}$ almost surely. As a result, Mallory can select a suitable precoding matrix (among infinite possibilities). In practice it will be difficult for Mallory to determine the

correct precoding matrix, but here we assume that Mallory was able to do this. Assuming a total power constraint $\mathbb{E}[|\mathbf{P}\mathbf{x}_J|^2] \leq \Gamma$ for Mallory's transmission, P_2 should be chosen as

$$P_2 \leq \frac{\sqrt{\Gamma}}{\left|\frac{h_{B2}-h_{A2}}{h_{A1}-h_{B1}}\right|+1}. \tag{5.3}$$

This procedure, illustrated in Fig. 5.3, shows that it is possible to generalise the injection attack presented in [286], in which an attacker injected a strong signal whenever the RSS in the Mallory-Alice and Mallory-Bob links were similar. More importantly, the presented injection attack accounts not only for the RSS but for the full CSI, *i.e.*, it includes the signal phase.

The observations at Alice and Bob, denoted by $\mathbf{x}_A$ and $\mathbf{x}_B$, are

$$\mathbf{x}_A = \mathbf{y}\mathbf{h} + \mathbf{w} + \mathbf{z}_A \tag{5.4}$$

$$\mathbf{x}_B = \mathbf{y}\mathbf{h} + \mathbf{w} + \mathbf{z}_B, \tag{5.5}$$

where $\mathbf{w} = \mathbf{h_A}^T\mathbf{P}\mathbf{x}_J = \mathbf{h_B}^T\mathbf{P}\mathbf{x}_J$ denotes the observed injected signal at Alice and Bob which is identical at both due the precoding matrix $\mathbf{P}$; and, $\mathbf{z}_A, \mathbf{z}_B$ denote zero-mean unit variance i.i.d. complex circularly symmetric Gaussian random noise variables, *i.e.*, $\mathbf{z}_A, \mathbf{z}_B \sim \mathcal{CN}(0,1)$. The secret key rate controlled by Mallory is upper bounded by [282]

$$L \leq I(\mathbf{x}_A, \mathbf{x}_B; \mathbf{w}). \tag{5.6}$$

Identifying the optimal injection signal $\mathbf{w}$, corresponds to finding the capacity achieving input signal of the *two-look Gaussian channel* in (5.4)-(5.5). This signal is known to be Gaussian [288]; hence, a good choice for $\mathbf{x}_J$ is to be constant, so that, the overall injected signal is an optimal complex zero-mean circularly symmetric Gaussian signal, $\mathbf{w} \sim \mathcal{CN}(0, \sigma_J^2\Gamma)$.

A countermeasure to injection attacks can be built by randomising the pilot sequence exchanged between Alice and Bob [282], [287]. This study, proposes to randomise the pilots by drawing them from a (scaled) QPSK modulation, as follows: instead of trans-

mitting the same probing signal $\mathbf{y}$, Alice and Bob transmit independent, random probe signals $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively, drawn from i.i.d. zero-mean discrete uniform distributions $\mathcal{U}(\{\pm r \pm jr\})$, where $j = \sqrt{-1}, r = \sqrt{P/2}$, so that, $\mathsf{E}[\mathbf{y}_1] = \mathsf{E}[\mathbf{y}_2] = 0$, $\mathsf{E}[|\mathbf{y}_1|^2] = \mathsf{E}[|\mathbf{y}_2|^2] = P$ and $\mathsf{E}[\mathbf{y}_1\mathbf{y}_2] = 0$, *i.e.*, the pilots are randomly chosen QPSK signals. Alice's and Bob's observations $\mathbf{x}_A, \mathbf{x}_B$, respectively, are modified accordingly as

$$\mathbf{x}_A = \mathbf{y}_1\mathbf{h} + \mathbf{w} + \mathbf{z}_A, \tag{5.7}$$

$$\mathbf{x}_B = \mathbf{y}_2\mathbf{h} + \mathbf{w} + \mathbf{z}_A. \tag{5.8}$$

To establish shared randomness in spite of the pilot randomisation, Alice and Bob post-multiply $\mathbf{x}_A$ and $\mathbf{x}_B$ by their randomised pilots, obtaining local observations $\tilde{\mathbf{x}}_A$ and $\tilde{\mathbf{x}}_B$ (unobservable by Mallory), expressed as:

$$\tilde{\mathbf{x}}_A = \mathbf{y}_1\mathbf{x}_A = \mathbf{y}_1\mathbf{y}_2\mathbf{h} + \mathbf{y}_1\mathbf{w} + \mathbf{y}_1\mathbf{z}_A, \tag{5.9}$$

$$\tilde{\mathbf{x}}_B = \mathbf{y}_2\mathbf{x}_B = \mathbf{y}_1\mathbf{y}_2\mathbf{h} + \mathbf{y}_2\mathbf{w} + \mathbf{y}_2\mathbf{z}_B. \tag{5.10}$$

Lemma 5.1. *The source of shared randomness, when the pilots are randomised QPSK symbols, is a circularly symmetric zero mean Gaussian random variable, following the distribution* $\mathbf{y}_1\mathbf{y}_2\mathbf{h} \sim C\mathcal{N}(0, P^2\sigma^2)$.

Proof. The two orthogonal axes (real and imaginary) are treated independently. Looking only at the real values of the pilots and of the channel coefficient $\mathbf{y}_1, \mathbf{y}_2, \mathbf{h}$ denoted here by $\mathbf{y}_{1,\mathbb{R}} = \mathrm{Re}(\mathbf{y}_1)$, $\mathbf{y}_{2,\mathbb{R}} = \mathrm{Re}(\mathbf{y}_2)$ and $\mathbf{h}_{\mathbb{R}} = \mathrm{Re}(\mathbf{h})$, the underlying discrete uniform pdf expressed as $f_{\mathbf{y}_{1,\mathbb{R}}}(\mathbf{y}_1)$ and $f_{\mathbf{y}_{2,\mathbb{R}}}(\mathbf{y}_2)$ and the continuous pdf $f_{\mathbf{h}_{\mathbb{R}}}(\mathbf{h})$ as

$$f_{\mathbf{y}_{1,\mathbb{R}}}(\mathbf{y}_1) = \frac{1}{2}\delta(\mathbf{y}_1 - r) + \frac{1}{2}\delta(\mathbf{y}_1 + r), \tag{5.11}$$

$$f_{\mathbf{y}_{2,\mathbb{R}}}(\mathbf{y}_2) = \frac{1}{2}\delta(\mathbf{y}_2 - r) + \frac{1}{2}\delta(\mathbf{y}_2 + r), \tag{5.12}$$

$$f_{\mathbf{h}_{\mathbb{R}}}(\mathbf{h}) = \frac{1}{\sqrt{\pi}\sigma}e^{-\frac{\mathbf{h}^2}{\sigma^2}}. \tag{5.13}$$

The pdf of the product $\mathbf{y}_{1,\mathbb{R}}\mathbf{h}_{\mathbb{R}}$ is given as

$$\begin{aligned} f_{\mathbf{y}_{1,\mathbb{R}}\mathbf{h}_{\mathbb{R}}}(\mathbf{x}) &= \int_{-\infty}^{\infty} f_{\mathbf{y}_{1,\mathbb{R}}}(\mathbf{y}_1) f_{\mathbf{h}_{\mathbb{R}}}(\mathbf{x}/\mathbf{y}_1) \frac{1}{|\mathbf{y}_{1,}|} \mathrm{d}\mathbf{y}_1 \\ &= \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\pi}\sigma|\mathbf{y}_1|} \delta(\mathbf{y}_1 - r) e^{-\frac{(\mathbf{x}/\mathbf{y}_1)^2}{\sigma^2}} \mathrm{d}\mathbf{y}_1 \\ &+ \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\pi}\sigma|\mathbf{y}_1|} \delta(\mathbf{y}_1 + r) e^{\frac{-(\mathbf{x}/\mathbf{y}_1)^2}{\sigma^2}} \mathrm{d}\mathbf{y}_1 \\ &= \frac{\sqrt{2} e^{-\frac{2\mathbf{x}^2}{P\sigma^2}}}{\sqrt{\pi P}\sigma} \end{aligned} \tag{5.14}$$

by substituting $r = \sqrt{P/2}$ at the last derivation, *i.e.*, $\mathbf{y}_{1,\mathbb{R}}\mathbf{h}_{\mathbb{R}} \sim \mathcal{N}(0, \frac{P\sigma^2}{4})$. A similar result holds for the products involving also the imaginary parts of $\mathbf{y}_1$ and $\mathbf{h}$: $\mathbf{y}_{1,\mathbb{I}}\mathbf{h}_{\mathbb{I}}$, $\mathbf{y}_{1,\mathbb{I}}\mathbf{h}_{\mathbb{R}}$ and $\mathbf{y}_{1,\mathbb{R}}\mathbf{h}_{\mathbb{I}}$, so that $\mathbf{y}_1\mathbf{h} \sim \mathcal{CN}(0, P\sigma^2)$. Extending this result, it can be fount that $\mathbf{y}_1\mathbf{y}_2\mathbf{h} \sim \mathcal{CN}(0, P^2\sigma^2)$. ■

Furthermore, due to the fact that $\mathbf{y}_1$ and $\mathbf{y}_2$ are independent and have zero mean, the variables $\mathbf{y}_1 W$ and $\mathbf{y}_2 W$ are uncorrelated, circularly symmetric zero-mean Gaussian random variables, and, therefore independent, while the same holds for $\mathbf{y}_1\mathbf{z}_A, \mathbf{y}_2\mathbf{z}_B$, *i.e.*, $(\mathbf{y}_1 W, \mathbf{y}_2 W) \sim \mathcal{CN}(\mathbf{0}, \sigma_J^2 P\Gamma\mathbf{I}_2)$ and $(\mathbf{y}_1\mathbf{z}_A, \mathbf{y}_2\mathbf{z}_B) \sim \mathcal{CN}(\mathbf{0}, P\mathbf{I}_2)$. Alice and Bob extract the common key from the modified source of common randomness $\mathbf{y}_1\mathbf{y}_2\mathbf{h}$ as opposed to $\mathbf{y}\mathbf{h}$. On the other hand, since $\mathbf{y}_1\mathbf{w}, \mathbf{y}_2\mathbf{w}, \mathbf{y}_1\mathbf{z}_A, \mathbf{y}_2\mathbf{z}_B$ are i.i.d. complex circularly symmetric Gaussian random variables, the proposed scheme reduces injection attacks to uncorrelated jamming attacks, *i.e.*, using Lemma 5.1 follows:

$$L \leq I\left(\tilde{\mathbf{x}}_A, \tilde{\mathbf{x}}_B; \mathbf{w}\right) = 0. \tag{5.15}$$

Employing the proposed, within this section, mechanism could transform an injection attacks to a jamming attacks. As discussed earlier, MiM in the form of injection attack could be far more severe than jamming attacks, therefore, the presented solution is of great importance for the systems security.

5.5 Jamming Attacks on SKG

Building on the results of the previous section, this section examines in detail the scenario in which Mallory acts as a reactive jammer. Reactive jamming is a stealthy jamming approach in which the jammer first senses the spectrum and jams only when she detects an ongoing transmission. Due to the effectiveness and difficulty to be detected, reactive jammers are considered as the most harmful [265,266]. In this context, this work assumes that Alice and Bob perform SKG over N subcarriers. The notation introduced in Section 5.4 is extended with the introduction of a carrier index $i \in \{1, \ldots, N\}$, *i.e.*, $y_{1,i}, y_{2,i}$ denote the randomised pilots on the i-th subcarrier, h_i denotes the channel coefficient in the link Alice-Bob, W_i the signal injected by Mallory on the i-th subcarrier and $z_{A,i}, z_{B,i}$ noise variables. As a reactive jammer, Mallory senses the spectrum and jams a specific subcarrier only when the power on it exceeds a certain threshold p_{th}. Note that, Alice and Bob can get an estimate of p_{th} using the methods described in Sec. 5.2.2, in particular the jamming detector schemes presented in [265] and [271] can be employed: the former uses a pre-stored sequence of non-jammed bits, which are later compared to the received signals; the latter is based on estimating the variance of received pilot signals. Therefore, before the SKG process Alice and Bob can use any of these schemes to estimate the detection threshold p_{th} of the jammer Malory. For further details regarding these schemes please see to the corresponding reference. For the purpose of this study, two scenarios are considered: i) when p_{th} is fixed (determined in essence by the carrier sensing capability of Mallory's receiver); ii) when p_{th} is variable (its choice forms part of her strategy).

The expressions of Alice's and Bob's local observations on the i-th SKG subcarrier are reformulated as follows:

$$\tilde{x}_{A,i} = y_{1,i}y_{2,i}h_i + y_{1,i}w_i + y_{1,i}z_{A,i} \tag{5.16}$$

$$\tilde{x}_{B,i} = y_{1,i}y_{2,i}h_i + y_{2,i}w_i + y_{2,i}z_{B,i} \tag{5.17}$$

for $i = 1, \ldots, N$ with $h_i \sim \mathcal{CN}(0, \sigma^2)$, $w_i \sim \mathcal{CN}(0, \sigma_J^2\gamma_i)$, $z_{A,i} \sim \mathcal{CN}(0, 1)$, $z_{B,i} \sim$

$\mathcal{CN}(0,1)$. This work assumes that Alice and Bob use the same power p on all pilots, in agreement with common practice during the advantage distillation phase. Based on this assumption it holds that $\mathbb{E}[|y_{1,i}|^2] = \mathbb{E}[|y_{2,i}|^2] = p$ with $p \in [0, P]$.

On the other hand, Mallory chooses the power allocation vector to maximise the impact of her attack. The power Mallory uses on the i-th subcarrier is denoted by γ_i, so that $\mathbb{E}[|W_i|^2] = \sigma_J^2 \gamma_i$. Denoting the average available power for jamming by Γ and the power allocation of the jammer by $\underline{\boldsymbol{\gamma}} = (\gamma_1, \ldots, \gamma_N)$, the following short-term power constraint is assumed:

$$\underline{\boldsymbol{\gamma}} \in \mathbb{R}_+^N, \quad \sum_{i=1}^{N} \gamma_i \leq N\Gamma. \tag{5.18}$$

Assuming that h_i is uncorrelated with $h_{A,i}, h_{B,i}$, $i = 1, \ldots, N$ and that the pilot randomisation approach proposed in Section 5.4 is employed, the SKG rate $R_{SKG}(p, \gamma_i) = I\left(\tilde{x}_{A,i}; \tilde{x}_{B,i}\right)$ on the i-th subcarrier, can be expressed as a function of p and $\gamma_i, i = 1, \ldots, N$ as [289]:

$$R_{SKG}(p, \gamma_i) = \log_2 \left(1 + \frac{p\sigma^2}{2(1 + \gamma_i \sigma_J^2) + \frac{(1+\gamma_i \sigma_J^2)^2}{p\sigma^2}} \right). \tag{5.19}$$

Note that the rate in (5.19) is independent of the instantaneous realisations of the fading coefficients; instead, the variations of the channel gains expressed through the variances σ^2, σ_J^2 determine the rate of the secret keys that can be extracted from the wireless medium. The overall SKG sum-rate can then be simply expressed as follows:

$$C_{SKG}(p, \underline{\boldsymbol{\gamma}}) = \sum_{i=1}^{N} R_{SKG}(p, \gamma_i). \tag{5.20}$$

5.5.1 Optimal Power Allocation Strategies

Alice and Bob's common objective is to maximise $C_{SKG}(p, \underline{\boldsymbol{\gamma}})$ with respect to (w.r.t.) p, while Mallory wants to minimise $C_{SKG}(p, \underline{\boldsymbol{\gamma}})$ w.r.t. $\underline{\boldsymbol{\gamma}}$. Given the opposed objectives, a non-cooperative zero-sum game can be formulated to study the strategic interaction

between the legitimate users and the jammer: $\mathcal{G} = (\{L, J\}, \{\mathcal{A}_{\text{L}}, \mathcal{A}_{\text{J}}(p)\}, C_{SKG}(p, \underline{\boldsymbol{\gamma}}))$. The game $\mathcal{G}$ has three components. Firstly, there are two players: player L representing the legitimate users (Alice and Bob are considered to act as a single player) and player J representing the jammer (Mallory). Secondly, player L has a set of possible actions $\mathcal{A}_L = [0, P]$ while player J's set of actions is

$$\mathcal{A}_J(p) = \begin{cases} \{(0, \ldots, 0)\}, & \text{if } p \leq p_{\text{th}}, \\ \left\{\underline{\boldsymbol{\gamma}} \in \mathbb{R}_+^N | \sum_{i=1}^N \gamma_i \leq N\Gamma\right\}, & \text{if } p > p_{\text{th}}, \end{cases} \tag{5.21}$$

where p_{th} is maximum transmission power from either Alice or Bob that is still undetectable by Mallory. Finally, $C_{SKG}(p, \underline{\boldsymbol{\gamma}})$, denotes the payoff function of player L.

Due to the fact that Mallory first observes the transmit power of the legitimate users on the subcarriers and then decides which strategy to choose (a consequence of player J being a reactive jammer), this section studies a hierarchical game in which player L is the leader and player J is the follower. In this hierarchical game, the solution is the Stackelberg equilibrium (SE) – rather than Nash – defined as a strategy profile $(p^{\text{SE}}, \underline{\boldsymbol{\gamma}}^{\text{SE}})$ where player L chooses their optimal strategy first, by anticipating the strategic reaction of player J (*i.e.*, its best response). This can be rigorously written as:

$$p^{\text{SE}} \triangleq \arg\max_{p \in \mathcal{A}_L} \sum_{i=1}^N R_{SKG}(p, \underline{\boldsymbol{\gamma}}^*(p)), \text{ and } \underline{\boldsymbol{\gamma}}^{\text{SE}} \triangleq \underline{\boldsymbol{\gamma}}^*(p^{\text{SE}}), \tag{5.22}$$

where $\underline{\boldsymbol{\gamma}}^*(p)$ denotes the jammer's best response (BR) function to any strategy $p \in \mathcal{A}_L$ chosen by player L, defined as follows:

$$\underline{\boldsymbol{\gamma}}^*(p) \triangleq \arg\min_{\underline{\boldsymbol{\gamma}} \in \mathcal{A}_J(p)} \sum_{i=1}^N R_{SKG}(p, \underline{\boldsymbol{\gamma}}). \tag{5.23}$$

Note that $\gamma_i^*(p)$ is the i-th subcarrier of $\underline{\boldsymbol{\gamma}}^*(p)$.

