



UNIVERSITÀ
DEGLI STUDI
FIRENZE

UNIVERSITÀ DEGLI STUDI DI FIRENZE
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE (DINFO)
CORSO DI DOTTORATO IN INGEGNERIA DELL'INFORMAZIONE
CURRICULUM: TELECOMUNICAZIONI

PHYSICAL LAYER SECURITY IN 6-TH GENERATION MOBILE NETWORKS

Candidate

Giacomo Borghini

Supervisors

Prof. Lorenzo Mucchi

Dr. Stefano Caputo

PhD Coordinator

Prof. Stefano Ricci

CICLO XXXVIII, 2022-2025

Università degli Studi di Firenze, Dipartimento di Ingegneria
dell'Informazione (DINFO).

Thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Information Engineering. Copyright © 2026 by
Giacomo Borghini.

*“You were the lightning in that rain, you can still shine through the
darkness.”*

Dedicated to G.M, P.L, D.B, S.B. and all my relatives and friends.

Acknowledgments

I would like to begin by expressing my sincere gratitude to my supervisors, Prof. Lorenzo Mucchi and Dr. Stefano Caputo for the patience and guidance offered throughout my PhD journey. Their support has been instrumental in my growth both as a researcher and as a person. In particular, Dr. Stefano Caputo was the first one to encourage and inspire me to undertake a research career. I wish also to thank the other members of the “Laboratorio per l’Applicazione di Tecnologie dell’Informazione” (LATI), Sara Jayousi and Sara Guarducci.

A special thank goes to Prof. Massimiliano Pierobon for his hospitality during my stay in Lincoln (Nebraska, USA) and his kind and fundamental collaboration to several aspects of my research. Moreover, I wish to thank also Alex Enersen for his collaboration, during and after my period abroad.

Last but not least, I would like to express my heartfelt gratitude to my partner Giulia Mozzini, my family, Danilo Borghini, Paola Latil and Sara Borghini, and all my relatives and friends for their support during my PhD. A special thought goes to A.T., who is fighting a tough battle right now.

I hope that this thesis, and all the work that will stem from it, may serve as a legacy and a constructive contribution to this world. That it may help make it, in its own small way, a better place.

Abstract

The sixth generation (6G) of mobile networks envisions a hyper-connected, intelligent, and immersive digital ecosystem in which also human body will become a node of the network.

This ambitious vision, as outlined in the IMT-2030 framework, imposes exceptionally stringent requirements on 6G networks, particularly in achieving hyper-low latency and enabling the seamless integration and modeling of the human body as a functional node within the network.

To fulfill these demanding requirements, the adoption of novel transmission technologies is essential. Among the most promising candidates for enabling hyper-low latency communication are All-Optical Processing (AOP)-based Optical Communication via Fiber (OCvF) and Terahertz (THz) Communications. In parallel, Wireless Body Area Networks (WBANs) are envisioned to bridge in-body and on-body devices with the broader 6G infrastructure. Within the WBAN domain, emerging paradigms such as Molecular Communications (MCs), Galvanic Coupling (GC) Communications, and Visible Light Communications (VLCs) are being developed to acquire and transmit a diverse array of human-body-related data. These technologies aim to support a holistic representation of the human body, enabling its integration as a dynamic and intelligent node within the network.

However, these technologies introduce novel vulnerabilities and operational constraints. Traditional encryption techniques may prove excessively time-consuming and thus unsuitable for hyper-low latency transmission systems. Moreover, sensors and devices employed within the WBAN context are typically resource-constrained and not equipped to support complex and computationally intensive security techniques. Finally, considering the recent advancements of generative AI-algorithm in terms of complexity and usage, an additional challenge emerge in WBANs: the need to determine whether signals, data, and multimedia content truly originate from a human body.

Physical Layer Security (PLS) emerges as a promising solution to these challenges. By leveraging the inherent characteristics of communication channels, PLS enables fast and lightweight security mechanisms that operate directly at the physical layer. These techniques are particularly well-suited to

the computational and energy constraints of resource-limited devices and the stringent latency requirements of next-generation communication systems.

This thesis applies Physical Layer Security (PLS) techniques for confidentiality and authentication to address the security challenges posed by emerging transmission technologies in 6G networks. A unifying theme across all contributions is the need to secure communication systems that are both latency-sensitive and resource-constrained, particularly in the context of human-centric networking enabled by WBANs. The research spans multiple layers of the 6G architecture, from high-speed backbone technologies to body-level communications, and includes:

- Confidentiality for AOP-based OCvF and THz Communications [1, 2]. These technologies are key enablers of ultra-low latency transmission in 6G, and the proposed PLS techniques ensure data protection without compromising speed.
- Confidentiality of VLC-based Implantable Medical Devices (IMD) in WBANs [3]. Within WBANs, VLC offers a promising channel for intra-body communication. The thesis assesses reliability and security for VLC-based implantable devices.
- Authentication in GC Coupling communications in WBANs [4]. GC is a low-power technique for on-body data exchange. The work introduces PLS-based authentication methods that respect the energy and computational limitations of wearable devices.
- Security threats modeling of diffusion-based MCs in WBANs [5]. MCs represent a novel paradigm for intra-body data transmission. This contribution identifies and models potential vulnerabilities, laying the groundwork for future PLS applications.
- Authentication of AI-generated content in WBANs [6]. As generative AI becomes increasingly capable of mimicking human signals and behavior, this work proposes methods to distinguish genuine human-originated data from synthetic content, ensuring trust in human-body-as-a-node systems.

Contents

Contents	iii
List of Figures	vii
List of Tables	xi
List of Acronyms	xiii
Publications	xv
1 Introduction	1
1.1 6G Requirements	2
1.2 6G Security Challenges	5
1.3 Physical Layer Security as a Strategic Response	6
1.3.1 Advantages of Physical Layer Security Compared to Traditional Cryptography	7
1.3.2 Main Contributions	9
1.4 Organization of the Thesis	12
2 PLS Solutions for HRLLC Transmission Technologies	13
2.1 Confidentiality in OCvF: the Optical Noise-Loop Modulation	14
2.1.1 System Model	16
2.1.2 Reliability Analysis	19
2.1.3 Security Analysis	20
2.1.4 Hardware implementation	29
2.1.5 Results	34
2.1.6 Discussion	38
2.1.7 Conclusion	38

2.2	Confidentiality in THz communications: effects of phase and molecular noise on Noise-Loop modulation	39
2.2.1	Noise Sources in THz Communications	40
2.2.2	System Model	44
2.2.3	Reliability analysis	45
2.2.4	Results	47
2.2.5	Discussion	50
2.2.6	Conclusion	50
3	PLS Solutions for WBANs Transmission Technologies	51
3.1	Security framework for Wireless Body Area Networks: Comparison Between ETSI and IEEE approaches	53
3.1.1	Standards Suitable for WBANs	54
3.1.2	Security in IEEE BAN-related standards	56
3.1.3	Security in ETSI SmartBAN	58
3.1.4	SmartBAN Architecture	58
3.1.5	ETSI SmartBAN Approach to Security	60
3.1.6	Comparison of security approaches: ETSI SmartBAN Vs IEEE 802.15.1, 15.4 and 15.6 standards	63
3.1.7	Conclusion	63
3.2	PLS threats analysis in Molecular Communications	65
3.2.1	Modulation schemes in MCvD	67
3.2.2	Security Threats Analysis	69
3.2.3	Complexity Analysis of the Threats	75
3.2.4	Conclusion	78
3.3	Confidentiality in VLC: Secure Area analysis for VLC-based Implantable Medical Devices (IMD)	80
3.3.1	System Model	82
3.3.2	Results	89
3.3.3	Conclusion	96
3.4	PLA in GC Communications: GC-based Biometric Authentication in WBAN	97
3.4.1	Methodology	99
3.4.2	Experimental Campaign	104
3.4.3	Numerical Results and Discussion	106
3.4.4	Conclusion	110
3.5	Authentication of Outputs of Human Body: Measuring Aesthetic Information in AI-Generated and Human Artistic Images	112

3.5.1	System Model and InfAes Metrics	114
3.5.2	Analysis Methods to Evaluate Informational Aesthetics Metrics	126
3.5.3	Experimental setup and tests	129
3.5.4	Results	134
3.5.5	Comparison with Classifiers in Literature	143
3.5.6	Discussion	143
3.5.7	Conclusion	146
4	Conclusion	147
4.1	Summary of contribution	147
4.2	Directions for future work	149
A		151
A.1	Results of the pre-tests described in Sec. 3.5	151
A.1.1	Impact of generating the image with a size different from the native one	151
A.1.2	Impact of resizing the image	152
A.1.3	Impact of JPEG compression with different Q_F	154
	Bibliography	159

List of Figures

1.1	Representation of IMT 2030 capabilities, extensions and challenges.	2
2.1	ONL modulation scheme: (a) general scheme; (b) receiver scheme for the legitimate terminal Alice; (c) receiver scheme for the eavesdropper Terminal 3 (Eve).	17
2.2	Possible MITM attack scheme in a ONL context.	27
2.3	Real-world fiber optic based implementation scheme of the full ONL modulation system.	29
2.4	Histogram of the decision variable's values (2.55) with a number of loops $L = 5$ (a) and $L = 100$ (b), for 100 000 bits sent. For each approximation method, we report the optimal thresholds in Table 2.1.	35
2.5	Histogram of the values of Alice's [subfigs. (a) and (b)] and Eve's [subfigs. (c) and (d)] decision variable (2.55) with a number of loops $L = 100$ and 100 000 bits sent by Bob. The occurrences of the decision variable's values refer to the bits "+1" [subfigs. (a) and (c)] and "-1" [subfigs. (b) and (d)] sent by Bob.	36
2.6	Molecular noise PSD.	42
2.7	PN characteristics.	44
2.8	Noise Loop modulation scheme. The THz channel noise sources are included.	45
2.9	BER as a function of the number of loops L	46
2.10	The representation of the decision variable z_1 in the particular case where $\alpha = 0.9$, the thermal and molecular noise ($\sigma_n^2 = 10^{-3}$) is comparable to the phase noise. ($L = 12$)	49

3.1	WBAN Reference Architecture.	54
3.2	SmartBAN Architecture.	60
3.3	Molecular Communication system model.	65
3.4	Representation of the diffusion process.	67
3.5	General model of a MCvD system.	70
3.6	Example of jamming attack in a diffusion-based MC system.	71
3.7	Example of eavesdropping attack in a diffusion-based MC system.	72
3.8	Example of man in the middle attack in diffusion-based MC systems.	73
3.9	Example of spoofing attack in a diffusion-based MC system.	74
3.10	Visible Light Communication frequency spectrum.	81
3.11	VLC Implantable Use Case	82
3.12	Human skin absorption for wavelength at different skin layers: 1. Stratum Corneum; 2. Epidermis; 3. Dermis.	83
3.13	VLC System Model	83
3.14	Simulated Model	85
3.15	MIMO System Model	87
3.16	Transfer Function of In-body Channel: in (a) the distribution of light intensity along the skin surface; in (b) the refraction effect. Pattern emission of skin scaled to intensity and average direction: (c) Cartesian plot (d) Polar plot.	91
3.17	Bit error rate (BER) over the vertical plane for the single LED (a) and multi-LED (b)–(c) systems.	93
3.18	Evolution of the Secure Area A_{sec} moving RX position along the (a) vertical axis (b) horizontal axis.	95
3.19	Galvanic Coupling system model.	97
3.20	Architecture of proposed authentication system.	100
3.21	Experimental setup for GC signal acquisition, showing the connection between transmitter and receiver, and the placement of electrodes on the forearm.	105
3.22	Electrode placement scheme illustrating longitudinal distance (d_l) between TX and RX pairs, and transverse distance (d_t) within each pair.	106
3.23	Reliability and Security Performance.	107
3.24	Confusion Matrix.	108
3.25	Reliability and Security Performance.	109

3.26	Confusion Matrix.	111
3.27	The system model. The different InfAes metrics are illustrated at the bottom, depending on the parts of the image they act on.	115
3.28	Effects of generating images at various sizes with AI Stable Diffusion model 2.1 and comparison with the original paintings.	132
3.29	Global Aesthetics metrics distribution and Gaussian fitting curves: AI vs. Human comparison.	136
3.30	Compositional Aesthetics metrics distribution and Gaussian fitting curves: AI vs. Human comparison.	137
3.31	Mutual Information-based Partitioning metrics distribution and gaussian fitting: AI vs. Human comparison.	138
3.32	Classification of Human/AI artworks: confusion matrix obtained with the Fine Gaussian SVM.	141
3.33	Fine Gaussian SVM accuracy across the metric sets defined for the two experiments.	144
3.34	Classification of Human/AI artworks: confusion matrix obtained with the literature image classifier.	145
A.1	Effects of generating images at various sizes on the InfAes metrics.	153
A.2	Effects of resizing 768x768 images to different sizes on the InfAes metrics.	155
A.3	Effects of compressing images to JPEG with different compression quality factors (Q_F) on the InfAes metrics.	157

List of Tables

2.1	Threshold for the 4 approximation methods to compute the optimal decision threshold as in (2.59).	35
2.2	Bit Error Rate for the 4 approximation methods to compute the optimal decision threshold as in (2.59).	36
2.3	Parameter Settings for Fig. 2.6	42
2.4	Simulation Parameters Setting.	48
3.1	Security Features Comparison of WBAN-related Standards .	62
3.2	Complexity level of each PLS attack for each MC modulation scheme.	76
3.3	Skin Parameters used in the MC simulation	90
3.4	Order of magnitude (OOM) of the complexity of each InfoAes metric. The complexity is reported in base-10 logarithmic scale.	126
3.5	Human vs. AI comparison in terms of mean and standard deviation of InfoAes metrics.	135
3.6	Results of label estimation exploiting a Gaussian fitting of the histograms of the metrics.	139
3.7	Results of K-means clustering considering metrics set composed with increasing histogram-based label estimation accuracy metrics.	139
3.8	Results of K-means clustering considering metrics set composed with increasing complexity metrics.	140
3.9	Performance of the classifiers.	140
3.10	Results of Fine Gaussian SVM classifier considering metrics set composed with increasing histogram-based label estimation accuracy metrics	142

3.11 Results of Fine Gaussian SVM classifier considering metrics set composed with increasing complexity metrics	142
3.12 Comparison of the best classifier and K-mean clustering with literature image classifier. In round brackets we reported the metric set used for our K-means clustering and Fine Gaussian SVM classifier.	143

List of Acronyms

AI	Artificial Intelligence	MIMO	Multiple Input Multiple Output
ACL	Access Control List	MoCa	Monte Carlo
AES	Advanced Encryption Standard	NL	Noise-Loop
AOP	All-optical Processing	ONL	Optical Noise-Loop
AWGN	Additive White Gaussian Noise	OTA	Over-the-Air
BER	Bit Error Rate	OWC	Optical Wireless Communication
BnM	Bio-nanomachine	PLS	Physical Layer Security
CIR	Channel Impulse Response	PLA	Physical Layer Authentication
GC	Galvanic Coupling	PKE	Physical Key Extraction
IMD	Implantable Medical devices	PrN	Pseudo-random Noise
IoT	Internet of Things	PSD	Power Spectral Density
LSTM	Long Short-Term Memory	RF	Radio Frequency
MAC	Media Access Control	RV	Random Variable
ML	Machine Learning	SVM	Support Vector Machine
MCvD	Molecular Communication via Diffusion		

Publications

This research activity has led to the following publications in international journals and conferences.¹

International Journals

1. S. Caputo, S. Viciani, S. Gherardini, **G. Borghini**, F. Cataliotti, L. Mucchi. “Noise Level Modulation for Secure Optical Communications”, *IEEE Access*, vol. 12, pp. 148693-148705, 2024. [DOI: 10.1109/ACCESS.2024.3476520]
2. **G. Borghini**, A. Enersen, S. Caputo, M. Pierobon, L. Mucchi. “Measuring Aesthetic Information in AI-Generated Artistic Images: A Tool for Authorship Detection”, submitted to *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Oct 2025.

International Conferences and Workshops

1. **G. Borghini**, S. Caputo, L. Mucchi, A. Rashid, S. Jayousi, M. Hamalainen, T. Paso, M. Hernandez. “Security of wireless body area networks for healthcare applications: Comparison between ETSI and IEEE approaches”, in *Proc. of IEEE 17th International Symposium on Medical Information and Communication Technology (ISMICT)*, Lincoln, Nebraska (USA), May 10-12, 2023.
2. S. Caputo, **G. Borghini**, S. Jayousi, A. Rashid, L. Mucchi. “Visible light communications for healthcare applications: Opportunities and challenges”, in *Proc. of IEEE 17th International Symposium on Medical Information and Communication Technology (ISMICT)*, Lincoln, Nebraska (USA), May 10-12, 2023.

¹The author’s bibliometric indices are the following: H -index = 2, total number of citations = 20 (source: Scopus on October 28, 2025).