Stackelberg equilibrium with fixed p_{th}

In the following, the SE of the game $\mathcal{G}$ is evaluated assuming that the threshold p_{th} is predefined and fixed. The case $P \leq p_{\text{th}}$ is trivial as $\underline{\gamma}^{\text{SE}} = (0, \ldots, 0)$, whereas, the legitimate users will optimally use the maximum available power so that $(p^{\text{SE}} = P)$. Indeed, because of the badly chosen threshold or low sensing capabilities of Mallory, the legitimate transmission will never be detected on any of the subcarriers and hence will not be jammed. In the following, it is assumed that: $P > p_{\text{th}}$.

Lemma 5.2. *The BR of the jammer for any $p \in \mathcal{A}_L$ chosen by the leader defined in (5.23) is the uniform power allocation, such that:*

$$\underline{\gamma}^*(p) \triangleq \begin{cases} (\Gamma, \ldots, \Gamma), & \textit{if } p > p_{th}, \\ (0, \ldots, 0), & \textit{if } p \leq p_{th}. \end{cases} \tag{5.24}$$

Proof. Note that $R_{SKG}(p, \gamma_i)$ is a monotonically decreasing convex function w.r.t γ_i, $i = 1, \ldots, N$ for any $p > 0$. First this proof shows that the jamming power should be equally distributed on all of the subcarriers. First, Jensen's inequality is applied using $\delta_i > 0$, $\sum_{i=1}^{N} \delta_i = 1$, so that $R_{SKG}\left(p, \sum_{i=1}^{N} \delta_i x_i\right) \leq \sum_{i=1}^{N} \delta_i R_{SKG}(p, x_i)$. Substituting $\delta_i = 1/N$, $x_i = \Gamma/b_i$, we get:

$$\begin{aligned} &R_{SKG}\left(p, \sum_{i=1}^{N} \frac{\Gamma}{N b_i}\right) \leq \sum_{i=1}^{N} \frac{1}{N} R_{SKG}\left(p, \frac{\Gamma}{b_i}\right) \Rightarrow \\ &N R_{SKG}\left(p, \frac{1}{N} \sum_{i=1}^{N} \frac{\Gamma}{b_i}\right) \leq \sum_{i=1}^{N} R_{SKG}\left(p, \frac{\Gamma}{b_i}\right). \end{aligned} \tag{5.25}$$

Applying the power constraint $\sum_{i=1}^{N} \Gamma/b_i \leq N\Gamma$ on the LHS of (5.25), for any $p > p_{\text{th}}$ we have:

$$N R_{SKG}(p, \Gamma) < \sum_{i=1}^{N} R_{SKG}\left(p, \frac{\Gamma}{b_i}\right) \Rightarrow \quad C_{SKG}(p, (\Gamma, \ldots, \Gamma)) \leq C_{SKG}(p, \underline{\gamma}),$$

which shows that in order to minimise C_{SKG}, Mallory has to distribute her power equally on all subcarriers. ■

Note that, the results in Lemma 5.2 define the available actions to Mallory: i) if Alice and Bob transmit with power greater than the sensing threshold p_{th}, Mallory will jam over all subcarriers using equal power distribution; ii) if Alice and Bob transmit with power less than the sensing threshold p_{th}, Mallory will not jam on any of the subcarriers, *i.e.*, the transmission is not detected and Mallory will stay silent.

In light of this result, the SKG sum rate can have two forms:

$$C_{SKG}(p, \underline{\boldsymbol{\gamma}}^*(p)) = \begin{cases} NR_{SKG}(p, \Gamma), & \text{if } p > p_{\text{th}}, \\ \\ NR_{SKG}(p, 0), & \text{if } p \leq p_{\text{th}}, \end{cases} \tag{5.26}$$

which simplifies the players' options. Next, this work addresses the question of how Alice and Bob should choose their power p optimally by considering the actions available to the players in the game at the key points, *i.e.*, at P and p_{th}.

Theorem 5.1. *Depending on the available power P for SKG, player L will either transmit at P or p_{th} on all subcarriers. The SE point of the game is unique when $P \neq p_{th}(\sigma_J^2\Gamma + 1)$ and is given by*

$$(p^{SE}, \underline{\boldsymbol{\gamma}}^{SE}) = \begin{cases} \{(p_{th}, (0, \ldots, 0))\}, \textit{if } P < p_{th}(\sigma_J^2\Gamma + 1), \\ \\ \{(P, (\Gamma, \ldots, \Gamma))\}, \textit{ if } P > p_{th}(\sigma_J^2\Gamma + 1). \end{cases} \tag{5.27}$$

When $P = p_{th}(\sigma_J^2\Gamma + 1)$, the game $\mathcal{G}$ has two SEs, which are expressed as follows: $(p^{SE}, \underline{\boldsymbol{\gamma}}^{SE}) \in \{(p_{th}, (0, \ldots, 0)), (P, (\Gamma, \ldots, \Gamma))\}$.

Proof. Given the BR in (5.24) and the simplification in (5.26), player L wants to find the

optimal $p \in \mathcal{A}_L$ that maximises:

$$R_{SKG}(p, \gamma_i^*(p)) = \begin{cases} R_{SKG}(p, 0), & \text{if } p \leq p_{\text{th}}, \\ R_{SKG}(p, \Gamma), & \text{if } p > p_{\text{th}}. \end{cases} \tag{5.28}$$

Given that $R_{SKG}(p, \gamma)$ is monotonically increasing with p for fixed γ, two cases are distinguished: a) $p \in [0, p_{\text{th}}]$, b) $p \in (p_{\text{th}}, P]$. The optimal p in each case is given by

a) $\underset{p \in [0, p_{\text{th}}]}{\arg\max}\, R_{SKG}(p, \gamma_i^*(p)) = \underset{p \in [0, p_{\text{th}}]}{\arg\max}\, R_{SKG}(p, 0) = p_{\text{th}}$,

b) $\underset{p \in (p_{\text{th}}, P]}{\arg\max}\, R_{SKG}(p, \gamma_i^*(p)) = \underset{p \in (p_{\text{th}}, P]}{\arg\max}\, R_{SKG}(p, \Gamma) = P$.

From a) and b), it can be concluded that the overall solution is $p^{\text{SE}} =$

$$\underset{p \in \mathcal{A}_L}{\arg\max}\, R_{SKG}(p, \gamma_i^*(p)) = \begin{cases} p_{\text{th}}, & \text{if } R_{SKG}(P, \Gamma) < R_{SKG}(p_{\text{th}}, 0), \\ P, & \text{if } R_{SKG}(P, \Gamma) > R_{SKG}(p_{\text{th}}, 0), \\ \{p_{\text{th}}, P\}, & \text{if } R_{SKG}(P, \Gamma) = R_{SKG}(p_{\text{th}}, 0). \end{cases}$$

The three possibilities are simplified using the case when transmitting at full power $R_{SKG}(P, \Gamma)$ (hence being sensed and jammed) is equal to the case when player L is transmitting at threshold p_{th} (the jammer is silent) *i.e.*, $R_{SKG}(P, \Gamma) = R_{SKG}(p_{\text{th}}, 0)$. Using this equality, and by substituting appropriately into (5.19), it can be obtained as a quadratic equation in P:

$$P^2(2\sigma^2 p_{\text{th}} + 1) - P(2{p_{\text{th}}}^2\sigma^2 + 2\sigma_J^2\Gamma{p_{\text{th}}}^2\sigma^2) - (1 + \sigma_J^2\Gamma)^2{p_{\text{th}}}^2 = 0,$$

which has a unique positive root equal to $p_{\text{th}}(\sigma_J^2\Gamma + 1)$. Given that, the leading coef-

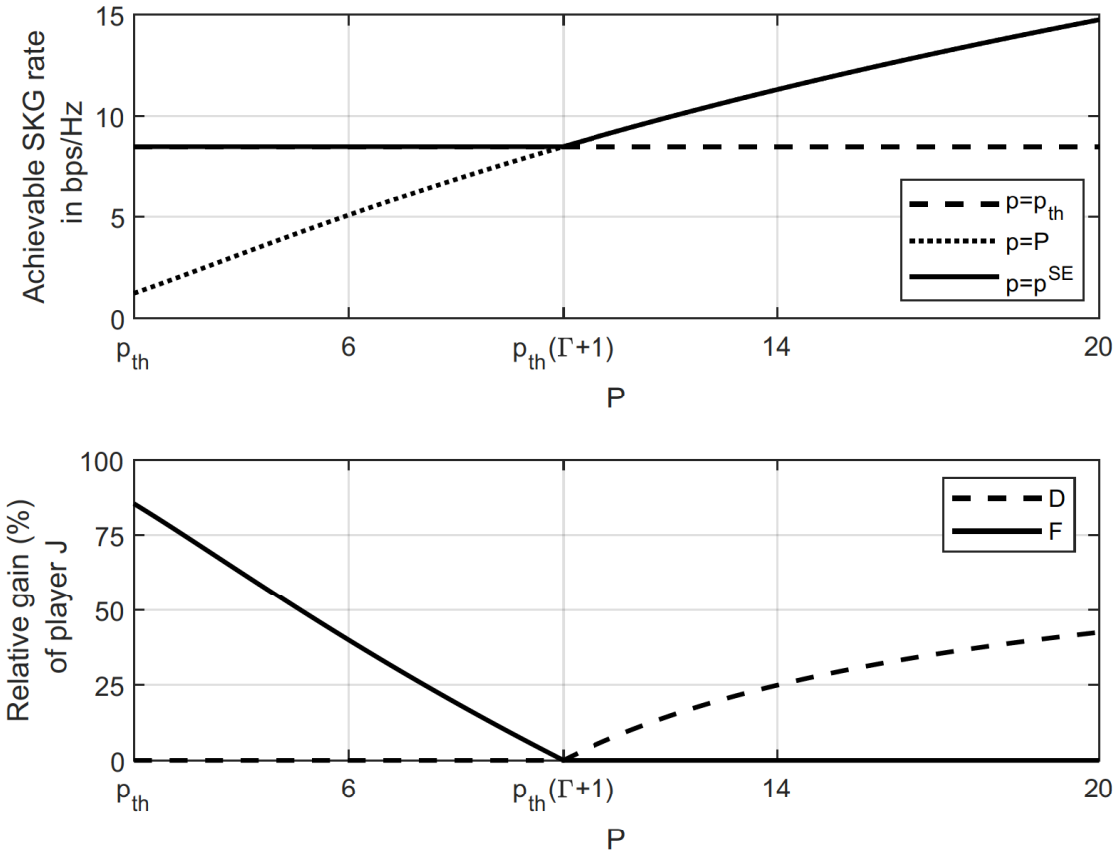

Figure 5.4: UP: SE policy compared to always transmitting with either full power or with p_{th}. DOWN: Functions D and F vs P. In both sub-figures, $p_{\text{th}} = 2, \Gamma = 4, N = 10, \sigma^2 = \sigma_J^2 = 1$.

ficient of (5.29): $(2\sigma^2 p_{\text{th}} + 1) \geq 0$ and that $P > 0$, it can be said that the inequalities $R_{SKG}(P, \Gamma) > R_{SKG}(p_{\text{th}}, 0)$ and $R_{SKG}(P, \Gamma) < R_{SKG}(p_{\text{th}}, 0)$ are equivalent to $P > p_{\text{th}}(\sigma_J^2 \Gamma + 1)$ and $P < p_{\text{th}}(\sigma_J^2 \Gamma + 1)$, respectively. ■

Some numerical results are presented in Fig. 5.4 for a total number of SKG subcarriers $N = 10$ (pertinent to narrowband IoT applications), $p_{\text{th}} = 2$, $\Gamma = 4$, and $\sigma^2 = \sigma_J^2 = 1$. The top figure compares the achievable rates of the SE strategy and of two alternative strategies consisting in transmitting with fixed $p = P$ or $p = p_{\text{th}}$. The bottom figure depicts the following quantities:

$$F = \frac{C_{SKG}(p^{\text{SE}}, \underline{\gamma}^{\text{SE}}) - C_{SKG}(P, (\Gamma, \ldots, \Gamma))}{C_{SKG}^{SE}}, \tag{5.29}$$

$$D = \frac{C_{SKG}(p^{\text{SE}}, \underline{\boldsymbol{\gamma}}^{\text{SE}}) - C_{SKG}(p_{\text{th}}, (0, \ldots, 0))}{C_{SKG}^{SE}}, \tag{5.30}$$

where F and D represent the jammer's gain (or legitimate users' loss) if player L deviates from the SE point (indeed, if player L transmits at $P > p_{\text{th}}$, the jammer will jam at $\gamma_i^*(P) = \Gamma$; and if player L transmits at p_{th} the jammer will not detect it and will remain silent). Both figures show that the SE is the optimal solution for Alice and Bob. Furthermore, deviating from the SE point can decrease their achievable sum-rates by up to 85%.

Stackelberg equilibrium with strategic p_{th}

Finally, this study investigates how Mallory could optimally adjust p_{th} and how her choice will impact Alice's and Bob's strategies. Allowing p_{th} to vary modifies the game under study as follows $\hat{\mathcal{G}} = (\{L, J\}, \{\mathcal{A}_L, \hat{\mathcal{A}}_J(p)\}, C_{SKG}(p, \underline{\boldsymbol{\gamma}}, p_{\text{th}}))$, where:

$$\hat{\mathcal{A}}_J(p) \triangleq \begin{cases} \{((0, \ldots, 0), p_{\text{th}}),\ p_{\text{th}} \geq 0\}, & \text{if } p_{\text{th}} \geq p, \\ \left\{(\underline{\boldsymbol{\gamma}}, p_{\text{th}}) \in \mathbb{R}_+^{N+1} \mid \sum_{i=1}^{N} \gamma_i \leq N\Gamma\right\}, & \text{if } p_{\text{th}} < p. \end{cases} \tag{5.31}$$

The BR of jammer can then be defined as:

$$(\widehat{\underline{\boldsymbol{\gamma}}}^*(p), \widehat{p_{\text{th}}}^*(p)) \triangleq \underset{(\underline{\boldsymbol{\gamma}}, p_{\text{th}}) \in \hat{\mathcal{A}}_J(p)}{\arg\min}\ C_{SKG}(p, \underline{\boldsymbol{\gamma}}, p_{\text{th}}). \tag{5.32}$$

Lemma 5.3. *The BR of player J in this case is a set of strategies:*

$$(\widehat{\underline{\boldsymbol{\gamma}}}^*(p), \widehat{p_{th}}^*(p)) \in \{\ ((\Gamma, \ldots, \Gamma), \epsilon),\ \epsilon \in [0, p)\}. \tag{5.33}$$

Proof. The problem that the jammer wants to solve is: $\min_{(\underline{\boldsymbol{\gamma}}, p_{\text{th}}) \in \hat{\mathcal{A}}_J(p)} C_{SKG}(p, \underline{\boldsymbol{\gamma}}, p_{\text{th}})$, which

can be split as follows:

$$\min_{p_{\text{th}} \geq 0} \min_{\underline{\gamma} \in \hat{\mathcal{A}}_J(p)} C_{SKG}(p, \underline{\gamma}(p), p_{\text{th}}). \tag{5.34}$$

The solution of the inner minimisation is already known from (5.24). For the outer problem the optimal $p_{\text{th}} \geq 0$ that minimises $C_{SKG}(p, \widehat{\underline{\gamma}}^*(p), p_{\text{th}})$ has to be found. Given that:

$$\min_{p_{\text{th}} \geq 0} C_{SKG}(p, \widehat{\underline{\gamma}}^*(p), p_{\text{th}}) = \begin{cases} NR_{SKG}(p, \Gamma, p_{\text{th}}), & \text{if } p_{\text{th}} < p, \\ NR_{SKG}(p, 0, p_{\text{th}}), & \text{if } p_{\text{th}} \geq p, \end{cases} \tag{5.35}$$

and that $R_{SKG}(p, \Gamma, p_{\text{th}}) < R_{SKG}(p, 0, p_{\text{th}})$ the jammer can optimally choose any threshold such that $p_{\text{th}} = \epsilon, \quad \forall \epsilon < p.$ meaning, any ongoing transmission is sensed and jammed. ■

Having considered the BR of Mallory, we now turn to the optimal strategy for Alice and Bob considering the response of Mallory.

Theorem 5.2. *The game $\hat{\mathcal{G}}$ has an infinite number of SEs:*

$$(\widehat{p}^{SE}, \widehat{\underline{\gamma}}^{SE}, \widehat{p_{th}}^{SE}) \in \{ (P, (\Gamma, \ldots, \Gamma), \epsilon), \ \forall \epsilon < P \}. \tag{5.36}$$

Proof. Given the BR of player J, this proof evaluates the SE of the game $\hat{\mathcal{G}}$. The definition for $\widehat{p}^{\text{SE}}$ is given as:

$$\widehat{p}^{\text{SE}} \triangleq \arg\max_{p \in \mathcal{A}_L} C_{SKG}(p, \widehat{\underline{\gamma}}^*(p), \widehat{p_{\text{th}}}(p)^*). \tag{5.37}$$

Since the jammer will act as in (5.33), follows that:

$$C_{SKG}(p, \widehat{\underline{\gamma}}^*(p), \widehat{p_{\text{th}}}(p)^*) = NR_{SKG}(p, \Gamma, \epsilon), \ \forall \epsilon < p, \tag{5.38}$$

and the fact that $R_{SKG}(p, \Gamma, \epsilon)$ is monotonically increasing with p results in $\widehat{p}^{\text{SE}} = P$. ■

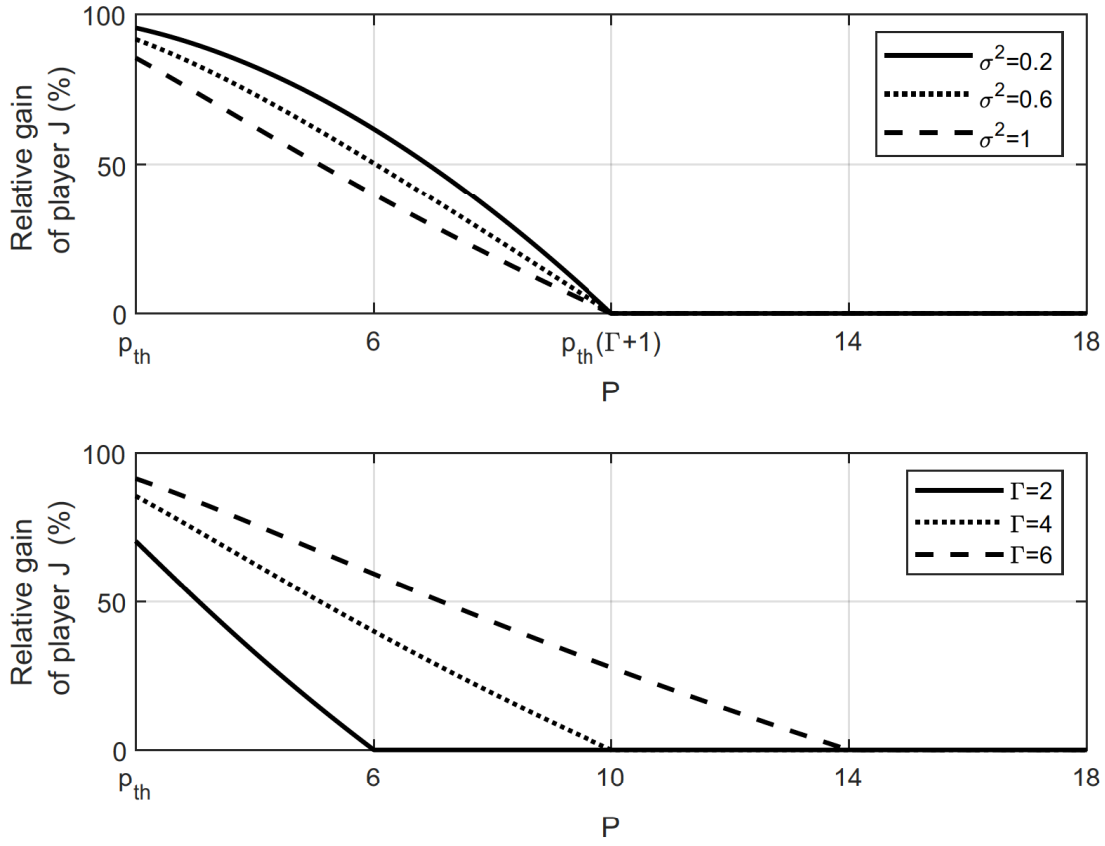

Figure 5.5: Relative gain of player J, evaluated by function E, for strategic p_{th} and fixed $p_{\text{th}} = 2$ when $N = 10$, $\sigma_J^2 = 1$ and UP: $\Gamma = 4$, DOWN: $\sigma^2 = 1$.

Fig. 5.5 and Fig. 5.6 illustrate the gain of the jammer (or the loss in SKG rate) when p_{th} is part of her strategy, with utility function $C_{SKG}(p, \underline{\gamma}, p_{\text{th}})$, compared to the case when it is not, with utility function $C_{SKG}(p, \underline{\gamma})$. This gain is evaluated by:

$$E = \frac{C_{SKG}(p^{SE}, \underline{\gamma}^{SE}) - C_{SKG}(\hat{p}^{SE}, \widehat{\underline{\gamma}}^{SE}, \widehat{p_{\text{th}}}^{SE})}{C_{SKG}(p^{SE}, \underline{\gamma}^{SE})}. \tag{5.39}$$

As in Fig. 5.4 the total number of subcarriers is $N = 10$ and $\sigma_J^2 = 1$. The non-strategic threshold on Fig. 5.5 is set to $p_{\text{th}} = 2$ and the quantity E is evaluated for different values of σ^2 and Γ. The numerical results demonstrate that when p_{th} is part of Mallory's strategy, she can be a significantly more effective opponent, compared to the case when p_{th} is fixed, confirming that reactive jammers can indeed pose a serious threat. This is also confirmed

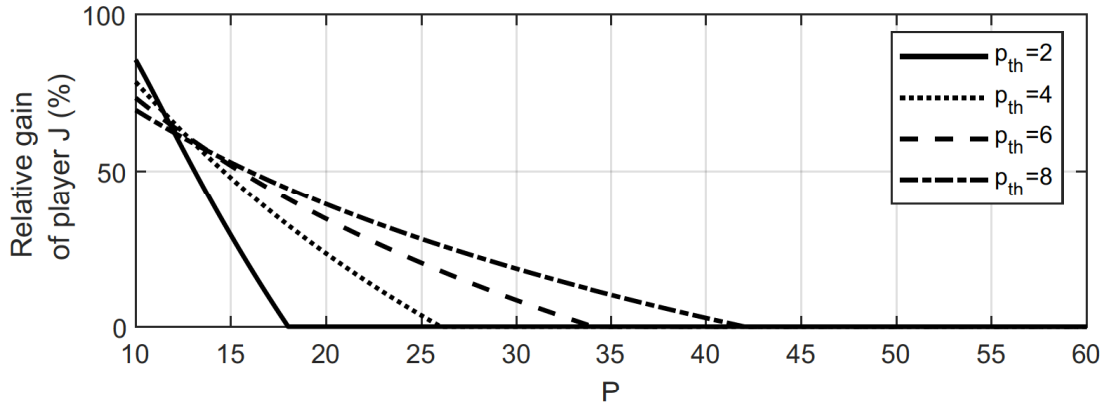

Figure 5.6: Relative gain of player J, evaluated by function E, for different values of p_{th} for $N = 10$, $\sigma_J^2 = 1$ and $\Gamma = 4$.

by the results on Fig. 5.6 where the relative gain of the jammer is presented for different p_{th}. As expected with decreasing the threshold her gain increases. On the other hand, in order to reduce Mallory's gain Alice and Bob must equally allocate full power across all subcarriers. While this is not a surprising result, the investigation above shows that it is the only optimal solution in the specific scenario considered.

5.6 Summary

To summarise this study, analysed injection and reactive jamming attacks in SKG systems and optimal power allocation policies were investigated in BF-AWGN channels. It was shown that pilot randomisation can reduce injection MiM attacks to less harmful jamming attacks. Next, it has been proven that, an intelligent reactive jammer should optimally jam with equal power on the whole spectrum. In the case when Mallory has a fixed sensing threshold the optimal strategy of the legitimate users might not be unique. Depending on their available power they may benefit transmitting at either low power level and keeping their communication unsensed or transmitting at full power and risk to be jammed. Finally, a strategically chosen jamming threshold just below the power level used by the legitimate users, allows the adversary to launch a much more effective attack. In this case, the legitimate users have no choice but to transmit at full power.

Chapter 6

Perspectives

The research on IoT communications is still in its inception and current technologies cannot offer solutions for some emerging applications that need strict requirements. For example, a 3GPP report on the security of URLLC systems notes: "for services with higher speed than 65kbps, the 3GGP Release 15 radio access network (RAN) cannot fulfil the quality of service (QoS) requirement while enforcing user plane integrity protection". Therefore, the research directions that were identified for this thesis were: 1) resource allocation within a short-block length regime; 2) cross-layer architecture supporting stringent QoS requirements; 3) lightweight and fast security protocols; 4) optimisation techniques reducing the computational complexity. Therefore, relevant problems investigated within the present thesis are:

Contributions

Lightweight authentication for IoT networks: Using PLS mechanisms such as PUFs, SKG and localisation the present thesis, Chapter 3 developed a lightweight multi-factor authentication protocol suitable for resource constrained IoT devices. Furthermore, a fast resumption protocol that can allow for data transmission within 0-RTT was proposed. The proposed mechanisms have the potential to pave the way

for a new breed of latency aware protocols.

Short block-length transmission: The size of the control information is often neglected in a system design with long block-length transmission. However, in a short-block length scenario its size might be equivalent to the data, hence better mathematical models must be proposed in order to optimally manage the wireless channel resources. In this regards, a contribution within in Chapter 3 uses probability theory to create a closed form upper-bound on the achievable SKG rate in a short-block length scenario.

Resource allocation in a cross-layer security design: Utilising optimisation tools such as convex optimisation, combinatorial optimisation (0-1 subset-sum Knapsack), dynamic programming and order statistics, Chapter 4 developed a novel AE SKG principle which optimises the resource allocation in a cross-layer setting. This parallel mechanism inter-weaves data and key information and can be applied to a NB-IoT frame type 2.

Stringent end-to-end delay constraints: In Chapter 4, using the theory of effective capacity, the thesis took a further step in the study to examine a pipelined approach using parallel transmission of SKG and data. This allowed the consideration of statistical delay QoS requirements within the solution. Using optimisation tools, the resource allocation that jointly optimises the data rate and the SKG rate was identified. Furthermore, observing the results helped to develop a novel algorithm that addresses a critical problem such as computational complexity.

Possible attacks and countermeasures: With high-reliability requirements, attacks at the physical layer, such as jamming can become a substantial threat to future IoT networks. To address this problem, using game theory analysis, the present thesis evaluated the Stackelberg equilibrium that gives the optimal power allocation in the presence of a reactive and active jammer.

Lightweight security key agreement for B5G applications: Despite the additional overhead in services such as URLLC systems, advanced CSI estimation techniques are employed in order to be able to satisfy the strict reliability requirements. In this regards, the present thesis, as a whole, showed how the CSI-based SKG mechanism can be used towards secure communication. In particular, Chapter 3 developed a fast authentication protocol where the physical layer SKG ensures forward secrecy of each session. Next, Chapter 4 employed the SKG in an optimised resource allocation mechanism which demonstrates both higher efficiency and lower complexity compared to alternative SKG solution. Finally, Chapter 5 how the SKG process could resist injection/jamming attacks providing a secure sesion establishment. To summarise, the proposed solutions within the present thesis show promising result and are ready for testing in a practical environment.

Further work

Continuing the research in the areas described above could lead to further major contributions for future IoT networks. Therefore, building on the methods proposed within the present thesis, the following research directions are proposed as promising next steps:

Flexible numerology: The usage of multiple numerologies within a single framework opens the door to new concerns in the spectral and scheduling efficiency. The multiplexing of URLLC and eMBB traffic in the same physical resource introduces non-orthogonality into the system and increased interference between users. Solutions such as pre-reserving radio resources and preemtive scheduling can lead to the wastage of radio resources or degraded decoding performance, respectively. Initial steps towards solving the above problem range from improvements on coding techniques with shorter decoding latency (as opposed to the heavy Turbo codes used in LTE) and deployment of non-orthogonal optimisation algorithms (such as transfer and meta learning) to flexible design for control and data channels.

Cross-layer resource control: Due to the hierarchical system architecture achieving low

latency and high-reliability simultaneously is a complex task. While separately optimising the physical and MAC layers focuses on improving transmission and queueing performance, respectively, the ultimate goal is to meet end-to-end stringent QoS requirements. The theory of effective capacity can be a helpful tool in identifying a fair trade-off between QoS and achievable rates. However, its application is limited to a constant service rate and does not guarantee an optimal solution. Hence, promising techniques that could enable optimised delivery of services are still sought. In this regard, cross-layer dynamic network slicing may provide the evolution roadmap to future networks. Finally, establishing efficient cross-layer optimisation is still a challenge and requires investigation on the fairness metrics between different network slices and services.

Context-aware optimisation: A further step towards enabling the IoT is building a context aware resource management model. Supporting critical non-scheduled traffic at the physical layer requires interference-free transmission along with the scheduled users. The flexible and shorter transmission time interval introduced for 5G will play an important role in solving this problem, however, its usage raises a set of further questions, such as: optimal subcarrier allocation and block length; reducing processing at the UE to avoid retransmission; and, real-time context-aware modification of the QoS constraints. Additionally, ensuring security for numerous IoT applications can no longer be a static process. In this sense, context-aware security can optimise the system performance using situational and application-based information. Overall, context-awareness could bring benefits ranging from optimised user experience to prevention of unauthorised access.

Strong but nevertheless, lightweight security: The present thesis has presented initial solutions for strong and lightweight security, however, the practical deployment of AE SKG within the IoT framework remains an open problem. Therefore, further research into the deployment of PLS solutions, such as the wireless secret key generation in actual NB-IoT devices and practical PUF authentication could lead to

major contributions for future networks.

mmWave communications and intelligent UE: Next generation networks could benefit from the employment of increased bandwidth with higher frequencies and a non-centralised architecture. On one hand, in mmWave communications methods such as beam selection could greatly improve the end-to-end performance using advanced channel estimation methods and interference mitigation techniques. Furthermore, short-range communication will empower higher localisation accuracy and therefore improving trust and flexibility within the network. On the other hand, systems today are built upon a centralised management architecture. This allows IoT devices to avoid heavy computation but degrades the network efficiency in terms of latency and reliability. In this regard, distributed and lightweight AI based algorithms among edge systems could contribute in aspects such as achieving stringent QoS requirements and enhancing privacy and security.

Closing comment

A simple look at today's communication systems reveals that the security is strictly concentrated in the upper layers of the networking architecture. Features such as authentication, confidentiality, and privacy rely on the complexity based public-key mechanisms. However, the continuous advance of wireless networks has brought fundamental problems making these techniques vulnerable. Furthermore, given the importance of security, it is now vital to take measures at all layers.

In this sense, the use of physical layer properties has been repeatedly ignored and not yet considered as a practical source of secrecy. However, this has recently changed and numerous results in literature suggest that the imperfections of the physical layer can be used as a valid and inherently secure source of randomness. As a result, this thesis was entirely devoted to exploring the PLS paradigm. All of the proposed mechanisms could be directly embedded at the physical layer, within the protocol stack, in a cost-effective manner and contributing towards the realisation of critical IoT applications.

APPENDIX A

Introduction to Tamarin prover

Tamarin prover [42, 290, 291] is a verification tool that can analyse security protocols using symbolic models. Within this thesis Tamarin is used for a formal analysis of the work presented in Chapter 3. Therefore, this section provides an introduction to its syntax and main functionalities.

Tamarin assumes a Dolev-Yao [196] (DY) type of adversary. (DY) model is captured by the following properties:

- The model assumes perfect cryptographic primitives, *i.e.*, the adversary cannot encrypt or decrypt a message without knowing the right key.
- The adversary can act as an eavesdropper, can block and modify messages that have been sent over the network and can inject messages.
- The adversary can inject messages that he/she constructed through eavesdropping and some initial pre-defined knowledge.
- The adversary can execute the protocol multiple times taking any role in the communication system, hence the roles of the involved agents can be changed, too.

Before beginning with the definitions a quick notation guideline of the Tamarin syntax is given in Fig. A.1. Next, to familiarise the readers with Tamarin a quick example of a protocol definition is introduced. The protocol models the scenario: Alice using a

```
~x          // x is a fresh random variable - usually used as a key
$A          // A is public variable - usually used as an agent ID
#i          // i is temporal variable - usually used to define time-point
m           // mesage m
<x,y>       // concatenation of x and y
!Fact       // persistent fact that is true for all states
@i          // at time-point i
.           // such that
==>         // implication
&           // conjunction
|           // disjunction
not         // negation
All         // all elements
F           // inequality
i1 < i2     // temporal variables ordering
#i1=#i2     // equality between temporal variables
m1=m2       // equality between messages
Reveal(A)   // Compromised agent (in this case agent A)
Fr(~x)      // Fresh variable (it has not been previously used)
In(x)       // Receipt of message x (it is available to the adversary)
Out(x)      // Sending of message x (it is available to the adversary)
K(x)        // Knowledge of the adversary; in this case variable x
```

Figure A.1: Example of common commands used in the Tamarin environment.

symmetric key $\mathbf{k}$ sends an encrypted message $\mathbf{m}$ to Bob who also posses the key $\mathbf{k}$. This is done by defining three sets of parameters: Cryptographic primitives, set of rules and set of lemmas.

1) Cryptographic primitives that are used in the protocol are defined in Fig. A.2. As shown every definition of a protocol begins with the keyword `theory` and name of the theory given by user's choice. Next each description of a protocol begins and ends with keywords `begin` and `end` respectively. Tamarim has already builtin message theories that can be easily included as shown on line 3. The theory `symmetric-encryption` models the functions `senc/2` and `sdec/2` that correspond to symmetric encryption and decryption, respectively. Note that the notation `f/n` defines function `f` with arity `n`. The two functions are related thought the equation: `sdec(senc(m,k),k)=m`, where `m` is message and `k` a symmetric key. Other built in theories that can be used are: `hashing`, `asymmetric-encryption`, `signing`, `xor` and more; including any of these will unlock functions and equations from the corresponding theory to the user. Next, a user

```
1 theory Name_of_the_theory
2 begin
3 builtins: symmetric-encryption
4 //functions: summation/2, subtraction/2,
5 //           zero/0
6 //equations: subtraction(x.1,x.1)=zero
7 ...
8 end
```

Figure A.2: Example of message theories in Tamarin

can define own functions and equations. Even though neither functions nor equations are required for the purpose here, in order to better illustrate the possibilities of Tamarin few are added to Fig. A.2 as comments. Two functions are given on line 4: `summation` and `subtraction` both with arity 2 (*i.e.*, can take two arguments as input). Finally, the equation on line 6 represents the following equation $x - x = 0$.