3. S. Caputo, **G. Borghini**, S. Jayousi, L. Mucchi. “Visible Light Communications for Implantable Medical Devices”, in *Proc. of EAI International Conference on Body Area Networks*, Varanasi, India, Dec 15-16, 2024.
4. S. Caputo, S. Jayousi, **G. Borghini**, M. Hamalainen, T. Paso, M. Hernandez, H. Tanaka, D. Anzai, T. Ito, L. Mucchi. “Etsi smartban vision on the use of visible light communications in body area networks”, in *Proc. of IEEE 17th International Symposium on Medical Information and Communication Technology (ISMICT)*, London, UK, May 15-17, 2024.
5. S. Caputo, **G. Borghini**, G. Pasolini, N. Longhi, G. Cuzzo, L. Mucchi. “Physical layer security at THz communications: the effects of phase and molecular noise”, in *Proc. of 2024 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd 4.0 & IoT)*, Florence, Italy, May 29-31, 2024.
6. **G. Borghini**, S. Caputo, S. Jayousi, M. Magarini, M. Pierobon, L. Mucchi. “Security Threats in Diffusion-based Molecular Communication Systems”, in *Proc. of IEEE 19th International Symposium on Medical Information and Communication Technology (ISMICT)*, Florence, Italy, May 28-30, 2025.
7. S. Caputo, A. Viziello, A. Coviello, M. Magarini, **G. Borghini**, P. Savazzi, S. Jayousi, L. Mucchi. “Biometric Authentication in Galvanic Coupling Communications: A Physical Layer Approach for Secure Wearable Medical Systems”, accepted at *EAI International Conference on Body Area Networks (BodyNets)*, Okayama, Japan, Dec 2-4, 2025.

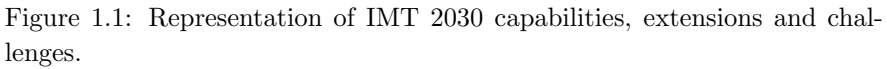
Chapter 1

Introduction

Following the recommendation expressed by the International Mobile Communications (IMT), known as IMT-2030 [7], the sixth generation (6G) of mobile networks are currently under development. IMT-2030 recommendations concern two main aspects: extend the key pillars defined by previous IMT-2020 [8] for 5G networks and the introduction of three novel capabilities envisioned for the 6G networks (see Figure 1.1).

The extension of 5G key pillars aims to provide better coverage, data-rates and lower latencies. In this context, several key components are undergoing significant transformation. In particular, the Ultra Reliable Low Latency Communications (URLLC) in 5G are extended to Hyper Reliable Low Latency Communications (HURLLC), the enhanced Mobile Broadband (eMBB) is extended to immersive communications and massive Machine Type Communications (mMTC) are extended to massive communications [9]. The three novel capabilities envisioned by IMT-2030 for 6G are the realization of an Ubiquitous Connectivity and the integration of Integrated Sensing and Communication (ISAC) and of Integrated AI and Communication (IAAC). Ubiquitous Connectivity is the capability of 6G network of providing seamless connection to the users, regardless of their position or mobility. ISAC refers to the fusion of advanced sensing technologies and reliable communication systems. Finally, IAAC involves Artificial Intelligence (AI) in communication to enable optimized spectrum and resource usage and enhance network security.

By satisfying the recommendations expressed by IMT-2030, the very concept of the network is expected to undergo a significant transformation:



1.1 6G Requirements

- Achieving the extremely low latencies needed to realize HRLLC.
- Modeling and integrating human body as a node of the network.

In order to satisfy these requirements is necessary to introduce new transmission technologies.

To realize HRLLC two promising transmission technologies are proposed: All Optical Processing (AOP)-based Optical Communications via Fiber (OCvF) and THz Communications.

- OCvF is a transmission technology, classified under Optical Communications (OCs), that exploits light waves, that propagates through fiber-optic cables, instead of traditional RF waves to transmit information. OCvF are already a cornerstone of communication networks since they are used to connect the backbone of current mobile networks. In order to reach the hyper-low latencies required for 6G, the introduction of AOP will be crucial. AOP is a technology in which signals are manipulated exclusively in the optical domain. Hence, the processing, storage, and transmission of information are performed using optical devices alone, without any intermediate electrical conversions. Exploiting AOP, optical signals are processed at the same speed as they travel in the optical fiber, allowing higher transmission speeds and lower latencies [10].
- THz communications refer to a transmission paradigm that leverages the 0.1-10 THz frequency band for data exchange. The vast spectral availability in this range enables a substantial reduction in communication latency, supporting ultra-fast transmission. Due to their extremely short wavelengths, THz waves are emitted in highly directional beams. This property contributes to interference mitigation and improved security. A heterogeneous network architecture incorporating high-performance, short-range THz links is therefore capable of meeting the IMT-2030 requirements for ultra-high data rates and extremely low latency.

Modeling human body and integrating it as a node of the network requires to realize an holistic vision of human being, which encompasses its physiological, psychological and social components. These components are defined by the data gathered by the devices and sensors placed both inside and on human body. However, many in-body and on-body sensors, lack the computational capabilities, energy resources, or physical space to support direct connectivity to 6G network. To bridge this gap, an intermediate gateway layer is required. In this thesis, Wireless Body Area Networks (WBANs) are proposed as a promising solution. WBAN architecture is composed by sen-

sors and devices placed inside and on human body which are connected to a central hub, such as a Smartphone, positioned on the body. The central hub collects and aggregates data from the sensors while providing gateway functionalities towards external network, in this case the 6G mobile network [11]. To realize the holistic vision of the human being needed to model it as a node within the 6G network, it is essential to collect a broader and more diverse range of physiological, psychological, and contextual data. Achieving this requires the adoption of novel transmission technologies specifically suited for intra-body and body-centric communications. Three promising approaches within Wireless Body Area Networks (WBANs) are considered here: Visible Light Communications (VLCs), Molecular Communications (MCs) and Galvanic Coupling (GC).

- VLCs are a transmission technology, classified within Optical Wireless Communications (OWCs), which exploits the visible light frequency spectrum. In the context of WBAN, VLC can be exploited as a new transmission technology for sensors and devices placed inside and on human body, as they offer unique advantages for short-range, high-speed and interference-resilient communications. A compelling use case for VLC in this framework are Implantable Medical Devices (IMDs), which refer to health-related sensor and actuators implanted inside human body. An implantable device, *e.g.*, a subcutaneous sensor, can leverage VLC to transmit signals, or relay in-body communication, to an on-body node.
- MCs are a novel transmission technology in which molecules, *i.e.*, a chemical signal, rather than electromagnetic or optical waves, serve as information carriers. Within the context of Wireless Body Area Networks (WBANs), MCs offer a compelling solution for scenarios where the artificial nature of current nano-scale devices, such as those based on nano-materials like graphene, raises concerns regarding biocompatibility and potential health risks. These limitations can be mitigated by replacing artificial devices with biological entities, such as biological cells. MCs are uniquely suited to facilitate communication between biological entities, thereby enabling access to a rich spectrum of physiological data from within the human body.
- GC is a promising technology proposed as an alternative to RF communication in the context of WBANs. GC exploits electrodes placed on

the skin or implanted inside the body to transmit and receive low-power electrical signals which propagates through biological tissues [12]. This communication technology offers intrinsic advantages in terms of reliability, as signals propagating through biological tissues are less susceptible to external interference. Finally, since GC exploits low-power electrical signals, it is a very suitable technology for low-resourced devices, *e.g.*, sensors, placed inside and on human body.

While these emerging transmission technologies are essential to meet the stringent performance and integration requirements of 6G, they also introduce a new layer of complexity in terms of security. The very features that make these technologies suitable for ultra-low latency and human-centric networking also render them vulnerable to novel attacks.

1.2 6G Security Challenges

The challenges raised by the introduction and implementation of the new transmission technologies, described in Sec.1.1, primarily concern two key aspects:

- **Security challenges for HRLLC transmission technologies** To ensure robust security while preserving the ultra-low latency characteristics of AOP-based OCvF and THz Communications it is imperative to adopt fast and lightweight security mechanisms which introduce only minimal latency. Classical encryption techniques could result too time-consuming to meet this requirements. Therefore, new security solution are needed.
- **Security challenges for WBANs transmission technologies** The data gathered and transmitted by devices and sensors placed inside or on the human body, via VLCs, GC Communications, and MCs, to achieve the holistic vision of the human being to model it as a node within the 6G network, are highly sensitive and demand a high level of protection.

The devices involved in GC Communications and VLC-based IMDs are typically resource-constrained and unable of implementing complex, computation-intensive security mechanisms.

Concerning MCs, very low attention has been given to security in current literature, which is mainly focused on modeling propagation

mechanisms and new modulation schemes. Nevertheless, MCs are vulnerable to various attacks that could not only put very sensitive data at risk, but also the life of a person itself. Therefore, it is mandatory to define a security framework also for MCs.

Furthermore, the outputs of human body, *i.e.*, signals and data that flow through WBANs, need to be identified as verifiably originate from a human being. In this context also multimedia products, like artworks or music, can be interpreted as a special kind of output of human body since they derive from the necessity of a human being of communicating its psycho-physical condition, *e.g.*, its thoughts or emotions. The recent increase in complexity and adoption of generative AI models rises the problem that AI models could attempt to emulate sensory, psychological and physiological outputs of human body. As a result, the boundary between human and machine will be blurred to the point of near invisibility. In this context, it is essential to define methods capable for distinguishing whether an output has been emitted by a human being or generated by an AI model.

As a result, the evolution of 6G must be accompanied by a rethinking of security paradigms where lightweight and adaptive solutions can be deployed to safeguard both data and identity. In this context, Physical Layer Security emerges as a promising approach, offering lightweight, channel-aware techniques that operate directly at the physical layer to ensure confidentiality and authentication without relying on traditional encryption.

1.3 Physical Layer Security as a Strategic Response

This thesis investigates how Physical Layer Security (PLS) can be a potential solution to address the security challenges, described in 1.2, raised by the introduction of new transmission technologies in 6G networks.

PLS encompasses a set of techniques that implement security mechanisms directly at the physical layer of the ISO/OSI stack, leveraging the inherent properties of the communication channel to protect data transmission. Security mechanisms are enabled to be tailored to the specific computational and energy constraints of the devices and environments in which they are deployed. Therefore, PLS techniques are suitable to address the security chal-

lenges derived by meeting the stringent requirements of 6G, both in terms of need of fast security techniques for HRLLC transmission technologies and of light-weight security solutions for low-resourced devices and authentication of the outputs of human body in WBANs.

Physical Layer Security techniques can be regrouped in three main application fields: confidentiality, physical layer authentication (PLA) and physical key extraction (PKE). Confidentiality refers to the techniques aimed to protect the privacy of the transmitted information. PLA techniques are aimed to verify the identity of the nodes connected to a network. Finally, PKE refers to techniques applied for the extraction of secret keys used for encryption algorithms. However, in this thesis only confidentiality and authentication techniques have been developed.

1.3.1 Advantages of Physical Layer Security Compared to Traditional Cryptography

The approach provided by PLS offers several distinct advantages over traditional cryptographic methods, particularly in the context of 6G networks and resource-constrained environments such as WBANs [13].

- **Lightweight and Low-Latency Operation.** Traditional cryptographic algorithms, such as AES, RSA, or ECC, require significant computational resources and introduce latency due to key exchange, encryption, and decryption processes. In contrast, PLS techniques operate directly at the physical layer, often relying on simple signal processing operations (*e.g.*, modulation). Therefore, the usage of PLS would allow the adoption of lighter cryptographic mechanisms with reduced computational complexity, since the missing security guarantees can be provided by PLS itself through low-latency physical-layer techniques which are more suitable to latency-sensitive applications, such as HRLLC in 6G [14, 15].
- **No Need for Key Distribution or Management.** Cryptographic systems rely on secure key distribution and management. These processes are complex and vulnerable to attacks, especially in dynamic or decentralized networks. PLS eliminates the need for pre-shared keys by exploiting channel reciprocity or randomness to derive security features. This simplifies protocol design and reduces the attack surface, particularly in mobile or ad hoc scenarios [12, 16].

- **Adaptability to Resource-Constrained Devices.** WBANs devices often have limited processing power, memory, and energy. Implementing complex cryptographic protocols on such devices is impractical. Unlike cryptographic algorithms that may require specialized hardware or high processing power, PLS can be realized through simple signal processing operations.

PLS techniques are well-suited to these environments, as they can be implemented with minimal hardware and computational overhead, enabling the implementation of lighter cryptographic mechanisms while keeping the communication secure [17, 18].

- **Resilience Against Computational Attacks.** Cryptographic security is based on computational hardness assumptions, which may be invalidated by advances in computing power, especially quantum computing. In particular, asymmetric cryptographic algorithms are particularly vulnerable to quantum-based attacks, since a sufficiently powerful quantum computer can exploit Shor’s algorithm to enable the decryption of protected communications [19]. In contrast, symmetric cryptography is not directly broken by quantum computation but is instead weakened: Grover’s algorithm effectively halves the security level of a symmetric key by providing a quadratic speed-up in brute-force search [20]. This reduction, however, can be mitigated by doubling the key length, for example by transitioning from AES-128 to AES-256, which restores an adequate security margin even in the presence of quantum adversaries. PLS relies on information-theoretic principles rather than computational hardness assumptions [21], therefore it provides security guarantees that are intrinsically independent of the adversary’s computational capabilities, whether classical or quantum.
- **Context-Aware and Environment-Sensitive Security.** PLS can adapt to the physical context of the communication environment. For example, channel fingerprinting and proximity-based authentication allow systems to dynamically adjust security levels based on spatial, temporal, or behavioral cues. This context-awareness is difficult to achieve with traditional cryptography, which treats the channel as a transparent medium [22, 23].

To conclude, while cryptography remains essential for many applications, PLS offers a complementary and often superior approach for securing next-

generation networks, especially where latency, low complexity, and adaptability are critical. In the context of 6G, PLS emerges not only as a technical solution but also as a strategic enabler of security.

1.3.2 Main Contributions

The main contributions of this thesis are the PLS solutions developed to address the security challenges outlined in 1.2, which are listed below:

- **PLS solutions for HRLLC transmission technologies:**

- *PLS confidentiality in AOP-based OCvF* - Since it operates at the first level of the ISO/OSI stack, PLS represent a quite natural approach to define security solutions for AOP-based OCvF, which also operates at the physical layer.

Leveraging an AOP approach, the Noise-Loop (NL) modulation, previously applied exclusively to RF systems, has been implemented in an OCvF system, resulting in the development of Optical Noise-Loop (ONL) modulation. ONL is a robust PLS technique that enhances confidentiality in optical networks by injecting random phase noise into the transmitted signal. The legitimate receiver can reliably recover the original information through a feedback loop, while any eavesdropping attempt fails, regardless of the attacker's location or computational capabilities.

- *PLS confidentiality in THz communications* - The implementation of the NL modulation has been evaluated also for THz Communications. Since THz communication channels present additional noise sources, with respect to RF and OC, a modeling of the THz-specific noise and analysis of its effects on the performance, in terms of reliability and security, of the NL modulation has been carried out.

- **PLS solutions for WBANs transmission technologies:**

- *Security threats analysis in Molecular Communication* - Since only the physical layer have been considered and modeled in MCs by current literature, PLS is the best approach to study how to protect information sent through this transmission technology. Focusing on Molecular Communication via Diffusion (MCvD), the

four main security threats known in classical telecommunications, *i.e.*, Jamming, Eavesdropping, Man in the Middle (MITM) and Spoofing, have been mapped in a diffusion-based MCs scenario. Additionally, an analysis of the complexity required by the hostile node to carry on these attacks is conducted.

- *Confidentiality for VLC-based IMD* - Considering both a single LED and a 2×2 Multiple Input Multiple Output (MIMO) scenario, the In-to-Out Body channel between an implanted transmitter to an out-body receiver has been modeled. Reliability and security performance of the system have been analyzed in terms of BER and secure area.
- *PLA in GC Communications* - In GC Communications the propagation channel presents a subject-specific nature, which means that they are linked to the individual's physiological and anatomical features and varies from person to person. PLS techniques can be implemented to exploit the features to enhance the security of the communication between low resourced devices placed on or inside human body
 A biometric-based PLA protocol has been applied to GC context. A Long Short-Term Memory (LSTM) neural network has been implemented to analyze raw GC data and authenticate the nodes of the system. Two typical types of attacks for GC Communications have been considered: Over-The-Air (OTA) and Direct Contact. The performance of the authentication protocol have been evaluated in terms of reliability and security.
- *Authentication of outputs of the body* - The information theoretic approach defined by PLS is implemented to define new methods to distinguish human created or AI generated products. The challenge of distinguishing between human-created and AI-generated outputs of human body has been addressed starting from the field of visual arts. Grounded in the theoretical framework of Informational Aesthetics (InfAes) [24], the creative process has been conceptualized as an information channel linking the artist to the viewer, with the hypothesis of an 'aesthetic gap' that AI systems are unable to fully bridge. Empirical analyses based on clustering and classification tasks to demonstrate that selected InfAes metrics are effective in differentiating human art from AI-generated

counterparts has been conducted. This contribution has been framed within the context of security in WBANs, since the artworks have been considered and modeled as a particular type of signals generated by the human body, and the authentication of their origin is consequently addressed as a PLS problem. Moreover, the integration into the WBAN scenario is motivated by the observation that, in future developments, the techniques devised for artworks may be adapted and extended to other categories of physiological signals.

Having established the strategic advantages of PLS, this chapter concludes by positioning PLS as a lightweight, adaptive, and context-aware alternative to traditional cryptographic approaches, particularly suited to the stringent requirements and novel architectures of 6G networks. The following chapters build upon this foundation by applying PLS techniques to a range of emerging transmission technologies and use cases. Together, the following chapters illustrate the versatility and relevance of PLS as a foundational security paradigm for 6G.

1.4 Organization of the Thesis

The rest of the thesis is structured:

- **Chapter 2** - reports the outcomes of the works conducted concerning the application of PLS solutions to HRLLC transmission Technologies (AOP-based OCvF and THz communications) [1, 2].
- **Chapter 3** - reports an analysis of security on WBANs, including a comparative evaluation of the approaches proposed by IEEE and ETSI standards [11]. It also reports the analysis of security threats in MCs [5] and the outcomes of the works concerning PLA in GC communications and the authentication of outputs of the body [4, 6].
- **Chapter 4** - Concludes the thesis.