2) A set of rules are used to represent the state transitions of the protocol, *i.e.*, a single rule contains a single state (Alice's or Bob's). Every rule has 3 components: a premise, action facts and a conclusion which as illustrated on Fig. A.3, are written in the form `[premise] --[action facts]-> [conclusion]`. In order to execute a specific rule all facts in the premise must be available in the current state; next, the conclusion contains set of facts that will overwrite the facts of the premise (excluding persistent facts which are denoted with exclamation mark as `!Fact`) and will generate a new state; finally, actions facts are used to relate the rule to a specific restriction or a property that has to be proven, they do not appear in the state but on the trace. The rules used to model the example of Alice sending symmetrically encrypted message to Bob are given in Fig. A.3. It can be seen that this is represented by 3 rules. As mentioned earlier, in order to execute a specific rule all facts in its premise must be available at the current state. Due to that the protocol begins with the rule `Initialisation` which is used to define the initial knowledge of each party and more specifically where these knowledge come from. In the premise we have two facts of type `Fr` one to generate a fresh key `k` and one to generate a fresh message `m`. The rule does not have any action facts, but does have two facts in its conclusion: these are Alice's (`AliceIn`) and Bob's (`BobIn`)

```
rule Initialisation:
[ Fr(~k), Fr(~m) ]
-->
[ AliceIn($A,~k,~m), BobIn($B,~k) ]

rule AliceSends:
[ AliceIn($A,~k,~m) ]
--[ Send($A,~m) ]->
[ Out(senc(~m,~k)) ]

rule BobReceives:
[ BobIn($B,~k), In(senc(m,~k))  ]
--[ Receive($B,m), Secret(m) ]->
[ ]
```

Figure A.3: Example of protocol definition in Tamarin

initial knowledge. `AliceIn` includes her ID, which is a public variable (denoted with dollar sign `$A`), the symmetric key and the message, which are fresh random variables `~k` and `~m`, respectively. `BobIn` includes his ID `$B` and the symmetric key `~k`. The next rule `AliceSends` models the process of sending the encrypted message. The premise contains of a single fact to define the initial knowledge required for Alice to send the encrypted message. The rule contains an action fact `Send` that will be used to relate the rule to specific security property (details of how action facts are used will be given in the next item). Finally, the conclusion of the rule models the fact of sending encrypted message to the network, which is done by using the built in fact `Out` for sending and the built in function `senc` available from the in-built theory `symmetric-encryption`. As mentioned earlier, anything within `Out` fact becomes available to the adversary. The last rule `BobReceives` models the receiving of the message at Bob's side. As in the previous rule, the premise recalls the initial knowledge required for the execution of the rule, *i.e.*, Bob's initial knowledge given by the fact `BobIn` and Bob receiving the message from Alice denoted by the built in fact `In`. Note that Tamarin assumes that all received messages are coming from the adversary, *i.e.*, as in this case Alice sends the message, the adversary captures it, performs an attack if possible (such as modifying the message), or just keeps a copy of the message and forwards it to Bob. This rule has two action

```
lemma secrecy:
    "All m #i1 #i2.
     Secret(m) @i1 & K(m) @i2 ==> F"

lemma executable: exists-trace
    "Ex A B m #i1 #i2.
    Send(A,m) @i1 & Receive(B,m) @i2
    & i1 < i2"
```

Figure A.4: Example of security properties in Tamarin

facts which will be used to ensure the security properties of the protocol in the next item. Finally, since Bob does not undertake any further actions and no facts will be used in another rule the conclusion is left empty.

3) A set of lemmas are used to model the properties that need to be proven. Particularly, lemmas are used to relate the rules to specific security properties. As mentioned above lemmas and rules are linked through action facts. Example is given in Fig. A.4: The first lemma `secrecy` tests if the message `m` is secret from Bob's perspective. It involves the action fact `Secret` from the rule `BobIn`. It implies that the input of the action fact `Secret` is available only to legitimate users and it is not part of adversary's knowledge denoted by the fact `K`. It can be read as follows: For all messages `m` and time-points `i1, i2`, the fact `Secret(m)` at time `i1` and `K(m)` at time `i2` imply inequality.

By default lemmas are proven true only if it holds for all possible traces, *i.e.*, all possible scenarios and actions of the adversary. However, if we want to prove that something could be true only in some of all traces Tamarin has an additional built-in function called `exists-trace`. The function is used in the second lemma `executable`. The lemma is used to verify that the protocol can be executed. It proves the following: there exists a trace where agent `A` sends message `m` at time-point `i1` and agent `B` receives it at time-point `i2`, the addition `i1<i2`, defines that `Send` happens before `Receive`.

As an addition to lemmas, one may add restrictions which can define specific properties of the action facts. An example which defines equality is given in Fig. A.5. The meaning of the restriction in Fig. A.5 is: whenever the action fact `Equal(x,y)` appears in the protocol definition it implies that `x = y` and this is true for all `x, y` and time-

```
restriction Equality:
    "All x y #i.
    Equal(x,y) @i ==> x = y"
```

Figure A.5: Restriction example in Tamarin

points `i`. When constructed as lemmas, all restrictions to create an effect must be linked to an action fact, which in this case is `Equal`. This restriction might be useful whenever one wants to verify another's signature. However, for the simple example presented above, this restriction is not linked with any of the action facts of the rules and hence, will not have an effect on the execution of the protocol.

To prove the security properties (lemmas) Tamarin provides an interactive graphical user interface (GUI). The proof itself begins with deduction of the knowledge, that Tamarin has about the property, and finishes with either a possible attack (counterexample) or a conclusion for correctness of the lemma. Now by combining the codes from Fig. A.2, A.3 and A.4, one we can try to prove lemma `secrecy`, *i.e.*, that the message `m` is secret from the adversary. The notation used by the Tamarin GUI is presented in Fig. A.6. As shown on the figure the GUI in Tamarin represents the rules of the protocol as a rectangle boxes where the premise, action facts and conclusion are separated on different levels. The adversary actions are represented as ellipses; black to illustrate his reasoning and grey to illustrate his goals (*i.e.*, facts needed to find a counterexample of lemmas). These boxes and ellipses are connected by arrows of different colours and type, where each has its own meaning as shown in the figure. Whenever a linear fact is consumed by the premise of a rule and is coming from a conclusion of another rule, this is shown by a black arrow (such facts are the `Alicein` and `BobIn` in Fig. A.3). Whenever a persistent fact (one labeled with exclamation mark `!Fact`) is consumed by the premise of a rule and is coming from a conclusion of another rule, this is shown by a grey arrow. All deductions of the adversary are represented by red arrows. Finally, the relation of time-points is illustrated using dotted arrows.

Following the above notation the proof of secrecy of the message `m` begins from Fig.

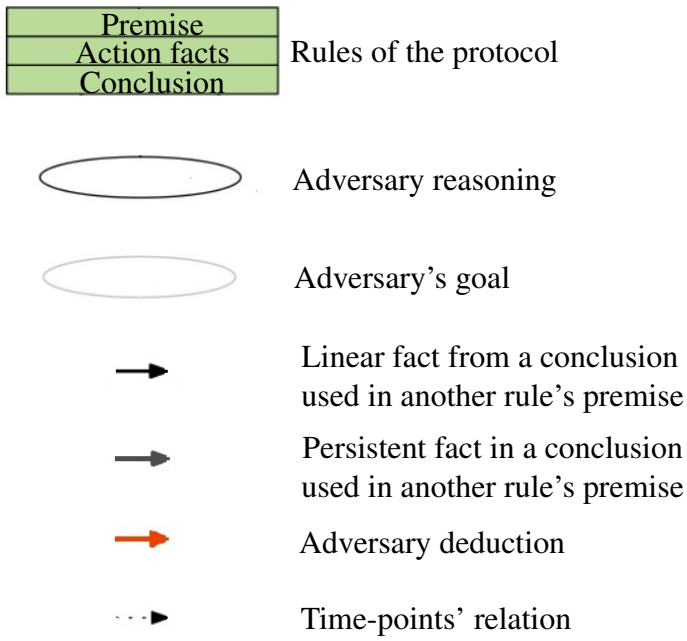

Figure A.6: Notation used by Tamarin's GUI

A.7. As discussed, the box in Fig. A.7 represents the rule linked to the lemma which has to be proven. There are two facts in the premise of the rule `BobIn` and `In`. On the lower level are the actions facts and the name of the rule, particularly action fact `Secret` is the one that links the rule with the lemma. The rule does not have any facts in the conclusion, hence Tamarin does not add anything at the lower level. As mentioned earlier, all received messages are assumed to be coming from the adversary. This is shown on the left of the figure. An arrow connects the premise and more specifically the `In` fact showing that the encrypted version of the message is sent by adversary, however the dotted line and the gray ellipse above show that, for this to happen the adversary should have got it from somewhere, so the first goal of the adversary here is to find where it comes from. Then, the right top side of the figure shows the final goal of adversary, *i.e.*, to know message `m`, represented by the fact `!KU(m)`. During the next steps of the proof the adversary starts a backwards search of and deducts the facts within the rules to check if it is possible to achieve the goals shown here.

Next, Tamarin starts looking for possible paths of the facts given in the premise, and

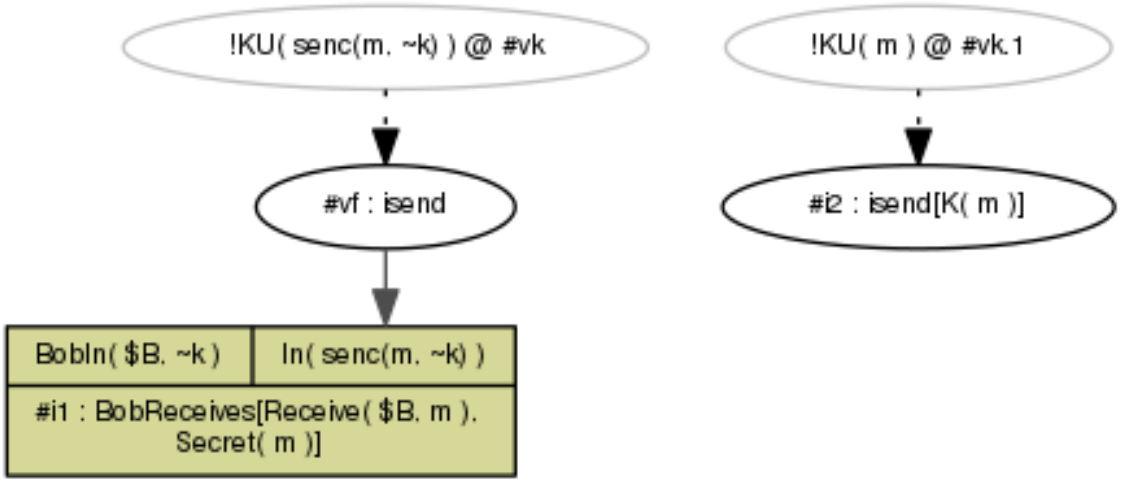

Figure A.7: Proving secrecy of message `m`: Part 1

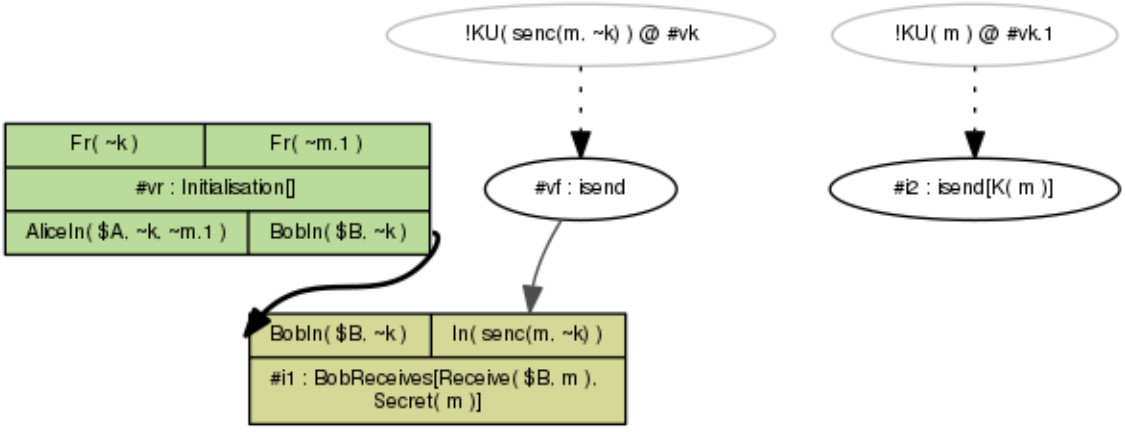

Figure A.8: Proving secrecy of message `m`: Part 2

where do they come from. As mentioned above the `In` fact comes from the adversary. Now, in Fig. A.8 Tamarin has found the origin of the other fact in the premise. This is illustrated by the black solid line; the conclusion of the rule `Initialisation` and rule `BobReceives` are linked through the linear fact `BobIn`. However, the origin of the message is still unknown.

Fig. A.9 shows the next step of the proof. Here, Tamarin has found the origin of the encrypted message. On the left side of the figure it can be see that before forwarding the message to Bob the adversary has received it from Alice. Alice sent the message using `Out` fact, and as mentioned earlier, anything sent into the network becomes available to

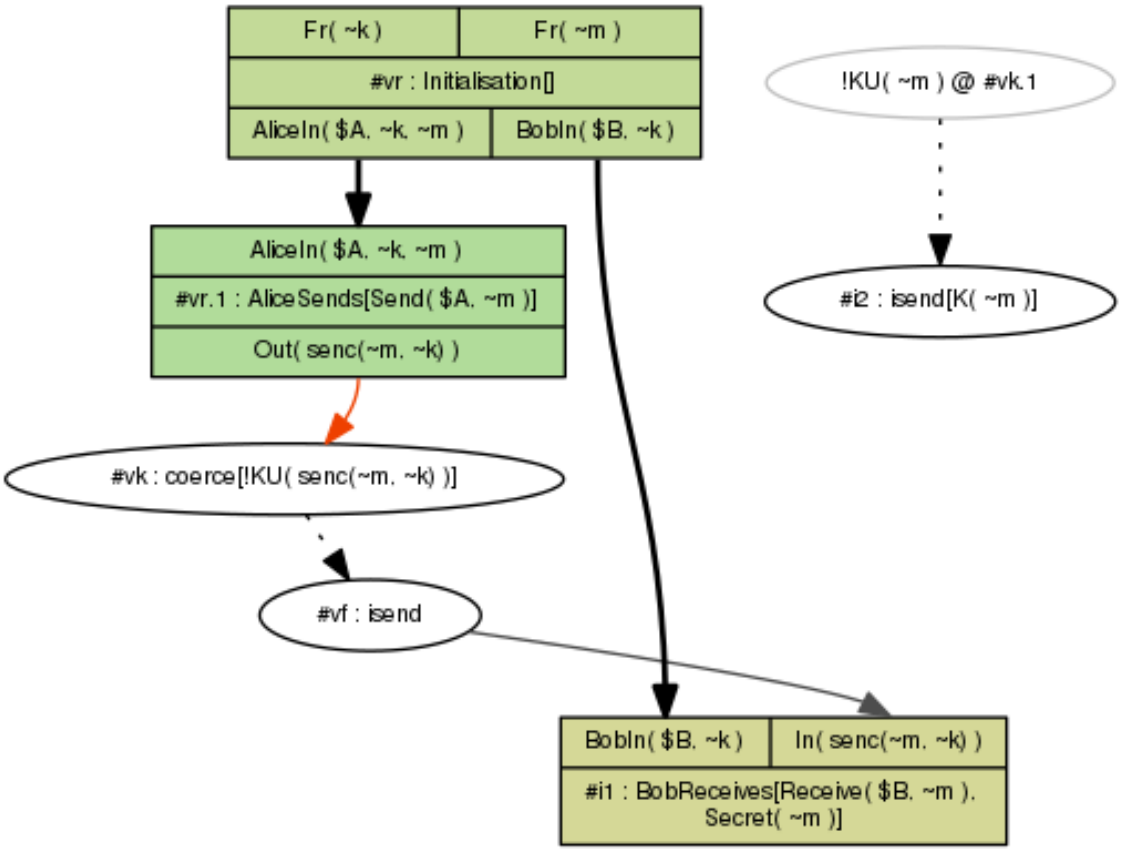

Figure A.9: Proving secrecy of message `m`: Part 3

the adversary; this is denoted with red arrow. In the middle line Tamarin gives the name of the rule `AliceSends` and the corresponding action facts, and on the top line is the premise which contains a single linear fact `AliceIn`. The origin of the linear fact has been found, which as in Bob's case is the conclusion of the rule `Initialisation`. It can be seen that, once the adversary has found the origin of a fact an ellipse becomes black. In this current stage the only goal left in order to find an attack and break the lemma `secrecy` is the one in the right side of the figure.

Tamarin continues looking for possible pieces of knowledge that could help the adversary learn the message `m`. This is illustrated on Fig. A.10. The program understands that the only possible source to learn that is the symmetrically encrypted message sent by Alice denoted as `senc(~m, ~k)`. Here, once the adversary receives the message at time-point `#vl`, on one hand he forwards the message to Bob whine on the other hand,

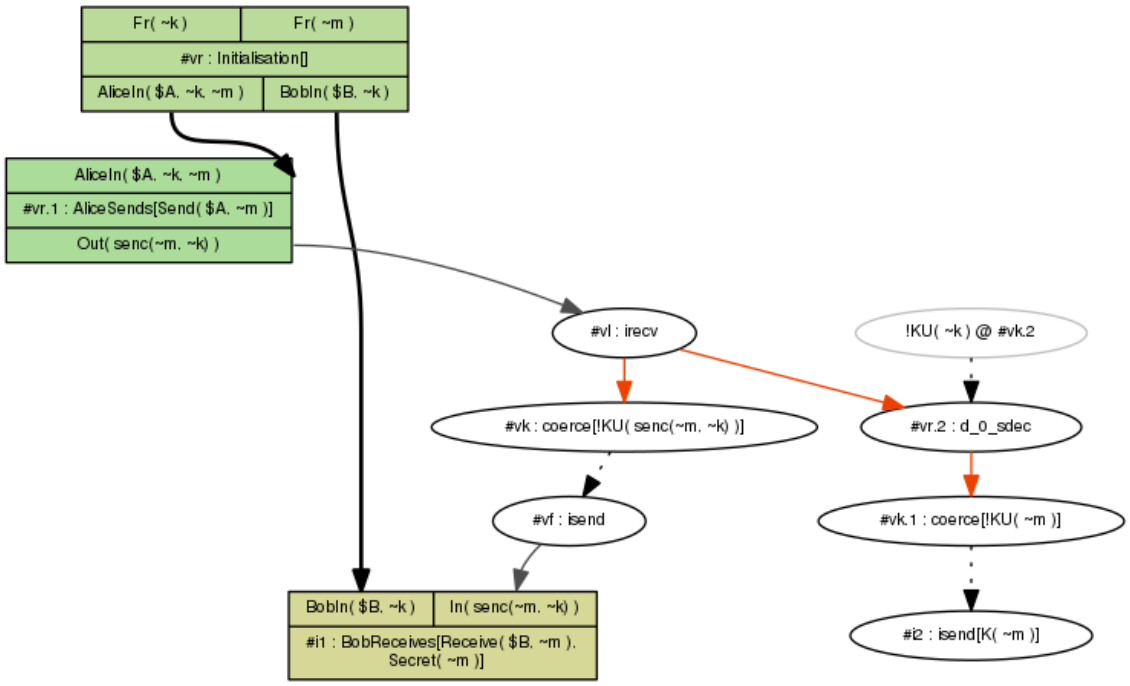

Figure A.10: Proving secrecy of message `m`: Part 4

tries to learn `m`. Tamarin understands that the only way to get `m` is to decrypt it using the key `k`. Given that the goal of the adversary has now changed and it will start looking for possible sources of the key.

The last part of the proof is presented on Fig. A.11. After all deductions shown on the previous figures Tamarin has concluded that in order for the adversary to know the message `m`, he should either know `m` itself in advance or the key `k`. However, there is no trace of the protocol which gives that possibility. As a result, the adversary's goals in the final part of the proof, *i.e.*, the ellipses on the top of Fig. A.11, are still in gray. Therefore, the simple example of a protocol provides secrecy to the message `m`.

Now, to finalise the introduction to Tamarin, a quick example of an insecure protocol is showed on Fig. A.12. The protocol is built by adding an extra fact within the `AliceSends` rule. The extra fact is `Out(~k)`, *i.e.*, along with the encrypted message `senc(~m, ~k)` Alice also sends her key `k` in clear text to the network. Fig. A.12 shows only the last step of the proof. It illustrates that in this case the adversary can obtain the secret message `m`. As in the previous case when Alice send the encrypted message, the ad-

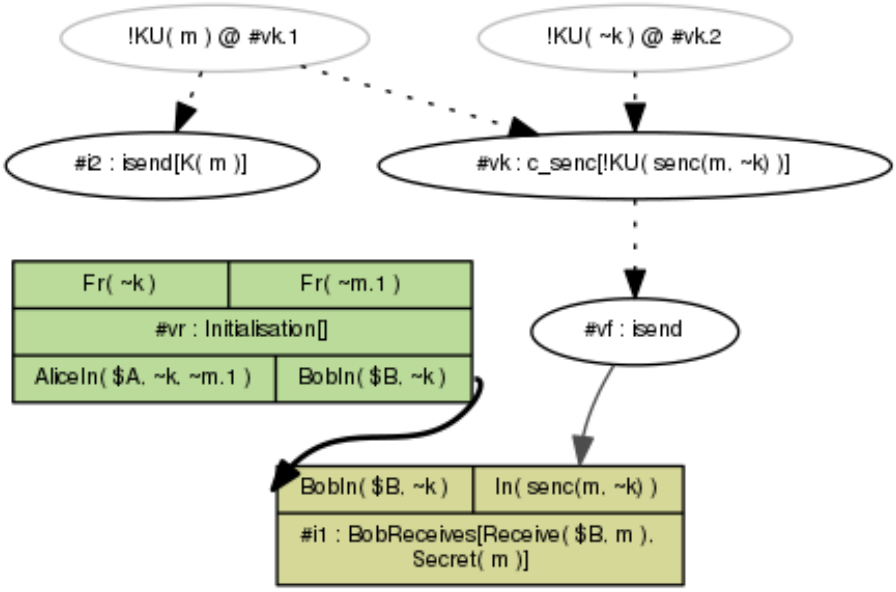

Figure A.11: Proving secrecy of message `m`: Final part

versary captures it at time-point `#vl`. However, Alice also sends her key in the network. The adversary captures that at moment `#vk.2`. Combining these two facts at moment `#vr.2` the adversary decrypts the encrypted message; as a result at moment `#vk.1` the adversary obtains the message `m`. This proves the lemma secrecy is not true anymore. In contrast to the previous case, all ellipses (facts used by the adversary) leading to the adversary knowing the messages, *i.e.*, `K(~m)` are in black. This means that all goals of the adversary have been achieved and Tamarin has found a counter example of lemma `secrecy`. This concludes the introduction to Tamarin-prover. Tamarin is used in Chapter 3 where it is used as a verification tool in order to prove security properties of the proposed authentication protocol.

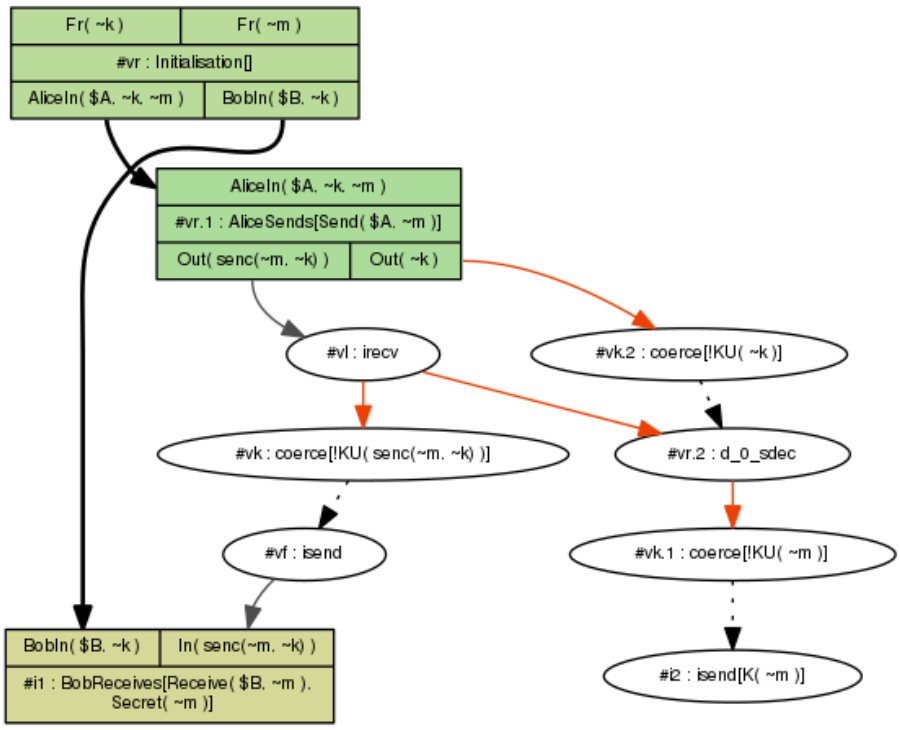

Figure A.12: Example of insecure protocol in Tamarin

APPENDIX B

Derivation of the channel dispersion in a finite block-length scenario

This section provides the steps towards finding the closed form expression of the dispersion of the channel model assumed in Section 3.3. The initial form of the dispersion is:

$$V = \mathsf{Var}\left[\log \frac{\mathsf{Pr}(x_A, x_B)}{\mathsf{Pr}(x_A)\mathsf{Pr}(x_B)}\right], \tag{B.1}$$

where the observations of Alice and Bob are expressed as:

$$x_A = h + z_A \tag{B.2}$$

$$x_B = h + z_B. \tag{B.3}$$

In the above $z_A, z_B \sim N(0,1)$ and $\mathsf{Var}(h) = P$. Therefore, one can conclude $\mathsf{Pr}(x_A) \sim N(0, P+1)$, $\mathsf{Pr}(x_B) \sim N(0, P+1)$ and

$$\mathsf{Pr}(x_A, x_B) \sim (\mu, \mathbf{\Sigma}) \sim \left(\begin{bmatrix} 0 \\ \\ 0 \end{bmatrix}, \begin{bmatrix} 1+P & P \\ \\ P & 1+P \end{bmatrix}\right), \tag{B.4}$$

where μ, represents mean and $\mathbf{\Sigma}$ is the covariance matrix.

Due to the fact x_A and x_B follow normal distribution we have [292]:

$$\mathsf{Pr}(x_A) = \frac{\exp(-\frac{1}{2}\left(\frac{(x_A-\mu)}{\sigma}\right)^2)}{\sigma\sqrt{2\pi}}, \tag{B.5}$$

$$\mathsf{Pr}(x_B) = \frac{\exp(-\frac{1}{2}\left(\frac{(x_B-\mu)}{\sigma}\right)^2)}{\sigma\sqrt{2\pi}}. \tag{B.6}$$

For the case observed here, *i.e.*, $\sigma = \sqrt{P+1}$ and $\mu = 0$, the above equals to:

$$\mathsf{Pr}(x_A) = \frac{\exp(-\frac{x_A^2}{2\sqrt{P+1}})}{\sqrt{2\pi(P+1)}}, \tag{B.7}$$

$$\mathsf{Pr}(x_B) = \frac{\exp(-\frac{x_B^2}{2\sqrt{P+1}})}{\sqrt{2\pi(P+1)}} \tag{B.8}$$

Next, due to the fact $\exp(x)\exp(y) = \exp(x+y)$, the product of (B.7) and (B.8) is:

$$\mathsf{Pr}(x_A)\mathsf{Pr}(x_B) = \frac{\exp(-\frac{x_A^2+x_B^2}{2\sqrt{P+1}})}{2\pi(P+1)} \tag{B.9}$$

Now, the multivariate normal distribution $\mathsf{Pr}(x_A, x_B)$ is [293]:

$$\mathsf{Pr}(x_A, x_B) = \frac{\exp(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu))}{\sqrt{(2\pi)^{|\mathbf{x}|}\det(\mathbf{\Sigma})}}, \tag{B.10}$$

where det, stands for determinant and $|\cdot|$ gives the size of a vector. Combining (B.7), (B.8) and (B.10) one can obtain the expression for $V = \mathsf{Var}\left[\log\frac{\mathsf{Pr}_{x_A,x_B}}{\mathsf{Pr}_{x_A}\mathsf{Pr}_{x_B}}\right]$ as:

$$V = \mathsf{Var}\left[\log\frac{\mathsf{Pr}_{x_A,x_B}}{\mathsf{Pr}_{x_A}\mathsf{Pr}_{x_B}}\right] = \mathsf{Var}\left[\log\frac{2\pi(P+1)[\exp(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu))]}{\exp\left(-\frac{x_A^2+x_B^2}{2\sqrt{P+1}}\right)\sqrt{(2\pi)^{|\mathbf{x}|}\det(\mathbf{\Sigma})}}\right] \tag{B.11}$$

For simplicity the denominator and the numerator of (B.11) are treated separately.

Starting with the denominator it is known that $|\mathbf{x}| = 2$, furthermore, the determinant of $\mathbf{\Sigma}$ easily can be evaluated as $(1+P)^2 - P^2 = 1 + 2P$. Therefore, the denominator of (B.11) is equal to:

$$\exp\left(-\frac{x_A^2 + x_B^2}{2\sqrt{P+1}}\right)\sqrt{4\pi^2(1+2P)} \tag{B.12}$$

$$= \exp\left(-\frac{x_A^2 + x_B^2}{2\sqrt{P+1}}\right)\sqrt{4\pi^2}\sqrt{1+2P} \tag{B.13}$$

$$= \exp\left(-\frac{x_A^2 + x_B^2}{2\sqrt{P+1}}\right)2\pi\sqrt{1+2P} \tag{B.14}$$

Next, one can expand the numerator as follows:

$$2\pi(P+1)\exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right) \tag{B.15}$$

$$= 2\pi(P+1)\exp\left(-\begin{bmatrix}\frac{x_A}{2} & \frac{x_B}{2}\end{bmatrix}\mathbf{\Sigma}^{-1}\begin{bmatrix}x_A \\ \\ x_B\end{bmatrix}\right) \tag{B.16}$$

The inverse of $\mathbf{\Sigma}$ is:

$$\mathbf{\Sigma}^{-1} = \frac{1}{(1+P)^2 - P^2}\begin{bmatrix}1+P & -P \\ \\ -P & 1+P\end{bmatrix} = \frac{1}{1+2P}\begin{bmatrix}1+P & -P \\ \\ -P & 1+P\end{bmatrix} = \begin{bmatrix}\frac{1+P}{1+2P} & \frac{-P}{1+2P} \\ \\ \frac{-P}{1+2P} & \frac{1+P}{1+2P}\end{bmatrix} \tag{B.17}$$

Combining (B.16) and (B.17), results in:

$$2\pi(P+1)\exp\left(-\begin{bmatrix}\frac{x_A}{2} & \frac{x_B}{2}\end{bmatrix}\begin{bmatrix}\frac{1+P}{1+2P} & \frac{-P}{1+2P} \\ \\ \frac{-P}{1+2P} & \frac{1+P}{1+2P}\end{bmatrix}\begin{bmatrix}x_A \\ \\ x_B\end{bmatrix}\right) \tag{B.18}$$

$$= 2\pi(P+1)\text{exp}\left(-\begin{bmatrix} \frac{x_A(1+P)}{2+4P} + \frac{-x_BP}{2+4P} & \frac{-x_AP}{2+4P} + \frac{x_B(1+P)}{2+4P} \end{bmatrix} \begin{bmatrix} x_A \\ \\ x_B \end{bmatrix}\right) \tag{B.19}$$

$$= 2\pi(P+1)\text{exp}\left(-\begin{bmatrix} \frac{x_A(1+P)-x_BP}{2+4P} & \frac{x_B(1+P)-x_AP}{2+4P} \end{bmatrix} \begin{bmatrix} x_A \\ \\ x_B \end{bmatrix}\right) \tag{B.20}$$

$$= 2\pi(P+1)\text{exp}\left(-\left[\frac{x_A^2(1+P)-x_Ax_BP}{2+4P} + \frac{x_B^2(1+P)-x_Ax_BP}{2+4P}\right]\right) \tag{B.21}$$

$$= 2\pi(P+1)\text{exp}\left(-\frac{x_A^2(1+P)-2x_Ax_BP+x_B^2(1+P)}{2+4P}\right) \tag{B.22}$$

$$= 2\pi(P+1)\text{exp}\left(-\frac{(x_A^2+x_B^2)(1+P)-2x_Ax_BP}{2+4P}\right) \tag{B.23}$$

Now, combining the expression for the numerator (B.23) and the expression for the denominator (B.14) the full expression for V becomes:

$$V = \mathsf{Var}\left[\log \frac{(P+1)\text{exp}(-\frac{(x_1^2+x_2^2)(1+P)-2x_1x_2P}{2+4P})}{\text{exp}\left(-\frac{x_1^2+x_2^2}{2\sqrt{P+1}}\right)\sqrt{1+2P}}\right] \tag{B.24}$$

Using the facts $\frac{\text{exp}(x)}{\text{exp}(y)} = \text{exp}(x-y)$, $\log(\text{exp}(x)) = x$, $\log(\frac{x}{y}) = \log(x) - \log(y)$ and $\mathsf{Var}(x + \text{const}) = \mathsf{Var}(x)$ one can expand Eq. (B.24) as:

$$\begin{aligned} V = \mathsf{Var}\Bigg(& \frac{1}{(1+2\,P)\,2\sqrt{P+1}}\Big[-\sqrt{P+1}{x_A}^2P - \sqrt{P+1}{x_A}^2 \\ & + 2\,\sqrt{P+1}x_Ax_BP - \sqrt{P+1}{x_B}^2P - \sqrt{P+1}{x_B}^2 + {x_1}^2 \\ & + 2\,{x_A}^2P + {x_2}^2 + 2\,{x_B}^2P + 2\,\log\left(\frac{P+1}{\sqrt{1+2\,P}}\right)\sqrt{P+1} \\ & + 4\,\log\left(\frac{P+1}{\sqrt{1+2\,P}}\right)\sqrt{P+1}P\Big]\Bigg) \end{aligned} \tag{B.25}$$

The above equation has the form $\mathsf{Var}(aX + aY + bZ)$, which is equivalent to:

$$\mathsf{Var}(aX + aY + bZ) = a^2\mathsf{Var}(X) + a^2\mathsf{Var}(Y) + b^2\mathsf{Var}(Z)$$
$$+ 2a^2\mathsf{Cov}(X, Y) + 2ab\mathsf{Cov}(X, Z) + 2ab\mathsf{Cov}(Y, Z)$$

where:

$$a = \frac{\left(-\sqrt{P+1} - \sqrt{P+1}P + 1 + 2\,P\right)}{2\,(1+2\,P)\,\sqrt{P+1}} \tag{B.26}$$

and:

$$b = \frac{P}{1+2P} \tag{B.27}$$

Therefore, V equals:

$$V = a^2\mathsf{Var}(x_A^2) + a^2\mathsf{Var}(x_B^2) + b^2Var(x_Ax_B)$$
$$+ 2a^2\mathsf{Cov}(x_A^2, x_B^2) + 2ab\mathsf{Cov}(x_A^2, x_Ax_B) + 2ab\mathsf{Cov}(x_B^2, x_Ax_B) \tag{B.28}$$

Now using the facts:

$$\mathsf{E}(x^2) = \mathsf{Var}(x) = P + 1, \tag{B.29}$$

$$\mathsf{E}(x) = \mu(x) = 0, \tag{B.30}$$

the following can be concluded:

$$\mathsf{Var}(x_A^2) = \mathsf{Var}(x_B^2) = 2(P+1)^2. \tag{B.31}$$

Next, for $\mathsf{Var}(x_A, x_B)$ we have:

$$\mathsf{Var}(x_A, x_B) = [\mathsf{Var}(x_A) + \mathsf{E}(x_A)^2] \cdot [\mathsf{Var}(x_B) + \mathsf{E}(x_B)^2] +$$
$$\mathsf{Cov}(x_A^2, x_B^2) - [\mathsf{Cov}(x_A, x_B) + \mathsf{E}(x_A)\mathsf{E}(x_B)]^2 \tag{B.32}$$

Recalling (B.29) and (B.30) follows:

$$\mathsf{Var}(x_A, x_B) = (P+1)^2 + \mathsf{Cov}(x_A^2, x_B^2) - \mathsf{Cov}(x_A, x_B)^2 \tag{B.33}$$