Chapter 2

Physical Layer Security Solutions for HRLLC Transmission Technologies

In this chapter PLS confidentiality techniques are applied to the HRLLC transmission technologies envisioned for 6G. In particular the application of the NL modulation, originally developed for RF systems, to AOP-based OCvF and THz communication systems is investigated.

The results presented in this chapter have been published in the following papers:

- S. Caputo, S. Viciani, S. Gherardini, **G. Borghini**, F. Cataliotti, L. Mucchi. “Noise Level Modulation for Secure Optical Communications”, *IEEE Access*, vol. 12, pp. 148693-148705, 2024 [1].
- S. Caputo, **G. Borghini**, G. Pasolini, N. Longhi, G. Cuozzo, L. Mucchi. “Physical layer security at THz communications: the effects of phase and molecular noise”, in *Proc. of 2024 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd 4.0 & IoT)*, Florence, Italy, May 29-31, 2024 [2].

2.1 Confidentiality in OCvF: the Optical Noise-Loop Modulation

Optical Noise-Loop is an AOP-based PLS confidentiality technique applied to OCvF systems. Using the ONL modulation, confidentiality is achieved by leveraging the inherent noise of transmitting and receiving devices to modulate the information. The objective of ONL modulation is to prevent an unintended receiver to correctly demodulate the information from the received signal. As a result, the need to rely on encryption algorithms, which are time-consuming and whose security depends on the computational power of potential attackers, can be eliminated. This means that information sent with ONL are secure also from quantum attacks that could break conventional encryption schemes [19].

To develop the ONL, an AOP-based approach is exploited. Therefore, all the operations needed to implement the ONL modulation are carried on by optical devices operating directly on the optical signal. Therefore this security technique introduces minimal latency and it's suitable for HRLLC applications.

The benefits of this technique can be summarized as follows:

- The proposed technique does not require any prior secret key exchange or key management between the sender and the receiver, as the noise is generated randomly and locally at the transmitter. Therefore, it simplifies the security protocol and reduces the cost and complexity of key distribution.
- The proposed technique does not suffer from physical attacks, such as fiber cutting, tapping, or jamming, as the noise masks the information and makes it indistinguishable from the background noise. Therefore, it enhances the resilience and robustness of the communication system.

The Noise-Loop modulation has already been investigated in current literature, although it has been applied only to RF domain [14]. Anyway, in current literature there are no attempts to adapt this technique to the context of OCvF.

The main challenge to realize ONL consists in modifying the classical Noise-Loop system in order to move from the Additive White Gaussian Noise (AWGN) used in RF, to the phase noise used in fiber optics. Also an insight on the equipment, materials and scheme to realize the ONL in

OCvF is presented, since its implementation as an AOP-based solution is not straightforward.

Related Works. In the current literature on Optical Communications, Physical Layer Security (PLS) is predominantly proposed to address security concerns in Optical Wireless Communications (OWCs), while its application to Optical Core via Fiber (OCvF) remains relatively limited [25]. In [26], the authors examine the trade-off between information rates and confidentiality, referred to as secrecy capacity in optical fiber systems, taking into account various channel dynamics.

The authors in [27] propose a new physical layer technique for secure key generation and distribution in passive optical networks, leveraging randomness and error correcting codes (ECC) to overcome vulnerabilities in standard upstream key handling. By intentionally inducing short bursts of random errors and exploiting ECC theory, the optical line terminal can reliably reconstruct the specific error pattern of the optical network units and derive secret keys that remain inaccessible to potential eavesdroppers.

Although in radio frequency (RF) communications, the noise (intrinsic in the equipment or in the channel or artificially generated) is often used to produce confidentiality, in optical fibers this method is not fully explored. In the following a review of papers in literature that propose a security mechanism for fiber optics communications, using any kind of noise sources is reported.

The work presented in [28] introduces a novel data encryption method in the physical layer using a broadband optical noise-like signal shared between Alice and Bob. This signal is employed to create a secret key for secure information transmission via the one-time-pad technique. The paper evaluates the scheme's features and assesses its compatibility with current fiber-optic communication infrastructure.

In [29], a novel scheme combining chaotic encryption and noise masking key-accompanying transmission is proposed to achieve high security and capacity over seven-core fiber. It uses a dual-polarization IQ modulator to generate a 3D-OFDM signal, encrypted by a chaotic signal and masked by a noise signal, which also carries the key information extracted at the receiver using a correlation algorithm.

The authors in [30] investigate an optical chaotic-based secure fiber-optic communication system using an externally modulated laser to generate a

chaotic optical signal that hides the data signal. The study in [31] focuses on secure communication in modern fiber optic networks, specifically with a multi-mode fiber (MMF) channel and a potential eavesdropper, leveraging artificial noise (AN) and statistical knowledge of the eavesdropper's channel to maximize the average secrecy rate.

An overview of physical-layer attacks in optical networks can be found in [32], and the fundamentals of PLS in optical wireless networks are discussed in [33, 34].

Current security approaches for optical communications are based on mathematical algorithms that assume the computational complexity of breaking the encryption is too high for any attacker to achieve. However, this assumption may not hold in the future, as quantum computers could potentially break these algorithms in polynomial time [20] [?].

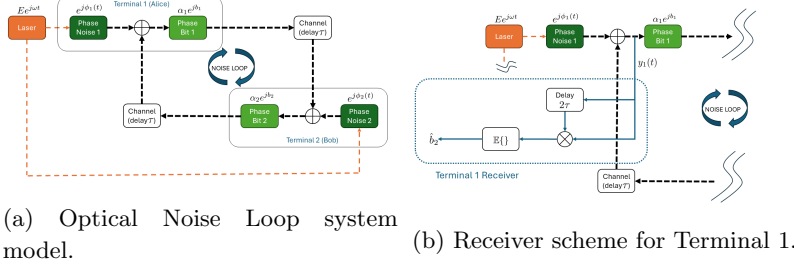
Other approaches rely on the secure distribution and management of keys, which can be challenging and costly in large-scale networks [35]. Moreover, the keys can be compromised by various attacks, such as eavesdropping, interception, or tampering. In addition, these methods introduce overhead and latency in the encryption and decryption processes, which can affect the performance and efficiency of the optical communication systems.

2.1.1 System Model

The ONL system model is depicted in Fig. 2.1, while the legitimate receiver scheme is reported in Fig. 2.1b and the eavesdropper receiver scheme is in Fig. 2.1c. The laser source, *i.e.*, the orange block in Fig. 2.1a, produces the optical signal which undergoes a phase shift (the dark green block). This signal is then added to the incoming signal from the optical channel transmitted by the other terminal). Subsequently the signal is modulated with the information bit of the terminal (the light green box) and then transmitted over the fiber optic to the other terminal. As it can be seen from Fig. 2.1a, the two legitimate terminals Alice and Bob exchange their bits at the same time over the fiber optic. Each bit, together with the received signal, is modulated by a phase noise generator before transmission. This mechanism draws a loop, where the two information bits are merged together and modulated by the two independent phase noises.

It is important to note that both legitimate terminals are transmitting (Tx) and receiving (Rx) at the same time. For the sake of simplicity, we can denote the Terminal 1 as Alice and the Terminal 2 as Bob, but both of them

can be either the Tx or the Rx.



(c) Receiver scheme of the eavesdropper (Eve). Eve is attempting to demodulate the bit transmitted from Alice to Bob. The parameter τ_3 denotes the delay between Alice and Eve.

Figure 2.1: ONL modulation scheme: (a) general scheme; (b) receiver scheme for the legitimate terminal Alice; (c) receiver scheme for the eavesdropper Terminal 3 (Eve).

Referring to the system model in Figure 2.1a, let's assume that the laser amplitude is constant, *i.e.*, $E(t) = E$ and that the laser frequency is constant, *i.e.*, $\omega(t) = \omega$. Moreover, let's assume that the constant phase of the laser source is negligible. Then, we define the signal transmitted by Alice and Bob, initially without closing the loop, as:

$$x_1(t) = \alpha_1 E e^{j\omega t} e^{j\phi_1(t)} e^{jb_1} \quad (2.1)$$

$$x_2(t) = \alpha_2 E e^{j\omega t} e^{j\phi_2(t)} e^{jb_2} \quad (2.2)$$

where:

- E : electric field intensity of the laser source.

- b_1, b_2 : information symbol (which takes the values 0 and π) exchanged by the legitimate nodes.
- α_1, α_2 : channel's attenuation.
- τ : channel's delay.
- ϕ_1, ϕ_2 : noise processes introduced by the phase modulators.

The phase noise from both the laser and the fiber, which includes fluctuations in the optical phase due to the laser's output and the fiber's transmission, is assumed to be included in the phase noise $\phi_1(t)$ and $\phi_2(t)$ of the ONL loop scheme. The parameters α_1 and α_2 attenuate the signal and they will play the role of loop stability controllers that can be implemented by using a splitter or a dissipating mirror.

Let's now close the loop between Alice and Bob. The signal received by Alice can be written as

$$r_1(t) = \alpha_2 e^{jb_2} \left[r_2(t - \tau) + E e^{j\omega(t-\tau)} e^{j\phi_2(t-\tau)} \right] \quad (2.3)$$

where $r_2(t - \tau)$ denotes the signal received by Bob, delayed by the time τ that is proportional to the physical distance from Bob to Alice (and vice versa in our setting). In addition, for Alice, the decision variable is

$$y_1(t) = r_1(t) + E e^{j\omega t} e^{j\phi_1(t)}. \quad (2.4)$$

After some calculations, and after one complete loop ($L = 1$), the expression of $y_1(t)$ becomes

$$\begin{aligned} y_1(t) = & \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} E e^{j\omega(t-2\tau)} e^{j\phi_1(t-2\tau)} \\ & + \alpha_2 e^{jb_2} E e^{j\omega(t-\tau)} e^{j\phi_2(t-\tau)} + E e^{j\omega t} e^{j\phi_1(t)}. \end{aligned} \quad (2.5)$$

Hence, if the following assumptions hold

$$\mathbb{E}\{e^{-j\phi_1(t)} e^{j\phi_2(t)}\} = 0 \quad (2.6)$$

$$\mathbb{E}\{e^{-j\phi_1(t)} e^{j\phi_1(t-\tau)}\} = \delta(\tau), \quad (2.7)$$

then the autocorrelation $R_{y_1 y_1}(2\tau) \equiv \mathbb{E}\{y_1^*(t) y_1(t - 2\tau)\}$ in the single loop becomes

$$R_{y_1 y_1}(2\tau) = \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} E^2. \quad (2.8)$$

Taking the real part

$$\text{Re}\{R_{y_1 y_1}(2\tau)\} \sim \alpha_1 \alpha_2 E^2 \cos(b_1 + b_2); \quad (2.9)$$

thus, since Alice knows the value of b_1 , it can easily estimate the angle b_2 in the range $[0, \pi]$, *i.e.*,

$$\hat{b}_2 = \arccos(\text{sgn}(\text{Re}\{e^{-jb_1} R_{y_1 y_1}(2\tau)\})). \quad (2.10)$$

After L loops, adopting a more comfortable notation, we have

$$\begin{aligned} y_1(t) = & \sum_{\xi=1}^{L-1} (\xi_1 \xi_2)^\ell n_1(t - 2\ell\tau) \\ & + \sum_{\xi=1}^{L-1} (\xi_1 \xi_2)^{\ell-1} \xi_2 n_2(t - (2\ell - 1)\tau) + n_1(t) \end{aligned} \quad (2.11)$$

where $\xi_i \equiv \alpha_i e^{jb_i}$ and $n_i(t) \equiv E e^{j\omega t} e^{j\phi_i(t)}$ with $i = 1, 2$.

2.1.2 Reliability Analysis

In this section a theoretical analysis of the reliability performances of the ONL modulation is provided, *i.e.*, it's proven that Alice and Bob are able to correctly demodulate also after L loops.

Without loss of generality, let's assume that Alice is trying to demodulate the symbol b_2 sent by Bob.

In order to demonstrate the reliability of the system, let's express $y_1(t)$ as an auto-regressive signal:

$$\begin{aligned} y_1(t) = & \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} y_1(t - 2\tau) \\ & + \alpha_2 e^{jb_2} E e^{j\omega(t-\tau)} e^{j\phi_2(t-\tau)} + E e^{j\omega t} e^{j\phi_1(t)}. \end{aligned} \quad (2.12)$$

Thus the autocorrelation of $y_1(t)$ becomes:

$$R_{y_1 y_1}(2\tau) = \mathbb{E}\{y_1^*(t) y_1(t - 2\tau)\} = \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} \sigma_{y_1}^2(t) \quad (2.13)$$

where also $\sigma_{y_1}^2(t) \equiv \mathbb{E}\{y_1^*(t) y_1(t)\}$ obeys an auto-regressive equation of order 1, *i.e.*,

$$\sigma_{y_1}^2(t) = E^2 + \alpha_2^2 E^2 + \alpha_1^2 \alpha_2^2 \sigma_{y_1}^2(t - 2\tau) \quad (2.14)$$

under the assumptions (2.6) and (2.7). Thus, by introducing $E^2 + \alpha_2^2 E^2 \equiv \lambda$ and $\alpha_1^2 \alpha_2^2 \equiv \mu$, one obtains

$$\sigma_{y_1}^2(2L\tau) = \lambda (1 + \mu + \dots + \mu^L) = \lambda \sum_{k=0}^L \mu^k. \quad (2.15)$$

The geometric series $\sum_{k=0}^L \mu^k$ is convergent for $|\mu| < 1$ and equal to

$$\sum_{k=0}^L \mu^k = \frac{1 - \mu^{L+1}}{1 - \mu}.$$

Accordingly, one finally has that

$$\sigma_{y_1}^2(t) = \frac{(E^2 + \alpha_2^2 E^2) (1 - (\alpha_1 \alpha_2)^{2(L+1)})}{1 - \alpha_1^2 \alpha_2^2} \quad (2.16)$$

with $|\alpha_1^2 \alpha_2^2| < 1$ for convergence purposes. It can be observed that, for L (number of loops) large, the value of $\sigma_{y_1}^2(t)$ converges to the stationary value

$$\sigma_{y_1}^2(\infty) = \frac{E^2(1 + \alpha_2^2)}{1 - \alpha_1^2 \alpha_2^2}. \quad (2.17)$$

From (2.13), since Terminal 1 knows b_1 , the angle b_2 can be estimated using (2.10) as in the $L = 1$ case.

This demonstrate that Alice is able to demodulate correctly the information bit sent by Bob, also in the generic case of L sufficiently large.

2.1.3 Security Analysis

In this section a theoretical analysis of the security performances of the ONL modulation is provided. It's proven that the information leakage to the eavesdropper is negligible, *i.e.*, the amount of information that Eve is able to extract from the received signal is zero. Two cases are considered here:

1. *Steady-state analysis* - Eve overhears the channel between Alice and Bob when the loop is already on.
2. *Transitory-state analysis* - Eve overhears the channel from the very first round of the loop.

The analysis of the two cases is carried on under the assumption of a single eavesdropper scenario and that Eve has full knowledge of the ONL system. Finally, a brief analysis of the vulnerability of the proposed system to the Man in the Middle (MITM) attack is provided.

Case 1: Steady-state Analysis

Let's assume, without loss of generality, that Eve is overhearing the channel between Alice and Bob (see Figure 2.1c), after the loop has started ($L \gg 0$).

Since both Alice and Bob decode information by performing the autocorrelation of the received signal, here is assumed that also Eve opts to compute the autocorrelation of the signal taken from channel. The received signal for Eve would be

$$y_3(t) = \alpha_1 e^{jb_1} y_1(t - \tau_3) \quad (2.18)$$

which can also be written as an auto-regressive signal

$$\begin{aligned} y_3(t) &= \alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)} \\ &+ \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{j\phi_2(t-\tau-\tau_3)} E e^{j\omega(t-\tau-\tau_3)} \\ &+ \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} y_3(t - 2\tau) \end{aligned} \quad (2.19)$$

under the hypothesis that Eve's receiver has no noise and that, without loss of generality, Eve is listening at the channel from Alice to Bob. The parameter τ_3 indicates the propagation delay between Alice and Eve, and thus denotes the asynchronism between the time when Alice sends information to Bob and the time when Eve takes $y_3(t)$.

The best that Eve can do is to extract the autocorrelation of the received signal. First of all, Eve has to try to get the autocorrelation peak, which is unknown. In such a case,

$$\begin{aligned} R_{y_3 y_3}(\tau + \tau_3) &= \mathbb{E}\{y_3(t) y_3^*(t - \hat{\tau})\} = 0, \\ \forall \hat{\tau} &\neq 2k\tau \text{ with } k = 1, 2, \dots \end{aligned} \quad (2.20)$$

However, if Eve is able to estimate exactly the propagation delay τ between Alice and Bob, thus by setting $\hat{\tau} = 2\tau$, then the autocorrelation that Eve can extract for L large is

$$\begin{aligned} R_{y_3 y_3}(2\tau) &= \mathbb{E}\{y_3(t) y_3^*(t - 2\tau)\} \\ &= \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} \sigma_{y_3}^2 = \alpha_1^3 \alpha_2 e^{jb_1} e^{jb_2} \sigma_{y_1}^2(\infty). \end{aligned} \quad (2.21)$$

Even in this worst case, the output is proportional to the XOR of the bits, thus it is not possible to come up with the value of each specific bit, without knowing one of them.

In order to try different strategies, Eve could decide to participate to the loop by injecting its own phase noise. Thus, if Eve injects the noise $n_3(t) = E_3 e^{j\omega_3 t} e^{j\phi_3(t)}$, its received signal becomes:

$$\begin{aligned} y_3(t) &= \alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)} \\ &+ \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{j\phi_2(t-\tau-\tau_3)} E e^{j\omega(t-\tau-\tau_3)} \\ &+ \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} y_3(t-2\tau) + \underbrace{E_3 e^{j\omega_3 t} e^{j\phi_3(t)}}_{n_3(t)}. \end{aligned} \quad (2.22)$$

Let's assume that the noises $\phi_1(t)$ and $\phi_2(t)$ are random stochastic processes that cannot be replicated by an adversary. Since the noise process $\phi_3(t)$ is uncorrelated to those injected by Alice $\phi_1(t)$ and Bob $\phi_2(t)$, Eve would not have any additional benefit, thus meaning that (2.21) holds. In addition, if Eve injects noise, it performs basically an active attack and the legitimate nodes could detect the presence of an attacker.

Information Leakage to Eve Let's assume that, after overhearing the channel between Alice and Bob, Eve tries to demodulate the symbol b_1 sent by Alice by exploiting the autocorrelation $R_{y_3 y_3}(2\tau)$, expressed as in (2.21). The symbols b_1 and b_2 are both independent binary random variables which takes values in $\{0, \pi\}$ with equal probability, which are used to map the bits B_1 and B_2 , respectively. B_1 and B_2 are independent Bernoulli (0.5) (Ber(0.5)) random variables. Eve tries to demodulate the bits by extracting the phase Z of $R_{y_3 y_3}(2\tau)$, which is expressed as:

$$Z = (b_1 + b_2) \bmod 2\pi. \quad (2.23)$$

Since the symbols b_1 and b_2 are binary random variables which takes values in $\{0, \pi\}$ with equal probability, so is Z . Eve demodulates the bits from Z by using a threshold comparator which maps $Z \in \{0, \pi\}$ into the decision variable $Y_3 \in \{0, 1\}$. Since, once the bits are demodulated, $(b_1 + b_2) \bmod 2\pi$ correspond to the XOR of B_1 and B_2 , Y_3 can be expressed as:

$$Y_3 = B_1 \oplus B_2. \quad (2.24)$$

Since B_1 and B_2 are Ber(0.5) random variables, also Y_3 is a Ber(0.5) random variable.

The amount of information on the bit B_1 that the eavesdropper can get from the decision variable Y_3 , is given by the mutual information:

$$I(B_1; Y_3) = H(Y_3) - H(Y_3 | B_1), \quad (2.25)$$

where $H(\cdot)$ and $H(\cdot|\cdot)$ denotes the Shannon entropy and conditional entropy, respectively [36]. $I(B_1; Y_3)$ is also known as *information leakage* to the eavesdropper. The Shannon entropy of the decision variable Y_3 is expressed as:

$$H(Y_3) = - \sum_{y \in \{0,1\}} p(Y_3 = y) \log p(Y_3 = y), \quad (2.26)$$

where $p(\cdot)$ denotes a probability and y denotes a possible realization of the decision variable Y_3 . The entropy of Y_3 conditioned by the value of the bit B_1 is expressed as:

$$H(Y_3 | B_1) = \sum_{\beta_1 \in \{0,1\}} p(B_1 = \beta_1) H(Y_3 | B_1 = \beta_1), \quad (2.27)$$

where β_1 denotes a possible realization of the random variable B_1 .

Since Y_3 is a $\text{Ber}(0.5)$ random variable, $p(Y_3 = y) = \frac{1}{2}$ thus $\log p(Y_3 = y) = -1$ for all y . It follows that the entropy expressed in (2.26) becomes:

$$H(Y_3) = \sum_{y \in \{0,1\}} \frac{1}{2} = 1 \quad (2.28)$$

Since B_1 is a $\text{Ber}(0.5)$ random variable, $p(B_1 = \beta_1) = \frac{1}{2}$ for all β_1 . Therefore, (2.27) becomes:

$$H(Y_3 | B_1) = \sum_{\beta_1 \in \{0,1\}} \frac{1}{2} H(Y_3 | B_1 = \beta_1), \quad (2.29)$$

The entropy of Y_3 conditioned by $B_1 = \beta_1$ can be expressed as:

$$H(Y_3 | B_1 = \beta_1) = - \sum_{y \in \{0,1\}} \underbrace{p(Y_3 = y | B_1 = \beta_1)}_{p(Y_3=y)=\frac{1}{2}} \log \underbrace{p(Y_3 = y | B_1 = \beta_1)}_{p(Y_3=y)=\frac{1}{2}} \quad (2.30)$$

$$= \sum_{y \in \{0,1\}} \frac{1}{2} = 1. \quad (2.31)$$

Where 2.30 is obtained from the fact that if Y_3 is conditioned to the observation $B_1 = \beta_1$ then:

$$p(Y_3 = y \mid B_1 = \beta_1) = p(B_1 \oplus B_2 = y \mid B_1 = \beta_1) \quad (2.32)$$

$$= p(\beta_1 \oplus B_2 = y) = p(B_2 = y) = \frac{1}{2} = p(Y_3 = y), \quad (2.33)$$

Therefore, conditioning the decision variable Y_3 to the observation of the bit B_1 does not change the probability distribution of Y_3 . By substituting (2.31) in (2.29) it follows that:

$$H(Y_3 \mid B_1) = \sum_{\beta_1 \in \{0,1\}} \frac{1}{2} \underbrace{H(Y_3 \mid B_1 = \beta_1)}_1 = 1, \quad (2.34)$$

Comparing (2.34) and (2.28), it follows that:

$$H(Y_3 \mid B_1) = H(Y_3) = 1. \quad (2.35)$$

Substituting this result in (2.25) yields:

$$I(B_1; Y_3) = H(Y_3) - H(Y_3 \mid B_1) = 0. \quad (2.36)$$

This demonstrate that the mutual information $I(B_1; Y_3)$ is zero for the eavesdropper, *i.e.*, Eve cannot recover any information about the legitimate bit from the decision variable Y_3 , no matter the computational power owned by the eavesdropper. This result also highlights that the proposed system provides a *full secrecy rate*, *i.e.*, each transmitted bit is a secure bit.

Case 2: Transitory-state Analysis

Let's assume, without loss of generality, that Eve is overhearing the channel between Alice and Bob, starting from the beginning, *i.e.*, when the loop is just started ($L = 0$). If the eavesdropper is overhearing the ONL channel from the very beginning of the transmission, it is able to get also the first sending from Alice to Bob, where only b_1 is modulated.

In this case, considering the worst case scenario where Eve equipment adds a negligible phase noise, the signal observed by Eve is

$$y_3^{[L=0]}(t) = \alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)} = \alpha_1 e^{j(\phi_1(t-\tau_3)+b_1)} E e^{j\omega(t-\tau_3)} \quad (2.37)$$

where τ_3 is the propagation delay of the eavesdropping channel from Alice to Eve. As it can be seen, (2.37) depends only on one bit b_1 , thus Eve could analyze this signal to extract the information sent by Alice to Bob. In order to see if Eve can extract information from that signal, the mutual information is derived.

In this case, Eve cannot use the autocorrelation as decision variable because it would be not proportional to the bit. In fact, under the assumptions (2.6) and (2.7), the autocorrelation is expressed as:

$$\mathbb{E}\{y_3^{[L=0]}(t)y_3^{*[L=0]}(t-\tau)\}_{|\tau=0} = \alpha_1^2 E^2 \quad (2.38)$$

The best Eve can do is to extract the phase of the signal $y_3^{[L=0]}(t)$, expressed as in (2.37), and use it as decision variable $Y_3^{[L=0]}$. $Y_3^{[L=0]}$ is expressed as:

$$Y_3^{[L=0]} = (\phi_1 + b_1) \bmod 2\pi \quad (2.39)$$

where ϕ_1 is an independent random variable uniformly distributed in $[-\pi, \pi)$ and b_1 is a binary random variable which takes values $\{0, \pi\}$, the time index is omitted for simplicity. The mutual information can be expressed in terms of Kullback-Leibler divergence:

$$I(b_1; Y_3^{[L=0]}) = \mathbb{E}_{(b_1, Y_3^{[L=0]})} \left[\log \frac{p(Y_3^{[L=0]} | b_1)}{p(Y_3^{[L=0]})} \right] \quad (2.40)$$

For any measurable set $A \subseteq [-\pi, \pi)$, the conditional probability $p(Y_3 \in A | b_1 = b)$, where b is a possible realization of the random variable b_1 , can be expressed as:

$$\begin{aligned} p(Y_3^{[L=0]} \in A | b_1 = b) &= p((\phi_1 + b_1) \bmod 2\pi \in A | b_1 = b) \\ &= p((\phi_1 + b) \bmod 2\pi \in A) \end{aligned} \quad (2.41)$$

$$= p(\phi_1 \in (A - b) \bmod 2\pi) \quad (2.42)$$

where (2.41) follows from the fact that both ϕ_1 and b_1 are independent random variables and (2.42) is a change of variables. Therefore:

$$p(Y_3^{[L=0]} \in A | b_1 = 0) = p(Y_3^{[L=0]} \in A | b_1 = \pi) = p(Y_3^{[L=0]} \in A) \quad (2.43)$$

As a consequence of (2.43), eq. (2.40) becomes:

$$I(b_1; Y_3^{[L=0]}) = \mathbb{E} \left[\log \frac{p(Y_3^{[L=0]} \in A | b_1)}{p(Y_3^{[L=0]} \in A)} \right] = \quad (2.44)$$

$$= \mathbb{E} \left[\log \frac{\overbrace{p(Y_3^{[L=0]} \in A | b_1 = 0)}^{p(Y_3^{[L=0]} \in A)} \overbrace{p(b_1 = 0)}^{\frac{1}{2}} + \overbrace{p(Y_3^{[L=0]} \in A | b_1 = \pi)}^{p(Y_3^{[L=0]} \in A)} \overbrace{p(b_1 = \pi)}^{\frac{1}{2}}}{p(Y_3^{[L=0]} \in A)} \right]$$

$$= \mathbb{E}[\log 1] = 0 \quad (2.45)$$

Therefore, Eve is unable to retrieve any information about the legitimate bit from the signal, including during the transient regime.

Additional possibilities for Eve to try to detect the information bits could be to correlate the start signal (2.37) with the steady-state signal (*e.g.*, after one loop) (2.19).

Let's assume without loss of generality that Eve tries to correlate the very first signal coming from Alice to Bob ($L = 0$) with the signal in the same channel once a complete loop is ended ($L = 1$).

$$\begin{aligned} & \mathbb{E} \left\{ y_3^{[L=0]}(t - \tau_3) y_3^{*[L=1]}(t - 2\tau - \tau_3) \right\} = \mathbb{E} \left\{ [\alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)}] \right. \\ & \cdot [\alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)} + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{j\phi_2(t-\tau-\tau_3)} E e^{j\omega(t-\tau-\tau_3)} + \\ & + \alpha_1^2 \alpha_2 e^{jb_2} e^{j\phi_1(t-2\tau-\tau_3)} E e^{j\omega(t-2\tau-\tau_3)}]^* \left. \right\} \\ & = \underbrace{\mathbb{E} \left\{ [\alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)}] \cdot [\alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)}]^* \right\}}_{=\alpha_1^2 E^2} \\ & + \underbrace{\mathbb{E} \left\{ [\alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)}] \cdot [\alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{j\phi_2(t-\tau-\tau_3)} E e^{j\omega(t-\tau-\tau_3)}]^* \right\}}_{=0 \text{ since (2.6)}} \\ & + \underbrace{\mathbb{E} \left\{ [\alpha_1 e^{jb_1} e^{j\phi_1(t-\tau_3)} E e^{j\omega(t-\tau_3)}] \cdot [\alpha_1^2 \alpha_2 e^{jb_2} e^{j\phi_1(t-2\tau-\tau_3)} E e^{j\omega(t-2\tau-\tau_3)}]^* \right\}}_{=0 \text{ since (2.7)}} = \alpha_1^2 E^2 \end{aligned} \quad (2.46)$$

Since $\alpha_1^2 E^2$ does not depend on the bits, Eve cannot recover any information.

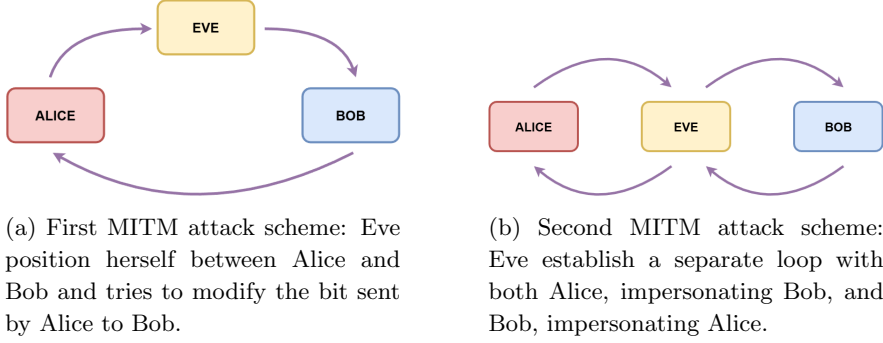


Figure 2.2: Possible MITM attack scheme in a ONL context.

Possible Vulnerabilities to MITM

While this work primarily focuses on analyzing the capability of NL modulation to enhance privacy in AOP-based OCvF networks by countering eavesdropping, a brief assessment of the vulnerability to MITM attacks is also provided.

Conducting a MITM attack requires the adversary not only to intercept but also to modify and relay the signals exchanged between legitimate nodes, which remain unaware that their communication has been compromised. Consequently, whereas eavesdropping constitutes a passive attack, a MITM attack is inherently active [37].

There are two possible attack strategies (see Figure 2.2) that Eve can employ to carry out a MITM attack in the context of ONL modulation.

The first possible MITM attack strategy is showed in Figure 2.2a. Let's assume that Alice and Bob have already established a loop and are communicating exploiting the ONL modulation. Without loss of generality, Eve is considered to be positioned between Alice and Bob, attempting to modify the bit B_1 transmitted from Alice to Bob. As described in the previous sections, by demodulating the phase of (2.21) Eve can obtain, in the worst case, the XOR of bits B_1 and B_2 , expressed as in (2.24). Since Eve doesn't know the value of B_2 that she would need to decode the XOR and then modify the bit B_1 , the best she can do is to alter the value of the XOR received by Bob by adding her own bit B_3 , where B_3 is an independent $\text{Ber}(0.5)$ random variable. In practice, Eve multiplies its symbol b_3 to the received signal $y_3(t)$, where b_3 is an independent binary random variable which takes

values in $\{0, \pi\}$ with equal probability. Therefore, assuming the worst case that Eve's equipment adds negligible phase noise and delay, Bob will receive a signal expressed as:

$$y_2(t) = \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{jb_3} e^{j\omega(t-2\tau)} e^{j\phi_1(t-2\tau)} + \quad (2.47)$$

$$+ \alpha_2 e^{jb_1} e^{jb_3} E e^{j\omega(t-\tau)} e^{j\phi_1(t-\tau)} + E e^{j\omega t} e^{j\phi_2(t)}. \quad (2.48)$$

In this case, the autocorrelation $R_{y_2 y_2}(2\tau)$ computed by Bob can be expressed as:

$$R_{y_2 y_2}(2\tau) = \mathbb{E}\{y_2(t) y_2^*(t - 2\tau)\} = \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{jb_3} \sigma_2^2(\infty) \quad (2.49)$$

In order to demodulate the signal, as in (2.10), Bob compute the product between its symbol b_2 and $R_{y_2 y_2}(2\tau)$:

$$e^{-jb_2} R_{y_2 y_2}(2\tau) = \alpha_1 \alpha_2 e^{jb_1} e^{jb_3} \sigma_2^2(\infty). \quad (2.50)$$

This means that, after demodulating the phase of $R_{y_2 y_2}(2\tau)$, Bob does not obtains the bit B_1 sent by Alice but $B_1 \oplus B_3$. Therefore, with this first possible MITM attack scheme, the best Eve can achieve is a Denial of Service attack, preventing Bob from correctly receiving the bits sent by Alice. In fact, $B_1 \oplus B_3$ have the same properties of (2.24) and it has already been proofed that no information on B_1 can be recovered by (2.24).

In the second MITM attack strategy (see Figure 2.2b), it is assumed that Eve is able to authenticate herself to Alice by impersonating Bob, and to Bob by impersonating Alice, thereby establishing a separate loop with each of them. Given that, in this case, Eve is treated as a legitimate node, she can effectively perform a MITM attack: she receives from Alice the bit B_1 intended for Bob and forwards a modified bit B_3 to him, without Bob having any knowledge of the alteration.

In conclusion, it is worth emphasizing that ONL is a physical-layer confidentiality technique and does not provide authentication by itself. Therefore, resistance to MITM attacks mainly relies on the presence of authentication mechanisms at higher layers; in their absence, as discussed earlier in this section an active adversary may cause denial-of-service (first MITM attack scheme) or even achieve to modify the information exchanged by Alice and Bob (second MITM attack scheme).