Now since h, z_A, z_B are independent in the channel model assumed here, it follows:

$$\begin{aligned}
\mathsf{Cov}(x_A, x_B) &= \mathsf{E}[x_A x_B] - \mathsf{E}[x_A]\mathsf{E}[x_B] \\
&= \mathsf{E}[(h+z_A)(h+z_B)] - 0 \\
&= \mathsf{E}[h^2] + \mathsf{E}[h]\mathsf{E}[z_B] + \mathsf{E}[h]\mathsf{E}[z_A] + \mathsf{E}[z_A]\mathsf{E}[z_B] \\
&= P
\end{aligned} \tag{B.34}$$

and:

$$\begin{aligned}
\mathsf{Cov}(x_A^2, x_B^2) &= \mathsf{E}[x_A^2 x_B^2] - \mathsf{E}[x_A^2]\mathsf{E}[x_B^2] \\
&= \mathsf{E}[(h+z_A)^2(h+z_B)^2] - (P+1)^2 \\
&= \mathsf{E}[h^4] + \mathsf{E}[h^2]\mathsf{E}[z_B^2] + \mathsf{E}[h^2]\mathsf{E}[z_A^2] + \mathsf{E}[z_A^2]\mathsf{E}[z_B^2] - (P+1)^2 \\
&= \mathsf{E}[h^4] + P + P + 1 - (P+1)^2 \\
&= 1 + 3P^2 + 2P - (P+1)^2
\end{aligned} \tag{B.35}$$

Using the results from Eq. (B.34) and (B.35) the final expression for $\mathsf{Var}(x_A, x_B)$ is:

$$\begin{aligned}
\mathsf{Var}(x_A, x_B) &= (P+1)^2 + 1 + 3P^2 + 2P - (P+1)^2 - P^2 \\
\mathsf{Var}(x_1, x_2) &= 1 + 2P^2 + 2P
\end{aligned} \tag{B.36}$$

The last missing pieces from the expression of V are: $\mathsf{Cov}(x_B^2, x_A x_B)$ and $\mathsf{Cov}(x_A^2, x_A x_B)$. Both of which can be evaluated as:

$$\begin{aligned}
\mathsf{Cov}(x_A^2, x_A x_B) &= \mathsf{E}[x_A^3 x_B] - \mathsf{E}[x_A^2]\mathsf{E}[x_A x_B] \\
&= \mathsf{E}[x_A^3 x_B] - (P+1)P
\end{aligned}$$

$$
\begin{aligned}
&= \mathsf{E}[(h+z_A)^3(h+z_B)] - (P+1)P \\
&= \mathsf{E}[(h^3+z_A^3+3h^2z_A+3hz_A^2)(h+z_B)] - (P+1)P \\
&= \mathsf{E}[h^4] + 3\mathsf{E}[h^2]\mathsf{E}[z_A^2] - (P+1)P
\end{aligned}
$$

Therefore we have:

$$
\mathsf{Cov}(x_A^2, x_A x_B) = \mathsf{Cov}(x_B^2, x_A x_B) = 3P^2 + 3P - P(P+1) \tag{B.37}
$$

Following from the above calculations the full expression of V can be written as:

$$
\begin{aligned}
V =& a^2\mathsf{Var}(x_A^2) + a^2\mathsf{Var}(x_B^2) + b^2\mathsf{Var}(x_A x_B) \\
&+ 2a^2\mathsf{Cov}(x_A^2, x_B^2) + 2ab\mathsf{Cov}(x_A^2, x_A x_B) + 2ab\mathsf{Cov}(x_B^2, x_A x_B) \\
=& a^2 2(P+1)^2 + a^2 2(P+1)^2 + b^2(1+2P^2+2P) + \\
& 2a^2(1+3P^2+2P-(P+1)^2) + 2ab(3P^2+3P-P(P+1)) + \\
& 2ab(3P^2+3P-P(P+1)) \\
=& 4a^2(P+1)^2 + b^2(1+2P^2+2P) + \\
& 2a^2(1+3P^2+2P-(P+1)^2) + 4ab(3P^2+3P-P(P+1)).
\end{aligned} \tag{B.38}
$$

By combining some of the terms the final result is:

$$
V = 8aP(aP+a+bP+b) + 2b^2P(P+1) + 4a^2 + b^2, \tag{B.39}
$$

where a and b are defined in Eq. (B.26) and Eq. (B.27), respectively.

Bibliography

[1] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] B. Zan, M. Gruteser, and F. Hu, "Key agreement algorithms for vehicular communication networks based on reciprocity and diversity theorems," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 8, pp. 4020–4027, Oct 2013.

[3] J. Wan, A. B. Lopez, and M. A. Al Faruque, "Exploiting wireless channel randomness to generate keys for automotive cyber-physical system security," in *2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS)*, April 2016, pp. 1–10.

[4] Y. Liu, J. Jing, and J. Yang, "Secure underwater acoustic communication based on a robust key generation scheme," in *2008 9th International Conference on Signal Processing*, Oct 2008, pp. 1838–1841.

[5] I. U. Zaman, A. B. Lopez, M. A. A. Faruque, and O. Boyraz, "Physical layer cryptographic key generation by exploiting pmd of an optical fiber link," *Journal of Lightwave Technology*, vol. 36, no. 24, pp. 5903–5911, Dec 2018.

[6] D. Tian, W. Zhang, J. Sun, and C. Wang, "Physical-layer security of visible light communications with jamming," in *2019 IEEE/CIC International Conference on Communications in China (ICCC)*, Aug 2019, pp. 512–517.

[7] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, pp. 614–626, 2016.

[8] H. Shehadeh and D. Hogrefe, "A survey on secret key generation mechanisms on the physical layer in wireless networks," *Security and Communication Networks*, vol. 8, no. 2, pp. 332–341, Jan. 2015.

[9] A. Chorti, K. Papadaki, and H. V. Poor, "Optimal power allocation in block fading channels with confidential messages," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 4708–4719, Sep. 2015.

[10] A. Chorti, S. Perlaza, Z. Han, and H. V. Poor, "On the resilience of wireless multiuser networks to passive and active eavesdroppers," *Selected Areas in Communications, IEEE Journal on*, vol. 31, pp. 1850–1863, 09 2013.

[11] A. Chorti, C. Hollanti, J.-C. Belfiore, and H. V. Poor, "Physical layer security: A paradigm shift in data confidentiality," *Lecture Notes in Electrical Engineering*, vol. 358, 01 2016.

[12] M. Mitev, A. Chorti, M. Reed, and L. Musavian, "Authenticated secret key generation in delay-constrained wireless systems," *Journal on Wireless Communication and Networking*, no. 122, June 2020.

[13] M. Mitev, A. Chorti, and M. Reed, "Optimal resource allocation in joint secret key generation and data transfer schemes," in *2019 15th International Wireless Communications Mobile Computing Conference (IWCMC)*, June 2019, pp. 360–365.

[14] M. Mitev, A. Chorti, and M. Reed, "Subcarrier scheduling for joint data transfer and key generation schemes in multicarrier systems," in *2019 IEEE Global Communications Conference (GLOBECOM)*, Dec 2019, pp. 1–6.

[15] M. Mitev, A. Chorti, V. Belmega, M. Reed, "Man-in-the-middle and denial of service attacks in wireless secret key generation," *Global Communications Conference (Globecom)*, Dec 2019.

[16] G. Rezgui, E. Belmega, and A. Chorti, "Mitigating jamming attacks using energy harvesting," *IEEE Wireless Communication Letters*, vol. 8, pp. 297–300, 2019.

[17] J. M. Hamamreh, H. M. Furqan, and H. Arslan, "Classifications and applications of physical layer security techniques for confidentiality: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 2, pp. 1773–1828, 2019.

[18] A. Mukherjee, "Physical-layer security in the internet of things: Sensing and communication confidentiality under resource constraints," in *Proc. IEEE*, vol. 103, no. 10, Oct 2015.

[19] A. Yener and S. Ulukus, "Wireless physical-layer security: Lessons learned from information theory," in *Proc. IEEE*, vol. 103, no. 10, Oct 2015.

[20] D. Karatzas, A. Chorti, N. M. White, and C. J. Harris, "Teaching old sensors new tricks: Archetypes of intelligence," *IEEE Sensors Journal*, vol. 7, no. 5, pp. 868–881, May 2007.

[21] "3GPP TR 33.825 V0.3.0, Study on the Security for 5G URLLC (Release 16)," 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects, available online https://www.3gpp.org/ftp/Specs/archive/33_series/33.825/.

[22] M. Latva-aho and K. Leppänen, "Key drivers and research challenges for 6G ubiquitous wireless intelligence," October 2019, published online by the University of Oulu.

[23] "International Network Generations Roadmap (INGR)," IEEE , available online https://futurenetworks.ieee.org/roadmap.

[24] F. Arute, K. Babbush, and *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, pp. 505–510, 2019.

[25] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1550–1573, Third 2014.

[26] "Elliptic Curves for Security," Internet Engineering Task Force (IETF) RFC 7748, available online https://tools.ietf.org/html/rfc7748.

[27] "PKCS 1: RSA Cryptography Specifications Version 2.2," Internet Engineering Task Force (IETF) RFC 8017, available online https://tools.ietf.org/html/rfc8017.

[28] "Diffie-Hellman Key Agreement Method," Internet Engineering Task Force (IETF) RFC 2631, available online https://tools.ietf.org/html/rfc2631.

[29] Y. Arjoune and S. Faruque, "Smart Jamming Attacks in 5G New Radio: A Review," in *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, 2020, pp. 1010–1015.

[30] A. Chorti and H. V. Poor, "Achievable secrecy rates in physical layer secure systems with a helping interferer," in *2012 International Conference on Computing, Networking and Communications (ICNC)*, Jan 2012, pp. 18–22.

[31] A. Weinand, M. Karrenbauer, and H. D. Schotten, "Security Solutions for Local Wireless Networks in Control Applications based on Physical Layer Security," *IFAC-PapersOnLine*, vol. 51, no. 10, pp. 32–39, 2018.

[32] Y. Kanaras and A. Chorti and M. Rodrigues and I. Darwazeh, "An optimum detection for a spectrally efficient non orthogonal FDM system," in *13th international OFDM-workshop*, Aug 2008, pp. 65–68.

[33] A. Chorti and H. V. Poor, "Faster than Nyquist interference assisted secret communication for OFDM systems," in *2011 Asilomar Conf. Signals, Systems and Computers (ASILOMAR)*, Nov 2011, pp. 183–187.

[34] A. Chorti, "Helping interferer physical layer security strategies for M-QAM and M-PSK systems," in *2012 46th Annual Conference on Information Sciences and Systems (CISS)*, March 2012, pp. 1–6.

[35] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[36] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[37] C. Ye, A. Reznik, and Y. Shah, "Extracting secrecy from jointly gaussian random variables," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jul 2006.

[38] A. Chorti, "A study of injection and jamming attacks in wireless secret sharing systems," in *Proc. Workshop on Communication Security (WCS), EUROCRYPT*, Mar 2017.

[39] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas, "Silicon physical random functions," in *Proc. of the 9th ACM Conference on Computer and Communications Security*, ser. CCS '02, NY, USA, 2002, p. 148–160.

[40] R. Pappu, "Physical one-way functions," *Science (New York)*, vol. 297, 10 2002.

[41] R. Maes and I. Verbauwhede, *Physically Unclonable Functions: A Study on the State of the Art and Future Research Directions*. Springer, 10 2010, pp. 3–37.

[42] The Tamarin team, "Tamarin-Prover Manual - Security Protocol Analysis in the Symbolic Model."

[43] M. Mitev, A. Chorti, and M. Reed, "Physical layer security in wireless networks with active eavesdroppers," Apr 2018, *invited poster.

[44] J. D. Vega Sánchez, L. Urquiza-Aguiar, and M. C. Paredes Paredes, "Physical layer security for 5g wireless networks: A comprehensive survey," in *2019 3rd Cyber Security in Networking Conference (CSNet)*, 2019, pp. 122–129.

[45] Y. Wu, A. Khisti, C. Xiao, G. Caire, K. Wong, and X. Gao, "A survey of physical layer security techniques for 5g wireless networks and challenges ahead," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 679–695, 2018.

[46] A. Riahi Sfar, E. Natalizio, Y. Challal, and Z. Chtourou, "A roadmap for security challenges in the internet of things," *Digital Communications and Networks*, vol. 4, no. 2, pp. 118 – 137, 2018.

[47] T. Xu, J. B. Wendt, and M. Potkonjak, "Security of IoT systems: Design challenges and opportunities," in *2014 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2014, pp. 417–423.

[48] I. Stellios, P. Kotzanikolaou, M. Psarakis, C. Alcaraz, and J. Lopez, "A survey of IoT-enabled cyberattacks: Assessing attack paths to critical infrastructures and services," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3453–3495, 2018.

[49] K. Piotrowski, P. Langendoerfer, and S. Peter, "How public key cryptography influences wireless sensor node lifetime," in *Proc. of the fourth ACM workshop on Security of ad hoc and sensor networks*, 01 2006, pp. 169–176.

[50] A. Teniou and B. A. Bensaber, "Efficient and dynamic elliptic curve qu-vanstone implicit certificates distribution scheme for vehicular cloud networks," *Security and Privacy*, Jan. 2018.

[51] J. Wang, Y. Shao, Y. Ge, and R. Yu, "A survey of vehicle to everything (V2X) testing," *Sensors*, vol. 19, p. 334, 01 2019.

[52] C. Herder, M. Yu, F. Koushanfar, and S. Devadas, "Physical unclonable functions and applications: A tutorial," *Proc. of the IEEE*, vol. 102, no. 8, pp. 1126–1141, Aug 2014.

[53] J. Delvaux, R. Peeters, D. Gu, and I. Verbauwhede, "A Survey on Lightweight Entity Authentication with Strong PUFs," Cryptology ePrint Archive, Report 2014/977, 2014.

[54] C. Böhm and M. Hofer, *Physical Unclonable Functions in Theory and Practice*. Springer Publishing Company, Incorporated, 2012.

[55] J. Delvaux, R. Peeters, D. Gu, and I. Verbauwhede, "A Survey on Lightweight Entity Authentication with Strong PUFs," *ACM Computing Surveys*, vol. 48, no. 2, Oct. 2015.

[56] N. Wang, P. Wang, A. Alipour-Fanid, L. Jiao, and K. Zeng, "Physical-Layer Security of 5G Wireless Networks for IoT: Challenges and Opportunities," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8169–8181, 2019.

[57] C. Wang, Z. Li, J. Shi, J. Si, and D. W. K. Ng, "Physical layer security of vehicular networks: A stochastic geometry approach," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2020, pp. 1–7.

[58] B. Li, Z. Fei, Y. Zhang, and M. Guizani, "Secure UAV communication networks over 5G," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 114–120, 2019.

[59] Q. Wu, W. Mei, and R. Zhang, "Safeguarding wireless network with UAVs: A physical layer security perspective," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 12–18, 2019.

[60] N. Miloslavskaya and A. Tolstoy, “Internet of things: information security challenges and solutions,” *Cluster Computing*, vol. 22, 03 2019.

[61] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, “Key generation from wireless channels: A review,” *IEEE Access*, vol. 4, pp. 614–626, 2016.

[62] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, “Radio-telepathy: extracting a secret key from an unauthenticated wireless channel,” *Proc. 14th Annual International Conference on Mobile Computing Networking (MobiCom)*, pp. 128–139, Sep. 2008.

[63] S. T. Ali, V. Sivaraman, and D. Ostry, “Eliminating reconciliation cost in secret key generation for body-worn health monitoring devices,” *IEEE Transactions on Mobile Computing*, vol. 13, no. 12, pp. 2763–2776, Dec 2014.

[64] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, “Ultra-reliable and low-latency communications in 5g downlink: Physical layer aspects,” *IEEE Wireless Communications*, vol. 25, no. 3, pp. 124–130, 2018.

[65] P. Popovski, C. Stefanovic, J. J. Nielsen, E. de Carvalho, M. Angjelichinoski, K. F. Trillingsgaard, and A. Bana, “Wireless access in ultra-reliable low-latency communication (URLLC),” *IEEE Transactions on Communications*, vol. 67, no. 8, pp. 5783–5801, 2019.

[66] H. Elayan, O. Amin, R. M. Shubair, and M. Alouini, “Terahertz communication: The opportunities of wireless technology beyond 5G,” in *2018 International Conference on Advanced Communication Technologies and Networking (CommNet)*, 2018, pp. 1–5.

[67] G. Caruso, F. Nucci, O. P. Gordo, S. Rizou, J. Magen, G. Agapiou, and P. Trakadas, “Embedding 5G solutions enabling new business scenarios in media and entertainment industry,” in *2019 IEEE 2nd 5G World Forum (5GWF)*, 2019, pp. 460–464.

[68] M. Maheswaran and E. Badidi, Eds., *Handbook of Smart Cities, Software Services and Cyber Infrastructure*. Springer, 2018.

[69] A. Garnaev and W. Trappe, "Bargaining Over the Fair Trade-Off Between Secrecy and Throughput in OFDM Communications," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 242–251, Jan 2017.

[70] Y. Arjoune and S. Faruque, "Smart jamming attacks in 5G new radio: A review," in *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, 2020, pp. 1010–1015.

[71] B. Kim, Y. E. Sagduyu, K. Davaslioglu, T. Erpek, and S. Ulukus, "How to make 5G communications "invisible": Adversarial machine learning for wireless privacy," 2020.

[72] V. N. Swamy, N. Naderializadeh, V. N. Ekambaram, S. Talwar, and A. Sahai, "Monitoring under-modeled rare events for URLLC," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2019, pp. 1–5.

[73] E. V. Belmega and A. Chorti, "Energy harvesting in secret key generation systems under jamming attacks," in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[74] E. V. Belmega and A. Chorti, "Protecting secret key generation systems against jamming: Energy harvesting and channel hopping approaches," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2611–2626, Nov 2017.

[75] C. Popper, M. Strasser, and S. Capkun, "Anti-jamming broadcast communication using uncoordinated spread spectrum techniques," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 5, pp. 703–715, 2010.

[76] X. Wang, J. Wang, Y. Xu, J. Chen, L. Jia, X. Liu, and Y. Yang, "Dynamic spectrum anti-jamming communications: Challenges and opportunities," *IEEE Communications Magazine*, vol. 58, no. 2, pp. 79–85, 2020.

[77] J. P. Vilela, M. Bloch, J. Barros, and S. W. McLaughlin, "Wireless secrecy regions with friendly jamming," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 2, pp. 256–266, 2011.

[78] I. Stanojev and A. Yener, "Improving secrecy rate via spectrum leasing for friendly jamming," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 134–145, 2013.

[79] Z. Liu, H. Liu, W. Xu, and Y. Chen, "Exploiting jamming-caused neighbor changes for jammer localization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 3, pp. 547–555, 2012.

[80] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," *Proc. of the National Academy of Sciences*, vol. 114, no. 1, pp. 19–26, 2017.