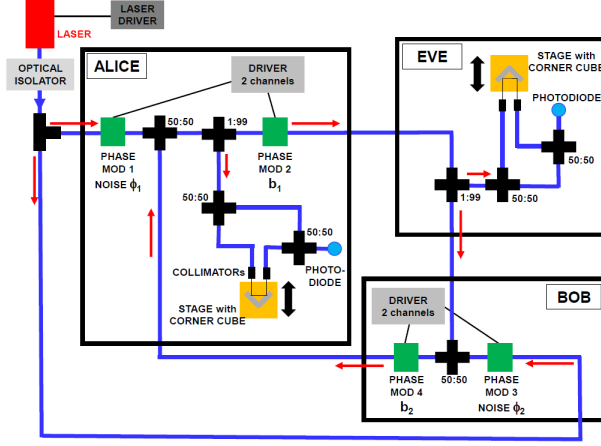


Figure 2.3: Real-world fiber optic based implementation scheme of the full ONL modulation system.

2.1.4 Hardware implementation

This section presents how the ONL modulation described in the previous section can be realized through as an All Optical Processing system, which is based on polarization-maintaining fiber components at telecom wavelength (1550 nm). The hardware implementation scheme is shown in Figure 2.3.

The light source used to carry the information is a single frequency diode laser, emitting at 1550 nm, thermally stabilized and controlled by a low-noise current driver, to provide a laser field with constant amplitude and frequency. The optical source is equipped with an isolator to mitigate feedback noise resulting from back-reflected light re-entering the device, which can disrupt its stable operation.

The optical signal is injected, by means of a 1×2 fiber coupler, into two terminals (Alice and Bob) that can be used both as transmitters and receivers. Before starting the communication loop, a phase noise ϕ is introduced in both terminals. The phase noise is inserted into the fiber network by using 2 fiber-coupled phase modulators (MOD1 for Alice and MOD3 for Bob) driven by a random signal. The loop is realized by two 2×2 fiber couplers, with 50:50 coupling ratio, which combine the signal coming from Alice with the signal coming from Bob. The contribution of phase noise due to both the laser and the fiber is included in the phase noise intentionally

inserted into the loop.

Alice and Bob encode two different information bits (b_1 and b_2 respectively) in the optical signal by changing its phase through two additional fiber-coupled phase modulators (MOD2 for Alice and MOD4 for Bob). Specifically, a binary 0 is represented by a phase shift of π degrees, while a binary 1 is represented by no phase shift.

In Fig. 2.3, the detection of the received message is shown only for the Alice terminal. A 2×2 fiber coupler, with 99:1 coupling ratio, is inserted inside the loop to collect a small portion of the signal and to send it to a phase-sensitive detection scheme that allows the measurement of the optical autocorrelation function. The auto-correlator setup is a Michelson interferometer, realized with two 2×2 fiber couplers with 50:50 coupling ratio, and a moving stage equipped with a corner cube mirror to introduce a variable path difference between the two interferometric arms. The received message b_2 is extrapolated by the measured autocorrelation when the time difference between the two interferometric paths is equal at 2τ .

The eavesdropping is simulated by the terminal Eve, where a phase-sensitive detection scheme (equal to the auto-correlator used by Alice) is employed to eavesdrop the signal transmitted by Alice and Bob.

Hardware implementation: Reliability analysis

It is worth observing that there is a small difference between the theory exposed in Secs. 2.1.2-2.1.2 and the hardware implementation of the ONL scheme in Fig. 2.3. This difference is due to the fact that, in the real implementation, the optical receiver can only extract the square module of the incoming signal. In the implementation scheme, the photo-diode at the legitimate node, *e.g.*, Alice, receives as input the signal

$$|y_1(t) + y_1(t - 2\tau)|^2,$$

which is different from receiving $y_1(t)$ only, as in Sec. 2.1.2. Thus, the successive expectation yields $\mathbb{E}\{|y_1(t) + y_1(t - 2\tau)|^2\}$ and not $\mathbb{E}\{y_1(t)y_1(t - 2\tau)\}$ as in (2.13). However, this difference can be easily overcome by noting that

$$\begin{aligned} y_1(t) + y_1(t - 2\tau) &= (1 + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2}) y_1(t - 2\tau) \\ &\quad + \alpha_2 e^{jb_2} E e^{j\omega(t-\tau)} e^{j\phi_2(t-\tau)} + E e^{j\omega t} e^{j\phi_1(t)}. \end{aligned} \quad (2.51)$$

Thus, reminding Eqs. (2.6) and (2.7), we have

$$\begin{aligned} \mathbb{E}\{|y_1(t) + y_1(t - 2\tau)|^2\} = & 2 \underbrace{\alpha_1 \alpha_2 \operatorname{Re}\{e^{jb_1} e^{jb_2}\}}_{\text{real part of } R_{y_1 y_1}(2\tau) \text{ as in (2.13)}} \sigma_{y_1}^2(t) + \\ & + \underbrace{(1 + \alpha_1^2 \alpha_2^2) \sigma_{y_1}^2(t) + E^2(1 + \alpha_2^2)}_{K > 0} \end{aligned} \quad (2.52)$$

where K is an always positive bias. In order to recover the decision variable, whose sign is related to the bit transmitted by the other node, it is enough to remove both the constant bias from (2.52) and the contribution of the own symbol b_1 , i.e.,

$$\hat{b}_2 = \arccos\left(\frac{\mathbb{E}\{|y_1(t) + y_1(t - 2\tau)|^2\} - K}{2\alpha_1 \alpha_2 \sigma_{y_1}^2(t)}\right) - b_1. \quad (2.53)$$

Therefore, the theory described in Sec. 2.1.2 is still valid.

The bias K can be either calculated as $(1 + \alpha_1^2 \alpha_2^2) \sigma_{y_1}^2(t) + E^2(1 + \alpha_2^2)$ or estimated at the receiver. It is important to note that K is not constant but it becomes constant after some loops, such that $\sigma_{y_1}^2(t)$ is well approximated by $\sigma_{y_1}^2(\infty)$ as in (2.16) and (2.17).

This means that the first estimated bit has to be neglected since it could be not reliable.

Actually, the expectation in (2.52) cannot be realized in a real-world scenario. In fact, the decision variable can be written as in (2.54).

Taking the expectation $\mathbb{E}\{\}$ of (2.54) would make the undesired terms to vanish, thanks to assumptions (2.6) and (2.7). Unfortunately, in a real scenario, the expectation has to be replaced by a time average. In a time average, in particular if it is calculated over a short amount of time (number of loops), the undesired terms cannot be neglected completely.

In addition, let's note that the desired term is summed to an always positive constant term $E^2(1 + \alpha_2^2)$, which has no relevance, and a term $(1 + \alpha_1^2 \alpha_2^2)|y_1(t - 2\tau)|^2$ that is an always positive RV with a Chi-Square distribution (since it comes from a product of two Gaussian RVs). This means that the probability density function (PDF) of the observed received signal is not symmetric. Thus, the optimal threshold to decide on the sign of the transmitted symbol is not in the middle of the PDFs, as we can see in Fig. 2.4a.

$$\begin{aligned}
z &= |y_1(t) + y_1(t - 2\tau)|^2 = \left[y_1(t) + y_1(t - 2\tau) \right] \cdot \left[y_1(t) + y_1(t - 2\tau) \right]^* \\
&= \left[(1 + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2}) y_1(t - 2\tau) + \alpha_2 e^{jb_2} E e^{j\omega(t-\tau)} e^{j\phi_2(t-\tau)} + E e^{j\omega t} e^{j\phi_1(t)} \right] \\
&\quad \cdot \left[(1 + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2}) y_1(t - 2\tau) + \alpha_2 e^{jb_2} E e^{j\omega(t-\tau)} e^{j\phi_2(t-\tau)} + E e^{j\omega t} e^{j\phi_1(t)} \right]^* \\
&= (1 + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2})^2 |y_1(t - 2\tau)|^2 + \alpha_2^2 E^2 + E^2 \\
&\quad + \underbrace{E \alpha_2 e^{jb_2} (1 + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2}) \cdot \left[y_1(t - 2\tau) e^{j\phi_2(t-\tau)} + y_1^*(t - 2\tau) e^{-j\phi_2(t-\tau)} \right]}_{\text{undesired term 1}} \\
&\quad + \underbrace{E (1 + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2}) \cdot \left[y_1(t - 2\tau) e^{j\phi_1(t)} + y_1^*(t - 2\tau) e^{-j\phi_1(t)} \right]}_{\text{undesired term 2}} \\
&\quad + \underbrace{2E^2 \alpha_2 \cos(b_2 + \phi_2(t - \tau) - \phi_1(t))}_{\text{undesired term 3}} \\
&= \underbrace{2\alpha_1 \alpha_2 e^{jb_1} e^{jb_2} |y_1(t - 2\tau)|^2}_{\text{desired term}} + \underbrace{(1 + \alpha_1^2 \alpha_2^2) |y_1(t - 2\tau)|^2}_{\text{always positive RV}} + \underbrace{E^2(1 + \alpha_2^2)}_{\text{constant term}} + \\
&\quad + \text{undesired terms 1, 2 and 3.}
\end{aligned} \tag{2.54}$$

Let's figure out how to find the optimal threshold in case of non symmetric PDFs of the two symbols. Let's note that our decision variable, after the time average, assumes the form

$$\bar{z} = \frac{1}{L} \sum_{l=1}^L z_l = (a \cdot b + c) \sigma_{y_1}^2 + d + w \tag{2.55}$$

where z_l is the decision variable at the end of the l th loop, L is the number of loops, $a = 2\alpha_1 \alpha_2$ is a constant, $b = e^{jb_1} e^{jb_2}$ is a binary random variable which takes values in $\{-1, 1\}$ with equal probability, $c = 1 + \alpha_1^2 \alpha_2^2$ is a constant, $d = E^2(1 + \alpha_2^2)$ is a constant, and w is a Gaussian RV that takes into account the non vanished contribution of the undesired terms in (2.54) after the time average. Since our goal is to find the sign of b_2 knowing the sign of b_1 , we can draw two hypothesis with equal probability: $\mathcal{H}_0 : b = 1$ and $\mathcal{H}_1 : b = -1$. In this case, the decision rule is to choose the hypothesis that has the highest posterior probability, i.e., decide for $b = +1$ if $P(\bar{z}|b = +1) > P(\bar{z}|b = -1)$, and decide $b = -1$ otherwise. The optimal threshold can be defined as the value of $\bar{z}$ that makes the posterior probabilities of both hypotheses equal, i.e., $P(\bar{z}|b = +1)P(b = +1) = P(\bar{z}|b = -1)P(b = -1)$, which indeed is equivalent to $P(\bar{z}|b = +1) = P(\bar{z}|b = -1)$ since the prior probabilities are

equal.

Let's focus on the case $b = -1$. In this case, we decide correctly if $\bar{z} - d \leq \epsilon_1$, where $\epsilon_1 = (a - c)\sigma_{y_1}^2 + w$. The RV ϵ_1 is the sum of two independent RVs, a Chi Square $(a - c)\sigma_{y_1}^2$ and a Gaussian w . Assuming that the threshold is z_0 , we decide correctly if $\bar{z} - d \leq z_0$, whose probability is given by

$$P(\bar{z} - d \leq z_0 | b = -1) = P(\epsilon_1 \leq z_0 | b = -1). \quad (2.56)$$

Analogously, in the case $b = +1$ we have a correct decision if $\bar{z} - d > z_0$, whose probability is given by

$$P(\bar{z} - d > z_0 | b = 1) = P(\epsilon_2 > z_0 | b = 1) \quad (2.57)$$

where $\epsilon_2 = (a + c)\sigma_{y_1}^2 + w$ is again the sum of two independent RVs, a Chi Square $(a + c)\sigma_{y_1}^2$ and a Gaussian w . The optimal threshold z_0 can be derived by imposing that

$$P(\epsilon_1 \leq z_0 | b = -1) = P(\epsilon_2 > z_0 | b = 1) \quad (2.58)$$

Taking in mind that the two RVs ϵ_1 and ϵ_2 both have a non symmetric PDF, the threshold z_0 is not in the middle point of the PDFs. To compute the threshold, it's necessary to impose that

$$\int_{-\infty}^{z_0} f_{\epsilon_1}(\epsilon_1) d\epsilon_1 = \int_{z_0}^{\infty} f_{\epsilon_2}(\epsilon_2) d\epsilon_2 \quad (2.59)$$

where $f_{\epsilon_1}(\epsilon_1)$ and $f_{\epsilon_2}(\epsilon_2)$ denote the PDF of ϵ_1 and ϵ_2 , respectively.

Eq. (2.59) does not have a closed-form solution, but it can be evaluated numerically. To simplify the evaluation, four possible approximations about the distribution of the RVs ϵ_1 and ϵ_2 are proposed, based on which type of assumptions are made on the RVs $\sigma_{y_1}^2$ and w :

1. The decision is computed as in (2.53), *i.e.*, we consider the RV $\sigma_{y_1}^2$ as constant and the RV w (the undesired terms) as negligible. This approach is analogous to the approach defined for radio communication systems [14].
2. We consider the RV w negligible, and thus ϵ_1 and ϵ_2 are Chi Square distributed RVs; the optimal threshold is computed as in (2.59).
3. We consider $\sigma_{y_1}^2$ approximated as a Gaussian RV (for the central limit theorem); thus, ϵ_1 and ϵ_2 are both Gaussian distributed RVs; the optimal threshold is computed as in (2.59).

4. We consider that ϵ_1 is a Gaussian RV (if w dominates the joint PDF $\epsilon_1 = (a - c)\sigma_{y_1}^2 + w$ when $b = -1$), while ϵ_2 is a Chi Square RV (in the case $\sigma_{y_1}^2$ dominates the corresponding joint PDF when $b = +1$); the optimal threshold is computed as in (2.59).

The accuracy of the proposed assumptions are compared with the optimal threshold in Monte Carlo simulations as reported in Sec. 2.1.5.

Hardware implementation: Security analysis

Analogously, in the experimental scheme (Fig. 2.3), let's consider that Eve detects, in the very worst case, the quantity $|y_3(t) + y_3(t - 2\tau)|^2$ and not $y_3(t)y_3(t - 2\tau)$ as in (2.21). Again, this is not a problem, in fact

$$\begin{aligned} \mathbb{E}\{|y_3(t) + y_3(t - 2\tau)|^2\} &= 2\alpha_1\alpha_2 \operatorname{Re}\{e^{jb_1}e^{jb_2}\} \sigma_{y_3}^2(t) \\ &\quad + (1 + \alpha_1^2\alpha_2^2)\sigma_{y_3}^2(t) + \alpha_1^2E^2(1 + \alpha_2^2). \end{aligned} \quad (2.60)$$

After some loops, $\sigma_{y_3}^2(t)$ converges to

$$\sigma_{y_3}^2(\infty) = \frac{\alpha_1^2E^2(1 + \alpha_2^2)}{1 - \alpha_1^2\alpha_2^2} = \alpha_1^2\sigma_{y_1}^2(\infty).$$

Therefore, even if Eve could remove the bias from (2.60), it will still obtain something that is related to the XOR of the two information bits. In conclusion, the theory described in Sec. 2.1.3 is still valid.

2.1.5 Results

In this section the numerical results of the proposed AOP-based ONL system, are presented. In particular, the performances of the system in terms of both reliability and security are showed. Let's start with the reliability results for the decision variables given, respectively, by the theoretical analysis (2.13) and the implementation scheme (2.55). The numerical results are taken from Monte Carlo simulations, implemented in MATLAB, with 100 000 bits sent and a variable number L of loops. We recall that in the ONL scheme, to produce the decision variable, the loop is run L times for each bit.

Figures 2.4a and 2.4b show the occurrences of the values of the decision variable $\bar{z}$ in (2.55) over 100 000 bits. The optimal thresholds are reported in the figures, in order to show that as the number of loops L increases, the two

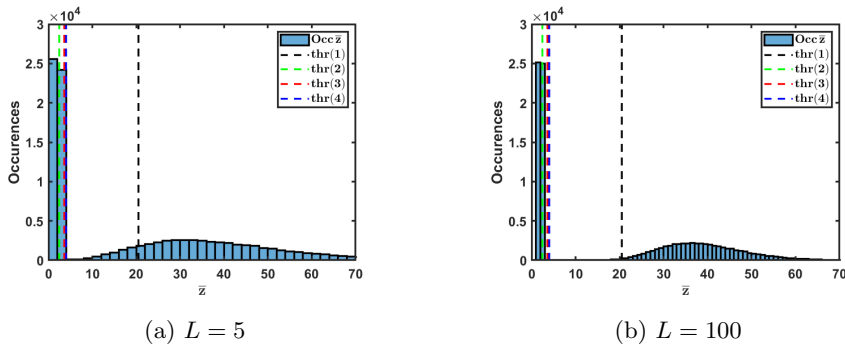


Figure 2.4: Histogram of the decision variable's values (2.55) with a number of loops $L = 5$ (a) and $L = 100$ (b), for 100 000 bits sent. For each approximation method, we report the optimal thresholds in Table 2.1.

Table 2.1: Threshold for the 4 approximation methods to compute the optimal decision threshold as in (2.59).

Methods	Number of Loops [L]						
	5	10	20	50	100	250	400
(1)	20.51	20.51	20.51	20.51	20.51	20.51	20.51
(2)	2.45	2.45	2.5	2.5	2.5	2.5	2.5
(3)	3.6	3.3	3.15	3.05	3.05	3	3
(4)	4	4.05	4.05	4.05	4.05	4.05	4.05

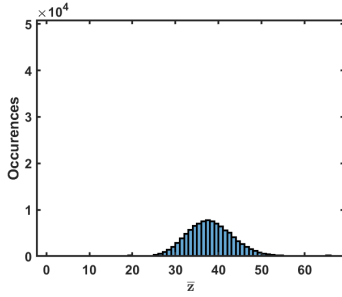
branches of the PDF of $\bar{z}$ shrink and any threshold between 2.5 and 20.51 becomes optimal (Table 2.1), since for a large L the two branches of the PDF are very well separated (see Fig. 2.4b in comparison with Fig. 2.4a).

Let us now discuss briefly the thresholds shown in Figures 2.4a–2.4b:

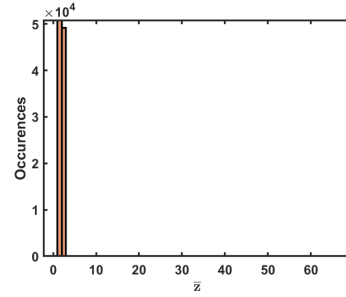
- Threshold (1) is calculated as the mean point of the two mean points of the distributions of $b = -1$ and $b = +1$; this threshold is located much more towards the mean point of the distribution of $b = +1$ (see Fig. 2.4a);
- Threshold (2) is located more towards the mean of the distribution of $b = -1$ since the undesired terms w is not negligible when $b = -1$;
- Threshold (3) considers both the distribution as Gaussian, while the

Table 2.2: Bit Error Rate for the 4 approximation methods to compute the optimal decision threshold as in (2.59).

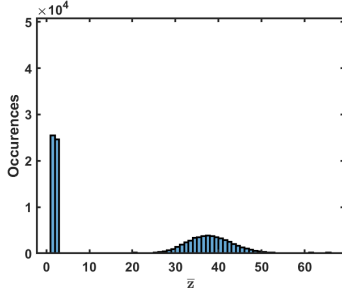
Methods	Number of Loops [L]						
	5	10	20	50	100	250	400
(1)	$6.2 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$	$1.0 \cdot 10^{-2}$	$4.0 \cdot 10^{-3}$	$1.5 \cdot 10^{-4}$	$< 10^{-5}$
(2)	$1.2 \cdot 10^{-1}$	$8.4 \cdot 10^{-2}$	$3.3 \cdot 10^{-2}$	$4.3 \cdot 10^{-3}$	$1.9 \cdot 10^{-4}$	$< 10^{-5}$	$< 10^{-5}$
(3)	$4.4 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	$1.5 \cdot 10^{-4}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$
(4)	$7.2 \cdot 10^{-4}$	$1.0 \cdot 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$



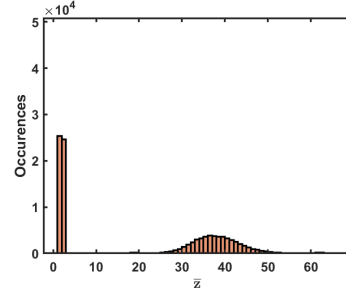
(a) Alice's decision variable.
Bob sends bits "+1".



(b) Alice's decision variable.
Bob sends bits "-1".



(c) Eve's decision variable. Bob sends bits "+1".



(d) Eve's decision variable.
Bob sends bits "-1".

Figure 2.5: Histogram of the values of **Alice's** [subfigs. (a) and (b)] and **Eve's** [subfigs. (c) and (d)] decision variable (2.55) with a number of loops $L = 100$ and 100 000 bits sent by Bob. The occurrences of the decision variable's values refer to the bits "+1" [subfigs. (a) and (c)] and "-1" [subfigs. (b) and (d)] sent by Bob.

distribution of $b = +1$ tends to be Chi-Square instead of Gaussian since the contribution of w is negligible;

- Threshold (4) is set considering the accurate distribution of $b = -1$ (Gaussian) and $b = +1$ (Chi-Square).

Table 2.1 shows the optimal threshold for each of the four proposed approximation methods (1)–(4) described after Eq.(2.59). It is worth noting that for large L the method (1) is equivalent to the theoretical decision variable (2.53), in fact when $L \geq 400$ the term $\sigma_{y_1}^2(t)$ in (2.16) converges to $\sigma_{y_1}^2(\infty)$ in (2.17) and becomes constant. As we can appreciate in Table 2.1, method (4) converges very quickly to the final optimal value of the threshold, compared to the other methods apart method (1) whose threshold is constant since it considers the PDF of $\bar{z}$ as symmetric.

Table 2.2 shows the Bit Error Rate (BER) over 100 000 bits for each of the methods (1)–(4) by using the thresholds depicted in Table 2.1. Method (4) is the most accurate, even for low number of loops L , as expected. Method (1) requires $L = 400$ to reach the best performance ($\text{BER} \leq 1 \times 10^{-5}$), while method (4) just requires $L = 20$.

Let us now show the security performance of the proposed ONL system and let us assume the worst case scenario, *i.e.*, Eve is able to recover the exact propagation delay between Bob and Alice. Bob sends 100 000 bits, half of those are $+1$, other half are -1 , both randomly distributed.

Figs. 2.5a and 2.5c show the occurrences of the values of Alice’s and Eve’s decision variable (2.55), respectively, over the bits “ $+1$ ” sent by Bob and with $L = 100$.

Analogously, Figs. 2.5b and 2.5d shows the occurrences of the values of Alice’s and Eve’s decision variable (2.55), respectively, over the bits “ -1 ” sent by Bob.

As it can be appreciated, the PDF of Alice’s decision variable in Figures 2.5a, 2.5b has only one region, while the PDF of Eve’s decision variable in Figs. 2.5c, 2.5d has always two regions. This means that Eve can only apply a threshold and decide which bit has been sent by Bob. Since the two regions spanned by the PDFs have the same integral, Eve decides for $+1$ or -1 with equal probability, although only $+1$ or -1 are sent by Bob to Alice. This entails that Eve experiences a $\text{BER} \approx 0.5$, when a binary signaling is used by the legitimate nodes.

2.1.6 Discussion

As a summary, the following conclusions can be derived from the numerical results:

- *Reliability* (the capacity of Alice to correctly demodulate the bit sent by Bob):
 - Methods (4) can lead to a very low BER (over 100 000 bits sent) even with a low number of loops $L \geq 10$.
 - Method (1) can lead to a very low BER (over 100 000 bits sent) only with a large number of loops $L \geq 250$; this means that the theoretical result is valid only for large L , as expected.
- *Security* (the capacity of Eve to correctly demodulate the bit sent by Bob to Alice):
 - Even in the very worst-case scenario, Eve is not able to correctly recover the sign of the bit sent by the legitimate nodes.
 - Eve experiences a $\text{BER} \approx 0.5$, while Alice can reach a $\text{BER} \leq 10^{-5}$ over 100 000 bits sent with $L \geq 10$.
 - The ONL modulation can be regarded as a full secrecy rate scheme, *i.e.*, each transmitted bit is secure.

2.1.7 Conclusion

A novel security technique, known as Optical Noise-Loop (ONL), is introduced as a PLS approach that ensures confidentiality by modulating information with locally generated noise. Unlike traditional encryption, ONL modulation does not depend on attacker's computational limits and eliminates the need for the prior exchange of a secret key, thus simplifying security protocols and reducing costs. ONL also enhances system resilience by making data indistinguishable from background noise, protecting them against physical attacks. While the NL concept is known, the theoretical foundation for AOP-based NL modulation are established, demonstrating its reliability and security in a OCvF system. The challenges in adapting NL modulation to fiber optical channels are identified, providing insights on necessary equipment, materials, and schemes.

2.2 Confidentiality in THz communications: effects of phase and molecular noise on Noise-Loop modulation

NL represents a promising PLS technique capable of enhancing the confidentiality also of THz communication systems, providing reliable security while preserving the extremely low latency properties of this transmission technology. The general properties of the Noise-Loop modulation are consistent with those described in Sec. 2.1 of this thesis. In general, NL modulation exploits the inherent thermal noise of the transmitting and receiving devices to modulate the transmitted bits [14]. However, in THz communication systems there are two relevant additional noise sources alongside thermal noise, namely, *molecular noise* and *phase noise*. Each of these noise source exhibits unique features within the THz band, the understanding of which is crucial for the design and optimization of secure and efficient THz communication systems. To demonstrate the effectiveness of NL modulation in THz systems, its essential to consider the effects of all three noise sources in the reliability and security performances of NL modulation.

Here, the effects of the aforementioned THz noise sources on the NL modulation are investigated, considering a THz link operating with a carrier frequency $f_0 = 237.7$ GHz and a bandwidth $B = 5$ GHz. This carrier frequency corresponds to the reference oscillator frequency used in [38], while the chosen bandwidth aligns with the IEEE 802.15.3d standard [39], which covers nearby carrier frequencies.

Related Works. The state of the art in physical layer security for terahertz communications encompasses a variety of cutting-edge techniques. These approaches, collectively contribute to the robustness of THz communication security. As technology continues to evolve, these innovations play a pivotal role in safeguarding data transmission, ensuring the privacy and integrity of information in an increasingly interconnected world.

In [40], ultra-narrow beams and multipath communication, which employ highly focused directional beams to reduce the risk of signal interception and enable redundancy and diversification of signal routes, are exploited, making it extremely challenging for unauthorized entities to compromise the communication link.

The authors in [41] exploit Intelligent reflecting surfaces (IRS), which

manipulate the propagation of THz signals by intelligently reflecting and controlling them, thereby reducing the risk of eavesdropping. These surfaces can adapt to changing communication conditions, making them a powerful tool in ensuring the security of data transmission.

2.2.1 Noise Sources in THz Communications

Thermal Noise

In general, wireless communications are susceptible to thermal noise, which deserves specific discussion when dealing with THz communications. This type of noise, characterized by Gaussian statistics, adds to the useful signal at different points between the source and the destination. However, here a single originating source of thermal noise is considered and it's added to the useful signal at the receiver's input. The PSD of Gaussian noise is [42]:

$$N(f) = \frac{h_p |f|}{2 \left(e^{\frac{h_p |f|}{k T_{\text{sys}}}} - 1 \right)} \left[\frac{W}{Hz} \right], \quad (2.61)$$

where h_p is the Planck's constant, k is the Boltzmann's constant, and T_{sys} is the system noise temperature. When $|f| \ll \frac{k T_{\text{sys}}}{h_p}$, an approximated, simpler, expression of (2.61) is obtained, which is usually denoted as N_0 :

$$N_0 \simeq \frac{1}{2} k T_{\text{sys}} \left[\frac{W}{Hz} \right]. \quad (2.62)$$

Since (2.62) is independent of f , thus being spectrally white, and considering the Gaussian and additive nature of this noise contribution, the thermal noise is commonly referred to as AWGN. However, under reasonable values of T_{sys} , the condition $|f| \ll \frac{k T_{\text{sys}}}{h_p}$ is not met in the THz band, rendering (2.61) the sole accurate expression for thermal noise PSD within this frequency range.

However, when considering a practical system operating within the THz band with a carrier frequency f_0 , it may still be reasonable to assume a constant PSD $N(f) \simeq N(f_0)$, provided that (2.61) remains relatively constant within the signal bandwidth. For example, if $N(f) \simeq N(f_0)$ is assumed with a noise figure $F = 12\text{dB}$ and a bandwidth B of 5 GHz centered around the carrier frequency $f_0 = 237.7\text{ GHz}$, as considered in this work, the maximum relative variation within this bandwidth is calculated

$$N_m(d, f) = \frac{h_p \pi |f|^3}{c^2} \left(e^{\frac{h_p |f|}{k T_0}} - 1 \right)^{-1} \left(\frac{c}{\sqrt{4\pi} r_c f_0} \right)^2 \frac{S(|f|)}{2} (1 - e^{-k(|f|)d}) \left(\frac{c}{4\pi r d f} \right)^2 \left[\frac{W}{Hz} \right], \quad (2.63)$$

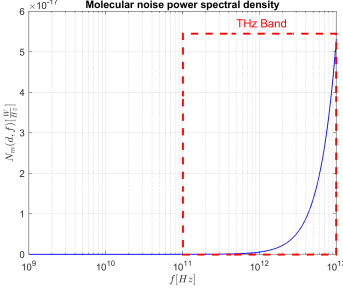
to be $\max_{f \in B} \frac{|N(f) - N(f_0)|}{N(f_0)} \cdot 100 = 0.0013\%$, confirming that the dependence on the frequency can be neglected under the given conditions. Therefore, our numerical results are based on the assumption of a flat spectrum for Gaussian thermal noise.

Molecular Noise

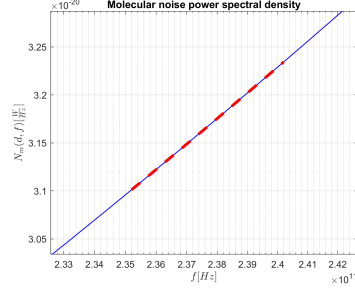
The molecular noise arises from the interactions between electromagnetic waves and the molecules of the medium through which it propagates. Specifically, the energy of the wave is partially absorbed by the molecules and then re-emitted into the channel, creating a disturbance. The extent of this phenomenon varies with frequency, resulting in a spectrum with distinct peaks at specific frequencies (known as resonance frequencies). Since the number of molecules, as well as the number of resonances per molecule, satisfies the assumptions of the central limit theorem, this noise exhibits Gaussian statistics. Its two-sided PSD is expressed as shown in equation (2.63) [43], where:

- d is the Tx-Rx distance,
- c is the speed of light in vacuum,
- T_0 is the reference temperature of the atmosphere,
- r_c is the index of refraction referring to the carrier frequency f_0 ,
- $S(f)$ is the one-sided PSD of the THz signal,
- r is the index of refraction seen by the THz wave,
- $k(f)$ is the medium absorption coefficient, representing the medium's absorption area of molecules per unit volume.

Fig. 2.6a shows an example of $N_m(d, f)$ for positive frequencies up to 10 THz. The parameter settings used for this plot are detailed in Table 2.3, where:



(a) Molecular noise PSD for positive frequencies up to 10 THz.



(b) Molecular noise PSD in a bandwidth $B = 5$ GHz centered at f_0 .

Figure 2.6: Molecular noise PSD.

Table 2.3: Parameter Settings for Fig. 2.6

Parameter	Value
f_0	237.7 GHz
B	5 GHz
d	1 m
T_0	290 K
P_T	-10 dBm
$k(f_0)$	$\simeq 5.91 \cdot 10^{-4} \text{ m}^{-1}$
r, r_c	1

- considering an array of CMOS-based THz sources utilizing frequency multipliers, capable of emitting around -10 dBm at 0.4 THz, a conservative approach sets $P_T = -10$ dBm at 0.2377 THz [44, 45];
- the value of $k(f_0)$ represents the average molecular absorption coefficient across a bandwidth of $B = 5$ GHz centered around a carrier frequency of $f_0 = 237.7$ GHz. Its value has been taken from the TeraSim simulator [46].

In Figure 2.6a, the THz band is highlighted by a dashed red box. It is evident that, unlike the thermal noise spectrum, $N_m(d, f)$ exhibits an increase within the THz band. For completeness, it should be noted that $N_m(d, f)$ approaches 0 for $f \gg 10^{13}$, which lies outside the THz band.

Let's focus on the spectrum characteristics within the band relevant to the presented work, *i.e.*, the 5 GHz band centered at $f_0 = 237.7$ GHz,

showed in Figure 2.6b. Within this band, it was found that the maximum deviation of the molecular noise's spectrum from $N_m(f_0)$ is approximately 2.1%, calculated as $\max_{f \in B} \frac{|N_m(f) - N_m(f_0)|}{N_m(f_0)} \cdot 100$. Given the small deviation within the considered band, it was deemed reasonable to assume $N_m(f) = N_m(f_0)$ within the bandwidth B that is of interest for this paper. Similar to the case of thermal noise, therefore, our numerical findings are based on the assumption of a flat spectrum for the Gaussian molecular noise.

Phase Noise

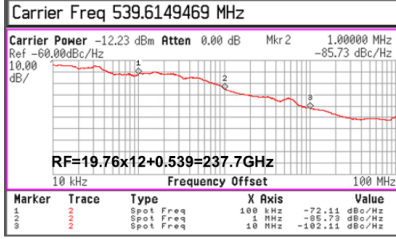
A key challenge in the design of THz communication systems is the need for high-frequency oscillators with low phase noise. The solution typically involves Phase-locked loop (PLL), which lock the output of a voltage controlled oscillator (VCO) with a highly stable reference input signal. This synchronization ensures that the oscillator's output matches the precision of the reference input [38, 47, 48]. However, despite their increased precision compared to free-running oscillators, PLL also exhibit pn. Specifically, when the oscillator experiences white noise, the phase noise is determined by the combination of a cumulative Wiener process (a continuous-time, one-dimensional Gaussian random walk) and an uncorrelated Gaussian process. The first type, characterized by a Lorentzian spectrum, *i.e.*, $1/f^2$ [49], arises from the integration of white noise within the PLL. The second type, on the other hand, stems from its amplification and is characterized by a white spectrum [50, 51].

In this regard, Fig. 2.7a depicts the PSD of the PLL-based frequency synthesizer operating at 237.7 GHz that serves as a reference in the presented work [38]. Notably, the spectrum exhibits a colored portion that decreases with frequency, along with a flat, *i.e.*, white, section, noticeable from approximately 30 MHz onwards. This shape is typically modeled as illustrated in Fig. 2.7b. This model employs two parameters: K_2 and K_0 , which respectively determine the slope of the PSD of the Wiener PN and of the uncorrelated, *i.e.*, white, portion of the spectrum, respectively [51].