[81] L. Sun and Q. Du, "A review of physical layer security techniques for internet of things: Challenges and solutions," *Entropy*, vol. 20, p. 730, 2018.

[82] M. Zoli, A. N. Barreto, S. Köpsell, P. Sen, and G. Fettweis, "Physical-layer-security box: a concept for time-frequency channel-reciprocity key generation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, pp. 1–24, 2020.

[83] B. Wu, J. Chen, J. Wu, and M. Cardei, *A Survey of Attacks and Countermeasures in Mobile Ad Hoc Networks*. Springer, Jan 2007, pp. 103–135.

[84] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. of the IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.

[85] A. Hassan, W. Stark, J. Hershey, and S. Chennakeshu, "Cryptographic key agreement for mobile radio," *Digital Signal Process*, vol. 6, no. 10, pp. 207–212, Oct 1996.

[86] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *Proc. of the 15th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 321–332.

[87] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. USA: Prentice Hall PTR, 2001.

[88] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe, "Using the physical layer for wireless authentication in time-variant channels," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2571–2579, 2008.

[89] L. Jin, S. Zhang, Y. Lou, X. Xu, and Z. Zhong, "Secret key generation with cross multiplication of two-way random signals," *IEEE Access*, vol. 7, pp. 113 065–113 080, 2019.

[90] J. K. Tugnait, Lang Tong, and Zhi ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 17–28, May 2000.

[91] S. Mathur, R. Miller, A. Varshavsky, W. Trappe, and N. Mandayam, "Proximate: Proximity-based secure pairing using ambient wireless signals," in *Proc. of the 9th International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '11. New York, NY, USA: Association for Computing Machinery, 2011, pp. 211–224.

[92] H. Liu, Y. Wang, J. Yang, and Y. Chen, "Fast and practical secret key extraction by exploiting channel response," in *2013 Proc. of IEEE INFOCOM*, April 2013, pp. 3048–3056.

[93] Wei Xi, Xiang-Yang Li, Chen Qian, Jinsong Han, Shaojie Tang, Jizhong Zhao, and Kun Zhao, "KEEP: Fast secret key extraction protocol for D2D communication," in *2014 IEEE 22nd International Symposium of Quality of Service (IWQoS)*, May 2014, pp. 350–359.

[94] S. Bakşi and D. C. Popescu, "Secret key generation with precoding and role reversal in mimo wireless systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3104–3112, June 2019.

[95] H. Jin, K. Huang, S. Xiao, Y. Lou, X. Xu, and Y. Chen, "A two-layer secure quantization algorithm for secret key generation with correlated eavesdropping channel," *IEEE Access*, vol. 7, pp. 26 480–26 487, 2019.

[96] W. Jakes, *Microwave Mobile Communications*. Wiley-IEEE Press, May 1994.

[97] W. Trappe, "The challenges facing physical layer security," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 16–20, June 2015.

[98] C. Chen and M. A. Jensen, "Secret key establishment using temporally and spatially correlated wireless channel coefficients," *IEEE Transactions on Mobile Computing*, vol. 10, no. 2, pp. 205–215, Feb 2011.

[99] J. Zhang, A. Marshall, R. Woods, and T. Q. Duong, "Efficient key generation by exploiting randomness from channel responses of individual OFDM subcarriers," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2578–2588, June 2016.

[100] J. Zhang, B. He, T. Q. Duong, and R. Woods, "On the key generation from correlated wireless channels," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 961–964, April 2017.

[101] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[102] A. Chorti and E. V. Belmega, "Secret key generation in Rayleigh block fading AWGN channels under jamming attacks," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.

[103] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *Proc. of the 15th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '09. New York, NY, USA: ACM, 2009, pp. 321–332.

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

[105] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-theoretically secret key generation for fading wireless channels," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 240–254, June 2010.

[106] C. Huth, R. Guillaume, T. Strohm, P. Duplys, I. A. Samuel, and T. Gneysu, "Information reconciliation schemes in physical-layer security," *Computational Networks*, vol. 109, no. P1, p. 84–104, Nov. 2016.

[107] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[108] C. Saiki and A. Chorti, "A novel physical layer authenticated encryption protocol exploiting shared randomness," in *Proc. IEEE Conference on Communication Networking Security (CNS)*, Florence, Italy, 2015.

[109] L. Guyue, Z. Zhang, Y. Yu, and A. Hu, "A hybrid information reconciliation method for physical layer key generation," *Entropy*, vol. 21, p. 688, 07 2019.

[110] P. Treeviriyanupab, P. Sangwongngam, K. Sripimanwat, and O. Sangaroon, "BCH-based Slepian-Wolf coding with feedback syndrome decoding for quantum key reconciliation," in *2012 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, May 2012, pp. 1–4.

[111] J. Etesami and W. Henkel, "LDPC code construction for wireless physical-layer key reconciliation," in *2012 1st IEEE International Conference on Communications in China (ICCC)*, Aug 2012, pp. 208–213.

[112] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1915–1923, Nov 1995.

[113] F. Zhan and N. Yao, "On the using of discrete wavelet transform for physical layer key generation," *Ad Hoc Networks*, vol. 64, pp. 22 – 31, 2017.

[114] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[115] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[116] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *2006 IEEE International Symposium on Information Theory*, 2006, pp. 356–360.

[117] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, 2008.

[118] N. Yang, M. Elkashlan, T. Q. Duong, J. Yuan, and R. Malaney, "Optimal Transmission With Artificial Noise in MISOME Wiretap Channels," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2170–2181, 2016.

[119] N. Yang, S. Yan, J. Yuan, R. Malaney, R. Subramanian, and I. Land, "Artificial noise: Transmission optimization in multi-input single-output wiretap channels," *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1771–1783, 2015.

[120] M. Hayashi and R. Matsumoto, "Construction of wiretap codes from ordinary channel codes," in *2010 IEEE International Symposium on Information Theory*, 2010, pp. 2538–2542.

[121] S. M. Shah and V. Sharma, "Enhancing secrecy rates in a wiretap channel," *Digital Communications and Networks*, vol. 6, no. 1, pp. 129 – 135, 2020.

[122] S. El Rouayheb, E. Soljanin, and A. Sprintson, "Secure network coding for wiretap networks of type II," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1361–1371, 2012.

[123] W. Yang, R. F. Schaefer, and H. V. Poor, "Finite-blocklength bounds for wiretap channels," in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 3087–3091.

[124] ——, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4069–4093, 2019.

[125] "3GPP TR 36.211, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects, available online https://www.3gpp.org/dynareport/36211.htm.

[126] "The Transport Layer Security (TLS) Protocol Version 1.3," Rescorla, E., RFC 8446 (2018), available online https://rfc-editor.org/rfc/rfc8446.txt.

[127] N. Aviram, K. Gellert, and T. Jager, "Session Resumption Protocols and Efficient Forward Security for TLS 1.3 0-RTT," Cryptology ePrint Archive, Report 2019/228, 2019.

[128] A. Babaei and G. Schiele, “Physical unclonable functions in the internet of things: State of the art and open challenges,” in *Sensors*, 2019.

[129] A. Maiti, I. Kim, and P. Schaumont, “A robust physical unclonable function with enhanced challenge-response set,” *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 333–345, Feb 2012.

[130] M. B. Yassein, S. Aljawarneh, E. Qawasmeh, W. Mardini, and Y. Khamayseh, “Comprehensive study of symmetric key and asymmetric key encryption algorithms,” in *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1–7.

[131] V. Mavroeidis, K. Vishi, M. D., and A. Jøsang, “The impact of quantum computing on present cryptography,” *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 3, 2018.

[132] A. S. Dr. Prerna Mahajan, “A Study of Encryption Algorithms AES, DES and RSA for Security,” *Global Journal of Computer Science and Technology*, 2013.

[133] P. Hamalainen, T. Alho, M. Hannikainen, and T. D. Hamalainen, “Design and implementation of low-area and low-power AES encryption hardware core,” in *9th EUROMICRO Conference on Digital System Design (DSD’06)*, 2006, pp. 577–583.

[134] B. Liu and B. M. Baas, “Parallel AES encryption engines for many-core processor arrays,” *IEEE Transactions on Computers*, vol. 62, no. 3, pp. 536–547, 2013.

[135] Federal Information Processing Standards Publication 197. United States National Institute of Standards and Technology (NIST), “Announcing the ADVANCED ENCRYPTION STANDARD (AES),” 2001.

[136] “Advanced Encryption Standard (AES) Key Wrap with Padding Algorithm,” Internet Engineering Task Force (IETF) RFC 5649, available online https://tools.ietf.org/html/rfc5649.

[137] M. Bellare, J. Kilian, and P. Rogaway, "The security of the cipher block chaining message authentication code," *Journal of Computer and System Sciences*, vol. 61, pp. 362–399, 12 2000.

[138] National Institute of Standards and Technology (NIST), "Message Authentication Codes (MAC)," National Institute of Standards and Technology (NIST), available online https://csrc.nist.gov/Projects/Message-Authentication-Codes.

[139] M. Naor and M. Yung, "Universal one-way hash functions and their cryptographic applications," in *Proc. of the Twenty-First Annual ACM Symposium on Theory of Computing*, ser. STOC '89, NY, USA, 1989, p. 33–43.

[140] G. Tsudik, "Message authentication with one-way hash functions," *SIGCOMM Computing Communications*, vol. 22, no. 5, p. 29–38, Oct. 1992.

[141] H. Krawczyk, M. Bellare, R. Canetti, "HMAC: Keyed-Hashing for Message Authentication," Internet Engineering Task Force (IETF), available online https://www.rfc-editor.org/rfc/rfc2104, 1997.

[142] T. Ylonen, SSH Communications Security Corp, C. Lonvick, Cisco Systems, "HMAC: Keyed-Hashing for Message Authentication," Internet Engineering Task Force (IETF), available online https://tools.ietf.org/html/rfc4253, 2006.

[143] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, 2002.

[144] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas, "Silicon physical random functions," in *Proc. of the 9th ACM Conference on Computer and Communications Security*, ser. CCS '02. New York, NY, USA: ACM, 2002, pp. 148–160.

[145] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *2007 44th ACM/IEEE Design Automation Conference*, June 2007, pp. 9–14.

[146] Daihyun Lim, J. W. Lee, B. Gassend, G. E. Suh, M. van Dijk, and S. Devadas, "Extracting secret keys from integrated circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 13, no. 10, pp. 1200–1205, Oct 2005.

[147] C. Marchand, L. Bossuet, U. Mureddu, N. Bochard, A. Cherkaoui, and V. Fischer, "Implementation and characterization of a physical unclonable function for IoT: A case study with the TERO-PUF," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 1, pp. 97–109, Jan 2018.

[148] J. Guajardo, S. S. Kumar, G.-J. Schrijen, and P. Tuyls, "FPGA intrinsic PUFs and their use for IP protection," in *Proc. of the 9th International Workshop on Cryptographic Hardware and Embedded Systems*, ser. CHES '07. Berlin, Heidelberg: Springer-Verlag, 2007, p. 63–80.

[149] J. Aarestad, P. Ortiz, D. Acharyya, and J. Plusquellic, "Help: A hardware-embedded delay PUF," *IEEE Design Test*, vol. 30, no. 2, pp. 17–25, April 2013.

[150] S. Devadas and M.-D. M. Yu, "Recombination of physical unclonable functions," in *Proc. of the 35th Annual Government Microcircuit Application Critical Technology Conference*, 03 2010.

[151] D. Ganta and L. Nazhandali, "Easy-to-build arbiter physical unclonable function with enhanced challenge/response set," in *International Symposium on Quality Electronic Design (ISQED)*, March 2013, pp. 733–738.

[152] U. Chatterjee, R. S. Chakraborty, and D. Mukhopadhyay, "A PUF-Based Secure Communication Protocol for IoT," *ACM Transactions on Embedded Computing Systems*, vol. 16, no. 3, pp. 67:1–67:25, Apr. 2017.

[153] M. N. Aman, M. H. Basheer, and B. Sikdar, "Two-Factor Authentication for IoT With Location Information," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3335–3351, April 2019.

[154] M. H. Mahalat, S. Saha, A. Mondal, and B. Sen, "A PUF based Light Weight Protocol for Secure WiFi Authentication of IoT devices," in *2018 8th International Symposium on Embedded Computing and System Design (ISED)*, Dec 2018, pp. 183–187.

[155] A. Braeken, "PUF based authentication protocol for IoT," *Symmetry*, vol. 10, 2018.

[156] Y. Yilmaz, S. R. Gunn, and B. Halak, "Lightweight PUF-Based Authentication Protocol for IoT Devices," in *2018 IEEE 3rd International Verification and Security Workshop (IVSW)*, July 2018, pp. 38–43.

[157] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," Cryptology ePrint Archive, Report 2003/235, 2003.

[158] J. W. Byun, "End-to-end authenticated key exchange based on different physical unclonable functions," *IEEE Access*, vol. 7, pp. 102 951–102 965, 2019.

[159] P. Gope and B. Sikdar, "Lightweight and Privacy-Preserving Two-Factor Authentication Scheme for IoT Devices," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 580–589, 2019.

[160] J. W. Byun, "PDAKE: a provably secure PUF-based device authenticated key exchange in cloud setting," *IEEE Access*, vol. 7, pp. 181 165–181 177, 2019.

[161] B. D. Deebak, F. Al-Turjman, M. Aloqaily, and O. Alfandi, "An Authentic-Based Privacy Preservation Protocol for Smart e-Healthcare Systems in IoT," *IEEE Access*, vol. 7, pp. 135 632–135 649, 2019.

[162] P. Gope, J. Lee, and T. Q. S. Quek, "Lightweight and practical anonymous authentication protocol for RFID systems using physically unclonable functions," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2831–2843, 2018.

[163] B. Fuller, L. Reyzin, and A. Smith, "When are fuzzy extractors possible?" *IEEE Transactions on Information Theory*, pp. 1–1, 2020.

[164] J. Delvaux, D. Gu, D. Schellekens, and I. Verbauwhede, "Helper data algorithms for PUF-based key generation: Overview and analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 6, pp. 889–902, 2015.

[165] P. Maurya and S. Bagchi, "A secure PUF-based unilateral authentication scheme for RFID system," *Wireless Personal Communications*, vol. 103, 05 2018.

[166] M. Yu, M. Hiller, J. Delvaux, R. Sowell, S. Devadas, and I. Verbauwhede, "A Lockdown Technique to Prevent Machine Learning on PUFs for Lightweight Authentication," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 2, no. 3, pp. 146–159, July 2016.

[167] W. Che, M. Martin, G. Pocklassery, V. K. Kajuluri, F. Saqib, and J. F. Plusquellic, "A privacy-preserving, mutual puf-based authentication protocol," *Cryptography*, vol. 1, p. 3, 2016.

[168] J. Calhoun, C. Minwalla, C. Helmich, F. Saqib, W. Che, and J. Plusquellic, "Physical unclonable function (PUF)-based e-cash transaction protocol (PUF-Cash)," *Cryptography*, vol. 3, p. 18, 07 2019.

[169] M. N. Aman, K. C. Chua, and B. Sikdar, "Mutual authentication in IoT systems using physical unclonable functions," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1327–1340, Oct 2017.

[170] W. Diffie, P. C. Van Oorschot, and M. J. Wiener, "Authentication and authenticated key exchanges," *Design Codes Cryptography*, vol. 2, no. 2, p. 107–125, Jun. 1992.

[171] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare internet of things: A survey of emerging technologies," *IEEE Communications Surveys Tutorials*, vol. 22, no. 2, pp. 1121–1167, 2020.

[172] F. John Dian, R. Vahidnia, and A. Rahmati, "Wearables and the internet of things (IoT), applications, opportunities, and challenges: A survey," *IEEE Access*, vol. 8, pp. 69 200–69 211, 2020.

[173] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[174] Y. Li, Y. Zhuang, X. Hu, Z. Gao, J. Hu, L. Chen, Z. He, L. Pei, K. Chen, M. Wang, X. Niu, R. Chen, J. Thompson, F. M. Ghannouchi, and N. El-Sheimy, "Location-Enabled IoT (LE-IoT): A Survey of Positioning Techniques, Error Sources, and Mitigation," *ArXiv*, vol. abs/2004.03738, 2020.

[175] C. Shao, Y. Kim, and W. Lee, "Zero-effort proximity detection with zigbee," *IEEE Communications Letters*, pp. 1–1, 2020.

[176] G. Welch and G. Bishop, "An introduction to the Kalman filter," University of North Carolina at Chapel Hill, USA, Tech. Rep., 1995.

[177] J. B. Andersen, T. S. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *IEEE Communications Magazine*, vol. 33, no. 1, pp. 42–49, 1995.

[178] H. A. Nguyen, H. Guo, and K. Low, "Real-time estimation of sensor node's position using particle swarm optimization with log-barrier constraint," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 11, pp. 3619–3628, 2011.

[179] M. Mitev and S. Stoyneva, "Indoor positioning for smart ambient assisted living services," Master's thesis, Aalborg University, 2016.

[180] G. Li, E. Geng, Z. Ye, Y. Xu, J. Lin, and Y. Pang, "Indoor positioning algorithm based on the improved rssi distance model," *Sensors*, vol. 18, p. 2820, 08 2018.

[181] B. Blanchet, *Modeling and Verifying Security Protocols with the Applied Pi Calculus and ProVerif*. IEEE, 2016.

[182] M. Burrows, M. Abadi, and R. Needham, "A logic of authentication," *ACM Transactions on Computing Systems*, vol. 8, no. 1, p. 18–36, Feb. 1990.

[183] P. F. Syverson and P. C. van Oorschot, "On unifying some cryptographic protocol logics," in *Proc. of 1994 IEEE Computer Society Symposium on Research in Security and Privacy*, 1994, pp. 14–28.

[184] L. Gong, R. Needham, and R. Yahalom, "Reasoning about belief in cryptographic protocols," in *Proceedings. 1990 IEEE Computer Society Symposium on Research in Security and Privacy*, 1990, pp. 234–248.

[185] M. Abadi and M. R. Tuttle, "A semantics for a logic of authentication (extended abstract)," in *Proc. of the Tenth Annual ACM Symposium on Principles of Distributed Computing*, ser. PODC '91, 1991, p. 201–216.