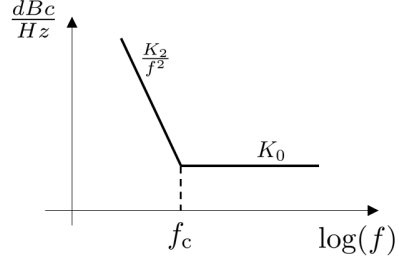
Since the PN here considered is the result of both a cumulative Wiener process and an uncorrelated Gaussian process, each p -th PN sample in our simulator is determined by

$$\phi_p = \phi_{w,p} + \phi_{g,p}, \quad (2.64)$$

where:



(a) PN spectrum of the PLL-based frequency synthesizer, operating at 237.7 GHz. This figure was taken from [38].



(b) Model of the PN spectrum at 237.7 GHz corresponding to the PLL proposed in [38].

Figure 2.7: PN characteristics.

- $\phi_{g,p} \sim \mathcal{N}(0, \sigma_g^2)$ represents the uncorrelated Gaussian PN at instant p , with $\sigma_g^2 = K_0 B$ [51];
- $\phi_{w,p}$ represents the Wiener PN at instant p and is given by $\phi_{w,p} = \phi_{w,p-1} + \delta\phi_{w,p}$, with $\delta\phi_{w,p} \sim \mathcal{N}(0, \sigma_w^2)$, $\sigma_w^2 = 4\pi^2 K_2 T_s$ [49, 51], being T_s the sampling interval.

Considering Fig. 2.7b, it's noteworthy that the variances σ_g^2 and σ_w^2 are constrained by the corner frequency $f_c = \sqrt{K_2/K_0}$.

As anticipated, the oscillator described in [38], designed to operate at 237.7 GHz, was taken as a reference in this work to model the PN. The corresponding PN spectrum, shown in Fig. 2.7a, was modeled as shown in Fig. 2.7b. In particular, comparing Fig. 2.7a and Fig. 2.7b, it is clear that, for this oscillator, it is $f_c = 30$ MHz and $K_0 = -110$ dBc/Hz. These values served as a reference for our investigation. Further details on phase noise modeling can be found in [52].

2.2.2 System Model

The Noise Loop in the THz communications scheme considered in this paper is illustrated in Figure 2.8. Without loss of generality, in Figure 2.8 the detailed models of phase noise and demodulation are shown for Alice, noting that Bob employs an analogous architecture. The working principle of NL modulation for RF systems has already been described in [53]. Here, the focus is solely on the novel mathematical treatment introduced by the

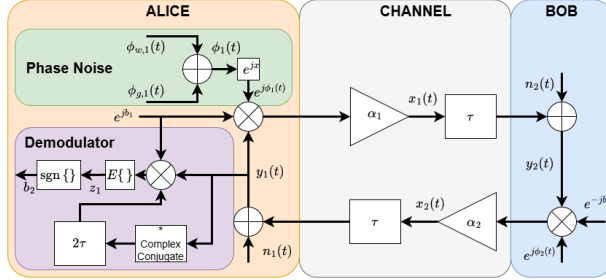


Figure 2.8: Noise Loop modulation scheme. The THz channel noise sources are included.

specific characteristics of noise sources in THz communications, as applied to the noise loop technique.

The system model is composed of two legitimate nodes, Alice and Bob, which simultaneously transmit their respective bits b_1 and b_2 . Each terminal receives a signal, *i.e.*, $y_1(t)$ and $y_2(t)$ for Alice and Bob respectively, that is a combination of the signal sent by the other terminal and the thermal noise (n). The channel's effects include an attenuation (α) and a time delay (τ). The received signal is modulated with binary phase-shift keying (BPSK) modulation, where b_1 and b_2 can be either 0 or π . In the context of THz communications, the signal, prior to being re-transmitted by another terminal, is also affected by a phase noise ($e^{j\phi}$).

Without loss of generality, let's focus the terminal Alice. To demodulate the information bit transmitted by Bob, Alice must first calculate the autocorrelation function, as described in the following subsection 2.2.3. In a discrete NL system, the expectation is defined as the average of a finite number of samples, corresponding to the finite number of loops that the two transmitters must repeat the same bit.

2.2.3 Reliability analysis

As discussed in Sec. 2.2.1, the molecular noise $n_m(t)$ is modeled as additional white Gaussian noise, which can be incorporated into the NL modulation scheme by adding it to the thermal noise $n_{th}(t)$, whose characteristics have been discussed in Sec. 2.2.1. In other words, in the NL scheme, our white Gaussian noise source consist of two components, *i.e.*, $n_i(t) = n_{th,i}(t) + n_{m,i}(t)$, where $i = 1, i = 2$ denotes Alice and Bob, respectively. The phase noise, on

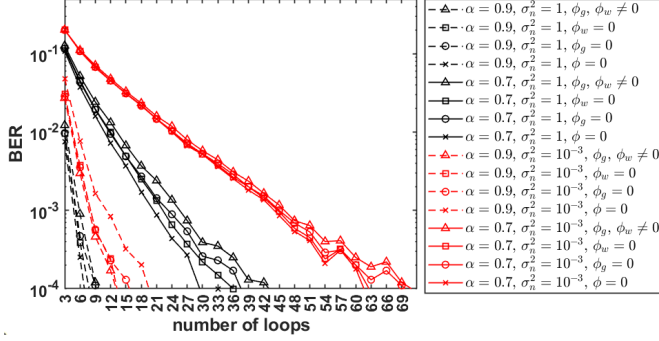


Figure 2.9: BER as a function of the number of loops L .

its hand, is modeled with two components: an uncorrelated Gaussian noise $\phi_g(t)$ and a Wiener noise $\phi_w(t)$, which are included into the information bits (see Figure 2.8) as a noise source $\phi_i(t) = \phi_{w,i}(t) + \phi_{g,i}(t)$, with $i = 1, 2$.

Without loss of generality, let's focus on Alice receiving stage. The signal $y_1(t)$ received by Alice at time t can be written as an auto-regressive model

$$y_1(t) = n_1(t) + \alpha_2 e^{jb_2} e^{j\phi_2(t-\tau)} n_2(t-\tau) + \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{j\phi_1(t-2\tau)} e^{j\phi_2(t-\tau)} y_1(t-2\tau). \quad (2.65)$$

In order to demodulate the information bit b_2 sent by Bob, Alice has first to compute the autocorrelation function

$$R_{y_1}(2\tau) = \mathbb{E}\{y_1(t)y_1^*(t-2\tau)\} = \alpha_1 \alpha_2 e^{jb_1} e^{jb_2} e^{j\phi_1(t-2\tau)} e^{j\phi_2(t-\tau)} \sigma_{y_1}^2. \quad (2.66)$$

The expectation operator $\mathbb{E}\{\cdot\}$, in practical implementation, is replaced by an average over a pre-determined number of loops L , i.e., $\mathbb{E}\{y_1(t)y_1^*(t-2\tau)\} = \frac{1}{L} \sum_{l=1}^L y_1(t)y_1^*(t-2\tau)$.

After multiplication by e^{-jb_1} to remove its own information bit, Alice can estimate Bob's bit by simply extracting the sign

$$\text{sgn}\{z_1\} = \text{sgn}\{e^{-jb_1} R_{y_1}(2\tau)\} = \text{sgn}\left\{\alpha_1 \alpha_2 e^{jb_2} e^{j\phi_1(t-2\tau)} e^{j\phi_2(t-\tau)} \sigma_{y_1}^2\right\} \quad (2.67)$$

where z_1 is the decision variable used by Alice to extract the information bit sent by Bob. Since α_1 , α_2 and $\sigma_{y_1}^2$ are always positive, the sign of (2.67) depends only by Bob's information bit b_2 (which can assume only values 0 or π), but it is corrupted by the THz phase noise $e^{j\phi_1(t-2\tau)}e^{j\phi_2(t-\tau)}$ of Alice and Bob.

Let us observe that the term $\sigma_{y_1}^2$, which appears in (2.67), can be written as follows

$$\sigma_{y_1}^2 = \mathbb{E}\{y_1(t)y_1(t)^*\} = \sigma_{n_1}^2 + \alpha_2^2\sigma_{n_2}^2 + \alpha_1^2\alpha_2^2\sigma_{y_1}^2 = \frac{\sigma_{n_1}^2 + \alpha_2^2\sigma_{n_2}^2}{1 - \alpha_1^2\alpha_2^2} \quad (2.68)$$

where $\sigma_{n_i}^2 = \sigma_{\text{th},i}^2 + \sigma_{\text{m},i}^2$ with $i = 1, 2$ is the sum of the variances of the thermal and molecular noise at node i , $\mathbb{E}\{\cdot\}$ is the expectation operator and $*$ is the conjugate operator. As it can be seen, $\sigma_{y_1}^2$ is not affected by the phase noise, but only by the thermal noise and by the molecular noise.

Still focusing on (2.67), in order to better understand the numerical results provided in the following, let us recall, that each phase noise contribution is composed of two components

$$\phi_1(t) = \phi_{1,\text{w}}(t) + \phi_{1,\text{g}}(t) \quad (2.69)$$

$$\phi_2(t) = \phi_{2,\text{w}}(t) + \phi_{2,\text{g}}(t) \quad (2.70)$$

as described in Sec. 2.2.1.

2.2.4 Results

This section describes the numerical results of the NL scheme. The results are focused on the reliability performance in order to check the robustness against the THz noise sources. Simulations have been carried on with Simulink on Matlab 2023b.

The numerical simulation parameters are the shown in Table 2.4.

In Figure 2.9 the performance in terms of BER (Bit Error Rate) as a function of the number of loops of the NL modulation is reported. Basically, the graph shows two group of curves: the red lines are related to low thermal and molecular noise in both the NL nodes $\sigma_{n_1}^2 = \sigma_{n_2}^2 = \sigma_n^2 = 10^{-3}$, while the black lines are related to high thermal and molecular noise $\sigma_n^2 = 1$.

When the additive white Gaussian noise (sum of thermal and molecular noises) is high, the presence of a phase noise of any kind (Wiener and/or Gaussian) degrades the performance of the legitimate NL receiver (black

Table 2.4: Simulation Parameters Setting.

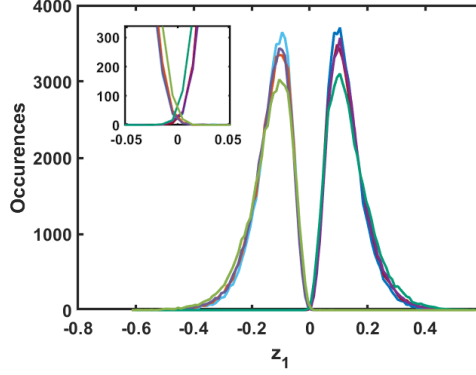
Parameter	Value
α	0.7 or 0.9
σ_n^2	1 or 10^{-3}
B	5 Ghz
K_0	-110 dB
f_c	30 MHz
σ_w^2	$5 \cdot 10^{-2}$
σ_g^2	$\sim 7 \cdot 10^{-5}$

solid lines). The Wiener phase noise component ϕ_w affects the performance more than the Gaussian phase noise component ϕ_g . Changing the loop gain $\alpha_1 = \alpha_2 = \alpha$ from 0.7 to 0.9 (black dashed lines) improves the performance, making the BER lines closer to each other without changing the previous conclusion.

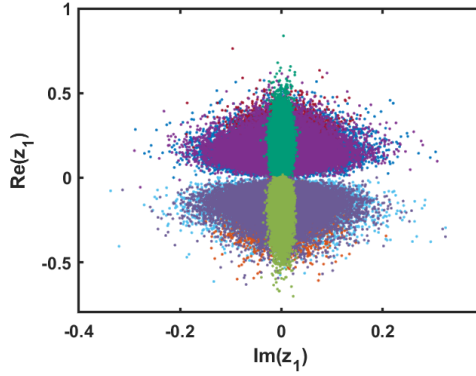
When the additive white Gaussian noise is low and the loop gain α is 0.7, any type of phase noise does also deteriorate the performance of the legitimate NL receiver (red solid lines). The Wiener phase noise component ϕ_w affects the performance more than the Gaussian phase noise component ϕ_g . Otherwise, changing the loop gain α from 0.7 to 0.9 (red dashed lines), it makes the presence of a phase noise beneficial for the legitimate NL receiver.

The benefits of this are illustrated in Fig. 2.10, which depicts the system's behavior with 12 loops ($L = 12$). In Fig. 2.10b, the impact of phase noise increases the imaginary component of the decision variable z_1 , more so than the real part of the decision variable. This effect narrows the distribution width of the decision variable's real part, as demonstrated in Fig. 2.10a and more clearly in the enlarged. Another consequence of phase noise is the shift of the distribution curve towards 0.

In summary, if the shift of the distribution curves exceeds the reduction in their width, as observed when α equals 0.7 or σ_n^2 is high, the phase noise impairs the performance of the legitimate NL receiver. Otherwise, when the additive noise is low and the decrease in the distribution width is more significant than the shift of the distribution curves, the phase noise actually enhances the performance of the NL system.



(a) Distribution of decision variable z_1



(b) In-Phase and In-Quadrature components of decision variable z_1

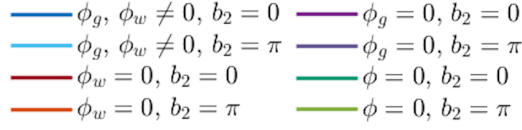


Figure 2.10: The representation of the decision variable z_1 in the particular case where $\alpha = 0.9$, the thermal and molecular noise ($\sigma_n^2 = 10^{-3}$) is comparable to the phase noise. ($L = 12$)

2.2.5 Discussion

In general the presented results can be summarized by defining two noise regimes:

- *Low additive gaussian noise*: receiver performances are degraded by phase noise, in particular by the Wiener component ϕ_w . Increasing the loop gain from $\alpha = 0.7$ to $\alpha = 0.9$ mitigates this degradation slightly, but does not reverse the trend.
- *High additive gaussian noise*: Phase noise degrades the performances of the receiver. However, raising the attenuation from $\alpha = 0.7$ to $\alpha = 0.9$ BER is improved by the effect of phase noise. This behavior is due to the increase of the imaginary component of the decision variable z_1 caused by the phase noise, which narrows the distribution of its real part and shifting it toward zero. If the narrowing outweighs the shift, performance improves.

Concerning security performances of NL in THz communications systems, the capability of the NL modulation to produce full secrecy rates has been already demonstrated in [14]. The presence of additional noise sources of the THz channel can only worsen the demodulation capabilities of the eavesdropper (Eve).

2.2.6 Conclusion

The implementation of the NL modulation for THz communications channels have been investigated, by analyzing the effects of THz communications-specific noise sources, *i.e.*, AWGN, molecular noise and phase noise. Through simulations and mathematical analyses, the conducted study reveals that molecular noise exhibits gaussian statistics, hence it is incorporated in the AWGN source. The main effects on the reliability of NL modulation in THz systems are associated to phase noise. Depending on the magnitude of AWGN and on the loop gain the degradation caused by the phase noise, expressed in terms of BER performance, varies. In case of a high AWGN and loop gain the BER of the system results improved.

Chapter 3

Physical Layer Security Solutions for WBANS Transmission Technologies

In this chapter the application of PLS to WBANS transmission technologies is investigated.

In particular, an analysis of the security framework of WBANS including a comparison of the security approaches proposed by IEEE and the European Telecommunications Standards Institute (ETSI) is provided.

An analysis of the physical layer security threats in MCvD is presented. The main security threats known in classical telecommunications are mapped in MCvD context. Specifically, the analysis focuses on how these security threats may manifest within the main modulation schemes of MCvD. Finally, an analysis of the complexity of these attacks, with reference to the various MCvD schemes, is provided.

The modeling of in In-to-Out Body channel model for a VLC-based IMD system is presented. The performance of reliability and PLS confidentiality of the system have been evaluated in terms of Bit Error Rate and Secure area, respectively.

The results of the application of a new GC-based biometric authentication protocol, based on the usage of an LSTM neural

network, are presented.

Finally, the problem of authenticating outputs of human body is approached. A PLS-based approach which exploits tools of the information theory is implemented to address the problem of distinguish AI and Human artworks.

The results presented in this chapter have been published or submitted in the following papers:

- **G. Borghini**, S. Caputo, L. Mucchi, A. Rashid, S. Jayousi, M. Hamalainen, T. Paso, M. Hernandez. “Security of wireless body area networks for healthcare applications: Comparison between ETSI and IEEE approaches”, in *Proc. of IEEE 17th International Symposium on Medical Information and Communication Technology (ISMICT)*, Lincoln, Nebraska (USA), May 10-12, 2023 [11].
- **G. Borghini**, S. Caputo, S. Jayousi, M. Magarini, M. Pierobon, L. Mucchi. “Security Threats in Diffusion-based Molecular Communication Systems”, in *Proc. of IEEE 19th International Symposium on Medical Information and Communication Technology (ISMICT)*, Florence, Italy, May 28-30, 2025 [5].
- S. Caputo, **G. Borghini**, S. Jayousi, L. Mucchi. “Visible Light Communications for Implantable Medical Devices”, in *Proc. of EAI International Conference on Body Area Networks*, Varanasi, India, Dec 15-16, 2024 [3].
- S. Caputo, A. Viziello, A. Coviello, M. Magarini, **G. Borghini**, P. Savazzi, S. Jayousi, L. Mucchi. “Biometric Authentication in Galvanic Coupling Communications: A Physical Layer Approach for Secure Wearable Medical Systems”, accepted at *EAI International Conference on Body Area Networks (BodyNets)*, Okayama, Japan, Dec 2-4, 2025 [12][Accepted].
- **G. Borghini**, A. Enersen, S. Caputo, M. Pierobon, L. Mucchi. “Measuring Aesthetic Information in AI-Generated Artistic Images: A Tool for Authorship Detection”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2025. [6][Submitted]

3.1 Security framework for Wireless Body Area Networks: Comparison Between ETSI and IEEE approaches

WBANs are among the key technologies for 6G mobile networks, as they are considered a potential candidate to serve as an intermediate gateway layer between in-body and on-body devices and the broader mobile network.