[186] W. Mao and C. Boyd, "Towards formal analysis of security protocols," in *Proc. of Computer Security Foundations Workshop VI*, 1993, pp. 147–158.

[187] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78 729–78 757, 2019.

[188] G. Pasolini, A. Guerra, F. Guidi, N. Decarli, and D. Dardari, "Crowd-Based Cognitive Perception of the Physical World: Towards the Internet of Senses," *Sensors*, vol. 20, p. 2437, 04 2020.

[189] O. Kanhere, S. Ju, Y. Xing, and T. S. Rappaport, "Map-assisted millimeter wave localization for accurate position location," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[190] "Intrinsic-id company," https://www.intrinsic-id.com/sram-puf.

[191] "ICTK holdings corporation," https://ictk-puf.com/puf-technology.

[192] A. Van Herrewege, S. Katzenbeisser, R. Maes, R. Peeters, A.-R. Sadeghi, I. Verbauwhede, and C. Wachsmann, "Reverse fuzzy extractors: Enabling lightweight mutual authentication for PUF-enabled RFIDs," in *Financial Cryptography and Data Security*, A. D. Keromytis, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 374–389.

[193] Y. Gao, Y. Su, L. Xu, and D. C. Ranasinghe, "Lightweight (reverse) fuzzy extractor with multiple reference PUF responses," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 7, pp. 1887–1901, 2019.

[194] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," in *2014 IEEE International Symposium on Information Theory*, 2014, pp. 1136–1140.

[195] "TOTP: Time-Based One-Time Password Algorithm," D. M'Raihi, S. Machani, M. Pei, J. Rydell. RFC 6238 (2011), available online https://tools.ietf.org/html/rfc6238.

[196] D. Dolev and A. Yao, "On the security of public key protocols," *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 198–208, 1983.

[197] H. Cha, K.-H. Kim, and S. Yoo, "Lbp: A secure and efficient network bootstrapping protocol for 6lowpan," *Proc. of the 5th International Conference on Ubiquitous Information Management and Communication, ICUIMC 2011*, p. 54, 01 2011.

[198] N. Anjum, Z. Yang, H. Saki, M. Kiran, and M. Shikh-Bahaei, "Device-to-device (D2D) communication as a bootstrapping system in a wireless cellular network," *IEEE Access*, vol. 7, pp. 6661–6678, 2019.

[199] P. Gope and B. Sikdar, "Lightweight and Privacy-Preserving Two-Factor Authentication Scheme for IoT Devices," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 580–589, 2019.

[200] P. Gope and B. Sikdar, "Privacy-aware authenticated key agreement scheme for secure smart grid communication," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3953–3962, 2019.

[201] T. Miki, N. Miura, H. Sonoda, K. Mizuta, and M. Nagata, "A Random Interrupt Dithering SAR Technique for Secure ADC Against Reference-Charge Side-Channel Attack," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 1, pp. 14–18, 2020.

[202] C. Bouillaguet, F. Martinez, and J. Sauvage, "Predicting the PCG Pseudo-Random Number Generator In Practice," Jun. 2020, working paper or preprint.

[203] C. Huth, D. Becker, J. G. Merchan, P. Duplys, and T. Güneysu, "Securing Systems With Indispensable Entropy: LWE-Based Lossless Computational Fuzzy Extractor for the Internet of Things," *IEEE Access*, vol. 5, pp. 11 909–11 926, 2017.

[204] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1915–1923, 1995.

[205] G. Lowe, "A hierarchy of authentication specifications," in *Proc. of 10th Computer Security Foundations Workshop*, 1997, pp. 31–43.

[206] Y. Lee, L. Batina, D. Singelée, and I. Verbauwhede, "Low-cost untraceable authentication protocols for RFID," in *ACM*, 01 2010, pp. 55–64.

[207] Q. Jiang, J. Ma, F. Wei, Y. Tian, J. Shen, and Y. Yang, "An untraceable temporal-credential-based two-factor authentication scheme using ecc for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 76, 10 2016.

[208] J. Liu, Z. Zhang, X. Chen, and K. S. Kwak, "Certificateless remote anonymous authentication schemes for wirelessbody area networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 332–342, 2014.

[209] Shiuh-Jeng Wang, "Anonymous wireless authentication on a portable cellular mobile system," *IEEE Transactions on Computers*, vol. 53, no. 10, pp. 1317–1329, 2004.

[210] C.-T. Li, M.-S. Hwang, and C.-Y. Liu, "An electronic voting protocol with deniable authentication for mobile ad hoc networks," *Computing Communications*, vol. 31, no. 10, p. 2534–2540, Jun. 2008.

[211] J. Camenisch, S. Hohenberger, M. Kohlweiss, A. Lysyanskaya, and M. Meyerovich, "How to win the clonewars: Efficient periodic n-times anonymous authentication," in *Proc. of the 13th ACM Conference on Computer and Communications Security*, ser. CCS '06, NY, USA, 2006, p. 201–210.

[212] M. Akhlaq, B. Aslam, M. A. Khan, and M. N. Jafri, "Comparative analysis of IEEE 802.1x authentication methods," in *Proc. of the 11th Conference on 11th WSEAS International Conference on Communications - Volume 11*, ser. ICCOM'07. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2007, p. 1–6.

[213] A. Chiornită, L. Gheorghe, and D. Rosner, "A practical analysis of EAP authentication methods," in *9th RoEduNet IEEE International Conference*, June 2010, pp. 31–35.

[214] S. Ahmad, A. H. Mir, and G. R. Beigh, "Latency evaluation of extensible authentication protocols in wlans," in *2011 Fifth IEEE International Conference on Advanced Telecommunication Systems and Networks (ANTS)*, Dec 2011, pp. 1–5.

[215] P. Gope and B. Sikdar, "Lightweight and Privacy-Preserving Two-Factor Authentication Scheme for IoT Devices," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 580–589, Feb 2019.

[216] U. Chatterjee, R. S. Chakraborty, and D. Mukhopadhyay, "A PUF-Based Secure Communication Protocol for IoT," *ACM Transactions on Embedded Computing Systems*, vol. 16, no. 3, Apr. 2017.

[217] A. Braeken, "PUF Based Authentication Protocol for IoT," *Symmetry*, vol. 10, no. 352, 2018.

[218] P. Gope and B. Sikdar, "Privacy-aware authenticated key agreement scheme for secure smart grid communication," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3953–3962, 2019.

[219] Q. Jiang, X. Zhang, N. Zhang, Y. Tian, X. Ma, and J. Ma, "Two-factor authentication protocol using physical unclonable function for IoV," in *2019 IEEE/CIC International Conference on Communications in China (ICCC)*, 2019, pp. 195–200.

[220] A. Ometov, P. Masek, L. Malina, R. Florea, J. Hosek, S. Andreev, J. Hajny, J. Niutanen, and Y. Koucheryavy, "Feasibility characterization of cryptographic primitives for constrained (wearable) IoT devices," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, March 2016, pp. 1–6.

[221] J. Cho and W. Sung, "Efficient Software-Based Encoding and Decoding of BCH Codes," *IEEE Transactions on Computers*, vol. 58, no. 7, pp. 878–889, July 2009.

[222] M. Bellare and C. Namprempre, "Authenticated encryption: Relations among notions and analysis of the generic composition paradigm," Cryptology ePrint Archive, Report 2000/025, 2000.

[223] T. Krovetz and P. Rogaway, "The software performance of authenticated-encryption modes," in *Fast Software Encryption*, A. Joux, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 306–327.

[224] S. Koteshwara and A. Das, "Comparative study of authenticated encryption targeting lightweight IoT applications," *IEEE Design Test*, vol. 34, no. 4, pp. 26–33, Aug 2017.

[225] E. Steinbach, S. Hirche, M. Ernst, F. Brandi, R. Chaudhari, J. Kammerl, and I. Vittorias, "Haptic communications," *Proc. of the IEEE*, vol. 100, no. 4, pp. 937–956, 2012.

[226] A. Aijaz, M. Dohler, A. H. Aghvami, V. Friderikos, and M. Frodigh, "Realizing the Tactile Internet: Haptic Communications over Next Generation 5G Cellular Networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 82–89, 2017.

[227] K. Antonakoglou, X. Xu, E. Steinbach, T. Mahmoodi, and M. Dohler, "Toward Haptic Communications Over the 5G Tactile Internet," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3034–3059, 2018.

[228] Peter W. Christensen and Anders Klarbring, *An Introduction to Structural Optimization*. Netherlands: Springer, Dordrecht, 2008.

[229] Zhi-Quan Luo and Wei Yu, "An introduction to convex optimization for communications and signal processing," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1426–1438, 2006.

[230] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[231] R. T. Rockafellar, "Lagrange multipliers and optimality," *SIAM Rev.*, vol. 35, no. 2, p. 183–238, Jun. 1993.

[232] W. Karush, "Minima of functions of several variables with inequalities as side conditions," Master's thesis, Department of Mathematics, University of Chicago, Chicago, IL, USA, 1939.

[233] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in *Proc. of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, Calif.: University of California Press, 1951, pp. 481–492.

[234] R. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations*, R. Miller and J. Thatcher, Eds. Plenum Press, 1972, pp. 85–103.

[235] H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Springer-Verlag Berlin Heidelberg, 2004.

[236] J. Gartner, "On large deviations from the invariant measure," *Theory of Probability & Its Applications*, vol. 22, no. 1, pp. 24–39, 1977.

[237] R. S. Ellis, "Large deviations for a general class of random vectors," in *Annual Probabilities*, 1984.

[238] A. Dembo and O. Zeitouni, *Large Deviation Principles and Applications*. USA: Jones and Bartlett Publishers, 1993.

[239] H. Touchette, *The Large Deviation Approach to Statistical Mechanics*. UK: School of Mathematical Science, Queen Mary University of London, 2009.

[240] Dapeng Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 630–643, July 2003.

[241] Cheng-Shang Chang and Joy A. Thomas, "Effective bandwidth in high-speed digital networks," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, Aug. 1995.

[242] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, pp. 641–657, 1999.

[243] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation over wireless links," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 3058–3068, 2007.

[244] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation for multichannel communications over wireless links," *IEEE Transactions on Wireless Communications*, vol. 6, no. 12, pp. 4349–4360, 2007.

[245] W. Yu, L. Musavian, and Q. Ni, "Statistical delay QoS driven energy efficiency and effective capacity tradeoff for uplink multi-user multi-carrier systems," *IEEE Transactions on Communications*, vol. 65, no. 8, pp. 3494–3508, 2017.

[246] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, "Ultra-reliable and low-latency communications in 5g downlink: Physical layer aspects," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 124–130, 2018.

[247] H. Khan, M. M. Butt, S. Samarakoon, P. Sehier, and M. Bennis, "Deep learning assisted csi estimation for joint urllc and embb resource allocation," 2020.

[248] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[249] C. E. Shannon, "Communication in the presence of noise," *Proc. of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[250] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.

[251] F. Yilmaz and M. Alouini, "Outage capacity of multicarrier systems," in *2010 17th International Conference on Telecommunications*, 2010, pp. 260–265.

[252] D. Tse and P. Viswanath, Eds., *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[253] D. P. K. and V. Sharma, "Finite blocklength rates over a fading channel with csit and csir," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7.

[254] P. Mary, J. Gorce, A. Unsal, and H. V. Poor, "Finite blocklength information theory: What is the practical impact on wireless communications?" in *2016 IEEE Globecom Workshops (GC Wkshps)*, 2016, pp. 1–6.

[255] R. Bellman and R. Kalaba, *Dynamic Programming and Modem Control Theory*. NY: Academic Press, 1966.

[256] O. Hölder, "Ueber einen mittelwertsatz," *Göttinger Nachr.*, pp. 38–47, 1889.

[257] J. Jensen, "Sur les fonctions convexes et les inégualités entre les valeurs moyennes," *Acta Mathematica*, pp. 175–193, 1906.

[258] H. Yang and M. Alouini, *Order Statistics in Wireless Communications*. NY: Cambridge University Press, 2011.

[259] S. Martello and P. Toth, *Knapsack problems: algorithms and computer implementations*. NY: John Wiley and Sons, 1990.

[260] V. Vazirani, *Approximation Algorithms*. Springer-Verlag Berlin Heidelberg, 2003.

[261] Cheng-Shang Chang, "Stability, queue length, and delay of deterministic and stochastic queueing networks," *IEEE Transactions on Automatic Control*, vol. 39, no. 5, pp. 913–931, May 1994.

[262] T. Abrão, S. Yang, L. D. H. Sampaio, P. J. E. Jeszensky, and L. Hanzo, "Achieving Maximum Effective Capacity in OFDMA Networks Operating Under Statistical Delay Guarantee," *IEEE Access*, vol. 5, pp. 14 333–14 346, 2017.

[263] C. Shahriar, M. La Pan, M. Lichtman, T. C. Clancy, R. McGwier, R. Tandon, S. Sodagari, and J. H. Reed, "PHY-Layer Resiliency in OFDM Communications: A Tutorial," *IEEE Communications Surveys Tutorials*, vol. 17, no. 1, pp. 292–314, Firstquarter 2015.

[264] C. Carlson, V. Nguyen, S. Hitefield, T. O'Shea, and T. Clancy, "Measuring smart jammer strategy efficacy over the air," in *2014 IEEE Conference on Communications and Network Security*, Oct 2014, pp. 7–13.

[265] S. Fang, Y. Liu and P. Ning , "Wireless communications under broadband reactive jamming attacks," *IEEE Transactions on Dependable Secure Computing*, vol. 13, no. 3, pp. 394 – 408, May 2016.

[266] M. Spuhler, D. Giustiniano, V. Lenders, M. Wilhelm and J. B. Schmitt, "Detection of reactive jamming in DSSS-based wireless communications," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1593 – 1603, May 2014.

[267] N. An and S. Weber, "Efficiency and detectability of random reactive jamming in wireless networks," in *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, June 2018, pp. 1–9.

[268] H. B. Salameh, S. Almajali, M. Ayyash, and H. Elgala, "Securing delay-sensitive cognitive radio IoT communications under reactive jamming attacks: Spectrum assignment perspective," in *2018 Fifth International Conference on Software Defined Systems (SDS)*, April 2018, pp. 20–24.

[269] X. Tang, P. Ren, Y. Wang, Q. Du, and L. Sun, "Securing wireless transmission against reactive jamming: A stackelberg game framework," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–6.

[270] Q. Yan, H. Zeng, T. Jiang, M. Li, W. Lou, and Y. T. Hou, "Jamming resilient communication using mimo interference cancellation," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 7, pp. 1486–1499, July 2016.

[271] M. Han, T. Yu, J. Kim, K. Kwak, S. Lee, S. Han, and D. Hong, "OFDM channel estimation with jammed pilot detector under narrow-band jamming," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 3, pp. 1934–1939, May 2008.

[272] A. O. F. Atya, A. Aqil, S. Singh, I. Broustis, K. Sundaresan, and S. V. Krishnamurthy, "Exploiting subcarrier agility to alleviate active jamming attacks in wireless networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 12, pp. 2488–2501, Dec 2015.

[273] K. Banawan, S. Ulukus, P. Wang, and B. Henz, "Secure rates in multiband broadcast channels with combating jammers," in *MILCOM 2016 - 2016 IEEE Military Communications Conference*, Nov 2016, pp. 390–395.

[274] A. Chorti, "Optimal signalling strategies and power allocation for wireless secret key generation systems in the presence of a jammer," in *IEEE International Conference on Communications (ICC)*, Paris, FR, May 2017.

[275] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[276] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991, translated into Chinesse by Renin University Press, Bejing: China.

[277] J. Hu, K. Yang, L. Hu, and K. Wang, "Reward-aided sensing task execution in mobile crowdsensing enabled by energy harvesting," *IEEE Access*, vol. 6, pp. 37 604–37 614, 2018.

[278] J. Hu, K. Yang, K. Wang, and K. Zhang, "A blockchain-based reward mechanism for mobile crowdsensing," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 1, pp. 178–191, 2020.

[279] J. Hu, M. Reed, M. Al-Naday, and N. Thomos, "Blockchain-aided flow insertion and verification in software defined networks," in *2020 Global Internet of Things Summit (GIoTS)*, 2020, pp. 1–6.

[280] J. F. Nash, "Equilibrium points in n-person games," *Proc. of the National Academy of Sciences*, vol. 36, no. 1, pp. 48–49, 1950.

[281] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[282] A. Chorti, "A study of injection and jamming attacks in wireless secret sharing systems," *Springer, Lecture Notes on Electronic Engineering*, pp. 1–14, Jan. 2018.

[283] Z. Feng, G. Ren, J. Chen, X. Zhang, Y. Luo, M. Wang, and Y. Xu, "Power control in relay-assisted anti-jamming systems: A bayesian three-layer stackelberg game approach," *IEEE Access*, vol. 7, pp. 14 623–14 636, 2019.

[284] Y. Li, L. Xiao, J. Liu, and Y. Tang, "Power control stackelberg game in cooperative anti-jamming communications," in *The 2014 5th International Conference on Game Theory for Networks*, Nov 2014, pp. 1–6.

[285] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *Proc. 15th Annual International Conference on Mobile Computing Networking*. ACM, 2009, pp. 321–332.

[286] S. Eberz, M. Strohmeier, M. Wilhelm and I. Martinovic, "A practical man-in-the-middle attack on signal-based key generation protocols," in *Proc. 17th ESORICS*, 2012, pp. 235–252.

[287] J. Rong and Z. Kai, "Physical layer key agreement under signal injection attacks," in *IEEE Conference on Communication Networking Security (CNS)*, 2015, pp. 254–262.

[288] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: John Wiley and Sons, Inc., 2006.

[289] E. V. Belmega and A. Chorti, "Protecting secret key generation systems against jamming: Energy harvesting and channel hopping approaches," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2611–2626, Nov. 2017.

[290] S. Meier, "Advancing automated security protocol verification," Ph.D. dissertation, ETH Zurich, 2012.

[291] B. Schmidt, "Formal analysis of key exchange protocols and physical protocols," Ph.D. dissertation, ETH Zurich, 2012.

[292] Weisstein, Eric W., "Normal Distribution. From MathWorld–A Wolfram Web Resource."

[293] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Annals of Statistics*, vol. 9, no. 6, pp. 1135–1151, 11 1981.