As a contribution of this thesis an overview of how security is approached in WBANs-related standards is presented, realizing a comparison between the standard TR 103.638, also known as SmartBAN, defined by the European Telecommunications Standards Institute (ETSI) for WBANs, and the other IEEE standards.

WBAN is a network used to monitor and track data gathered by various sensors positioned inside or outside the human body.

The general WBAN architecture is depicted in Fig.3.1. In WBANs, communications are divided in three tiers:

- *Body Tier* - It is responsible for managing the communications between in-body or on-body sensors and the hub. The hub node collects and aggregates data from the in-body or on-body sensors. The body tier is responsible for ensuring that the data collected from the sensors is transmitted accurately and efficiently to the coordinator device which is located at the Edge tier.
- *Edge Tier* - The tier between the Body and Cloud tiers in WBANs is responsible for initial data processing, real-time analysis, data filtering, and compression. It includes coordinator nodes such as laptops, smartwatches, and smartphones. This tier reduces latency and power consumption and enables quick response to critical events, making it crucial for patient safety in healthcare applications.
- *Cloud Tier* - This tier is where the WBAN data flows through a gateway and Internet connection. It provides additional processing power, storage capacity, long-term analysis, data backup, and visualization. This tier is crucial in medical applications, where substantial amounts of data need to be processed and analyzed for diagnosis and treatment purposes. It enables the development of advanced applications and algorithms for processing the data and enables the sharing of data and information between healthcare providers, caregivers, and patients.

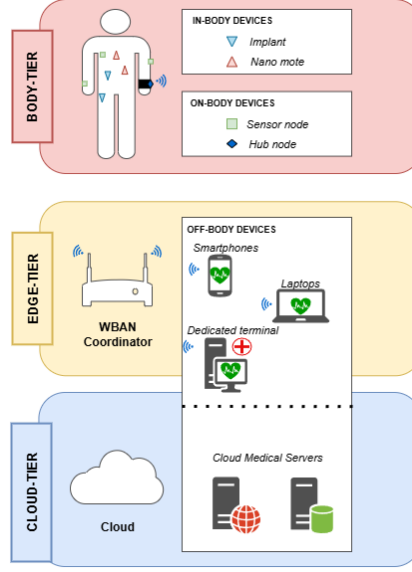


Figure 3.1: WBAN Reference Architecture.

The above mentioned tiers present various security challenges and threats, which can compromise the security of personal health data. The body tier presents unique security challenges, due to the low processing, low memory, low power capabilities of the devices. Security threats could involve threats such as eavesdropping, tampering, cloning, DoS, spoofing, and others. The other two tiers (edge and cloud) are considered to be significantly more secure than the body tier because in these tiers such constraints are not present on the participating devices and efficient cryptography and authentication techniques can be applied.

3.1.1 Standards Suitable for WBANs

Various standards can be used for WBANs, although not all of them have been developed explicitly for such kinds of networks. These standards may specify features, such as encryption algorithms, authentication protocols, access control policies, and data integrity mechanisms. Here is a list of reliable standards from the IEEE :

- **IEEE 802.15.1/Bluetooth** [54]: This standard specifies the Blue-

tooth wireless communication protocol, which can be used for short-range communication in WBANs.

- **IEEE 802.15.4** [55]: This standard defines the physical and Media Access Control (MAC) layer specifications for low-power, short-range wireless communication, which may be used in WBANs.
- **IEEE 802.15.6** [56]: This standard specifies the requirements for wireless communication in WBANs, including physical and MAC layers specifications, and security protocols.
- **IEEE 802.11** [57]: This standard specifies the Wi-Fi wireless communication protocol, which may be used as an access point for WBANs.
- **IEEE 11073** [58]: family of standards provides real-time plug-and-play interoperability for medical devices external computer systems, including personal devices, independent of the data link, which may be wired or wireless.

SmartBAN is an ETSI standard developed for wireless personal communication networks for medical and non-medical applications. [59]. The SmartBAN and IEEE 802.15.6 standards target the same personal network applications, including medical and healthcare use cases. The SmartBAN defines the physical and MAC layer specifications for wireless communication between medical devices, including body-worn sensors and implants, and gateways or other medical devices. SmartBAN is designed to support low-power and low-latency communications, with a focus on minimizing interference and power consumption. SmartBAN aims to include specific security measures for healthcare applications, such as a standardized mechanism for authenticated device pairing and encryption. Moreover, it aims to include data confidentiality, data integrity, and data origin authentication, to ensure that medical data is transmitted securely and protected from unauthorized access or malicious attacks.

These standards provide a framework for the development of WBAN devices, ensuring that they can communicate with each other effectively and securely. They are continually being updated and expanded to keep pace with the latest technological advancements and address emerging challenges in the field of healthcare.

The study proposed in this thesis is limited to the IEEE 802.15.4 and IEEE 802.15.6 standards, because they both share the 2.4 GHz Industrial,

Scientific, and Medical (ISM) frequency band with the ETSI SmartBAN standard. A comparison between security approaches of ETSI SmartBAN and IEEE 802.15.4 and 15.6 standards is reported in Section. 3.1.6.

3.1.2 Security in IEEE BAN-related standards

IEEE 802.15.4 Security

This standard defines the physical and MAC layers, two communications modes, and three security modes such as unsecured mode, secured mode, and Access Control List (ACL) mode, where a list of devices allowed to communicate with the node is exploited to filter incoming data, for Low-Rate Wireless Personal Area Networks [55]. The communications mode is based on a non-beacon-enabled mode where each node operates independently and does not require time synchronization with other nodes. The second communications mode is the beacon-enabled mode which requires time synchronization, with the coordinator periodically transmitting beacons to maintain network synchronization.

In the unsecured mode, no security services are provided by the devices. Limited security services for communications are provided in ACL mode where the higher layer may choose to reject frames based on whether the MAC sublayer indicates that a frame is aimed to originate from a specific device, but cryptographic protection is not provided in the MAC sublayer. The higher layer should implement other mechanisms to ensure the identity of the sending device. In the Secure mode, devices are fully secure at the MAC layer by using the conventional Advanced Encryption Standard (AES) algorithm, which encrypt and decrypt data using the same secret key which can have a length of 128,192 or 256 bit. Security services such as access control, data encryption, data authenticity, data integrity, replay protection, and sequential freshness are provided in this mode. Security services such as access control, data encryption, data authenticity, data integrity, replay protection, and sequential freshness are provided in this mode.

IEEE 802.15.9 [60] Note that [55] supports only datagram security, which is security techniques applied at the transport layer of the ISO/OSI stack. This standard defines Key Management Protocol (KMP) approaches that meet the critical needs of an effective security framework. It specifies the transportation of KMP datagrams within the IEEE 802.15.4 network. It

also provides guidelines for KMPs such as the Internet Engineering Task Force (IETF) host identity protocol (HIP), Internet Key Exchange version 2 (IKEv2), Protocol for carrying Authentication for Network Access (PANA)

IEEE 802.15.4y [61] This standard provides a mechanism for using other encryption algorithms in the IEEE Std 802.15.4 than AES-CCM-128. Depending on the KMP to be used this standard can be used to negotiate the algorithm.

In addition to the above work, the 802.15.4 standard has other noteworthy network capabilities such as CSMA-CA (Carrier Sensing Multiple Access - Collision Avoidance) mechanism, a multiple access protocol designed to avoid collisions by verifying the absence of any other transmission on the channel, *reliability* by using frame acknowledgments, where the receiver sends a confirmation message that the data has been received correctly, *data verification* by using ITU-T Cyclic Redundancy Check (CRC), exploited to check the integrity of the data and verify they have not been modified by an hostile node. It is important to note that this standard does not provide security at the physical layer.

IEEE 802.15.6 Security

This standard is based on three distinct levels of security, each offering unique security features, protection levels, and frame structures [56] [62].

The first is the *unsecured communication level* in which the exchange of frames is unsecured. This level is vulnerable and provides an opportunity for malicious actors to perform malicious activity in the network. The second is known as *authentication without encryption level*. It provides a moderate level of security such as authentication of the transmitted data. This level provisions message authenticity, integrity, and replay protection, but there is no confidentiality and privacy. The third level is known as *authentication with encryption level*. It offers the highest level of security at the MAC level. Frames are sent authenticated and encrypted by using the AES algorithm. This level provides message authenticity, integrity, confidentiality, privacy, and replay protection.

The security mechanisms of IEEE 802.15.6 are typically based on the Diffie-Hellman key exchange. 256-bit integers that are unique and independent should be used as private keys. To derive Key Message Authentication Codes (KMAC) and master key (MK), the Cipher-based Message Authen-

tication Code (CMAC) is utilized. Initially, a pre-shared MK is available to both the node and hub to facilitate the security association procedure. The node initiates the process by transmitting a security association frame request to the hub. The hub can choose to either join or abort the association procedure. If the hub responds by aborting, the node stops the current association procedure. On the other hand, if the hub responds by joining, the pre-shared MK is activated and shared between the node and hub after mutual agreement, which is then employed to generate a pairwise temporary key (PTK). It is important to note that IEEE 802.15.6 either does not provide security at the physical layer.

Both standards provide different levels of security at the MAC level but there is no mechanism to secure the data at the physical level. In addition to this, IEEE 802.15.4 [55] can form a star or peer-to-peer topologies, and together nodes can establish the cluster tree network where a single node *Coordinator* is responsible for network security and management of the whole network. Similarly, IEEE 802.15.6 [63] can form the star topology where a *HUB* is responsible for the entire network. According to [64], the *Availability* which is the fundamental security requirement is not addressed in both standards and the central node is at security risk because both standards have a single point of failure (SPoF).

3.1.3 Security in ETSI SmartBAN

The ETSI SmartBAN architecture is first outlined, after which attention is directed to the associated security mechanisms.

3.1.4 SmartBAN Architecture

Currently, the SmartBAN network topology is organized as a single-hop star. The center of the topology is the SmartBAN hub which acts like a sink node and gathers data from all sensing node and carries out an initial analysis and processing of those data. Then the hub is connected to the SmartBAN coordinator which acts as the gateway of the SmartBAN and provides connection with external networks, *e.g.*, 6G network. As a special case, a single device can contain both the hub and the coordinator. SmartBAN standard also provides for the creation of communication channels between neighboring hubs (for BAN to BAN communication). At the moment, hubs

communicates directly with each other without involving the coordinator. According to current development of SmartBAN standard, the hub implements the functionalities of the first two levels of the ISO/OSI stack, physical and MAC, and the coordinator implements the higher levels.

The sensor nodes are in general very low complex devices: in the simplest configuration, a node may include only the radio-frequency communication equipment and the dedicated sensing functionalities. It is important to notice that the presence of a central coordinator and a hub who deal with all the elaboration and the connections outside the SmartBAN allows updating of the SmartBAN functionalities simply by updating the coordinator. SmartBAN reference architecture also identifies four communications layers (ComLayer) placed between the network tiers as depicted in Fig. 3.2.

1. *ComLayer-1*: Given that WBAN envisions the presence of nano-scale devices or implants placed inside a body, this ComLayer includes the links between devices placed inside the body, which may also need to communicate with each other.
2. *ComLayer-2*: Similarly to ComLayer-1, this layer includes communications between devices of body tiers. In particular, ComLayer-2 contains communications between in-body and on-body devices. It can be imagined as a communication layer that allows data exchange between nano-devices or implants and a hub node placed outside the body.
3. *ComLayer-3*: This ComLayer comprehends the communications between on-body and off-body devices. It can be represented as communications between the sensors on-body or in-body, through a hub node, and the coordinator/gateway. In other words, this ComLayer interfaces the Body tier with the edge tier.
4. *ComLayer-4*: Through this ComLayer the data gathered in or on the body are sent outside the BAN to a cloud (e.g. a medical server). So, this layer links the Edge tier to the Cloud tier.

ETSI SmartBAN also defines two main categories to subdivide the ComLayers. *i) intra-WBAN*: includes communications between the sensors and the gateway; *extra-WBAN*: includes communications directed from the gateway outside the WBAN, to the Internet [65].

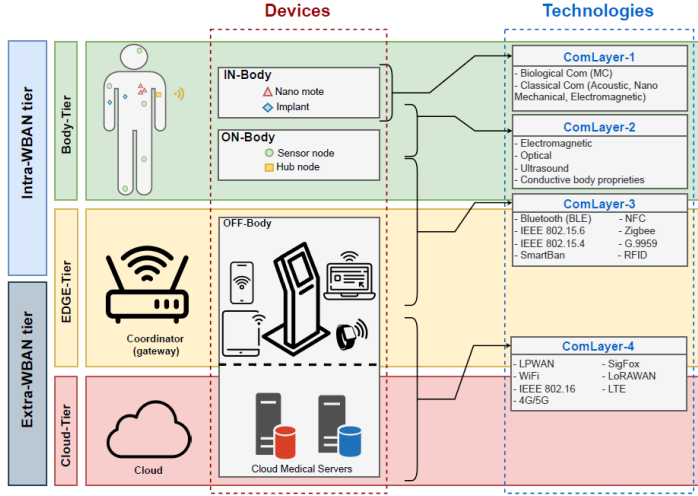


Figure 3.2: SmartBAN Architecture.

3.1.5 ETSI SmartBAN Approach to Security

The security specification for SmartBAN is currently under development. In this chapter, the current state of the work is reported. SmartBAN follows the security requirements defined on ETSI EN 303 645 developed by TC Cyber, which deals with baseline security requirements for consumer Internet of Things (IoT). The most relevant security requirements contained in ETSI EN 303 645 standards are listed below:

- *No universal default password* - When passwords are used to authenticate an IoT device, every consumer IoT device password must be unique per device or defined by the user. In the first case, it's also requested that the password generation procedure must be such to reduce the risk of automated attacks against a class or type of device.
- *Minimize exposed attack surfaces* - In accordance with the 'principle of least privilege' (a device should have access only to the resources it needs to complete its task), TC Cyber standard provides that the unused network and logical interfaces shall be disabled. Also, only software services used or required for the intended use or operation of the device must be enabled.

- *Ensure that personal data is secure* - Some IoT applications involve the transit of potentially sensitive information (such as medical information or payment information). So the devices must be designed in order to protect that information by means of best-practice cryptographic techniques.
- *Make system resilient to outages* - System resilience is another important aspect taken into account by TC Cyber. Firstly the design of IoT devices shall take into account the possibility of outages of data networks and power. The system shall be able to remain operational and locally functional in case of a loss of network access and should recover cleanly in the case of restoration of loss of power.

Starting from ETSI EN 303 645, the coming SmartBAN standard is defining additional security features. In particular, the SmartBAN standard is mainly focused on ComLayer-1 and ComLayer-2, which are less investigated in the current literature. In particular, a very critical aspect of the security of ComLayer-1 and ComLayer-2 is to thwart eavesdropping, which is one of the main security threats at the Body tier in WBAN.

Devices used to gather data from inside and outside the body may be simple and low-resourced and unable to implement high-level security techniques for data confidentiality, such as encryption. At the same time, the information sent or received at this level is extremely sensitive and needs a high degree of protection. Also, authentication is a security issue for the WBAN body layer. A strong security protocol should require devices connected to a HUB to be authenticated more than once, periodically. Finally, in order to reach a higher level of security, it's necessary to implement lightweight techniques for the generation and distribution of secret keys. To address this problem, SmartBAN standard envision the implementation of Physical Layer Security techniques. However, PLS techniques envisioned by the SmartBAN standard offer a new perspective that will help to solve those security issues.

Another very interesting feature envisioned in the ETSI SmartBAN standard is the integration of the Molecular Communications (MC) paradigm in the context of in-body communications.

SmartBAN security considerations Here are briefly listed and explained some other security features envisioned by the ETSI SmartBAN

Table 3.1: Security Features Comparison of WBAN-related Standards

Standards	Security Features					
	Encryption	Message Integrity	Secure Authentication	Availability	Security mode only	Physical Layer Security
IEEE 802.15.1	✓	✗	✓	✗	✓	✗
IEEE 802.15.4	✓	✗	✓	✗	✗	✗
IEEE 802.15.6	✓	✓	✓	✗	✗	✗
ETSI TR 103 638	✓	✓	✓	✓	✓	✓

standard.

- *Perfect forward security* - In general, in IEEE 802.15.6 standard the device and the coordinator of the WBAN establish a symmetric session key. This can cause problems if it does not provide perfect forward security. Since a key identifies the user, in case of disclosure of the long-term secret key, a hostile node may impersonate a legitimate node or decrypt messages. SmartBAN should prevent the above scenarios.
- *Device anonymity* - An adversary node that is observing communications, should not be able to learn the identity of any device in a SmartBAN network.
- *Session unlinkability* - An adversary node that is observing communications should not be able to link a communication session of any device to another completed session of the same device.

3.1.6 Comparison of security approaches: ETSI SmartBAN Vs IEEE 802.15.1, 15.4 and 15.6 standards

In this section a comparison between the IEEE standards analyzed in Section 3.1.2 and the ETSI SmartBAN showed in Section 3.1.3 is presented. A summary of the comparison can be found in Table 3.1. Let's start with a direct comparison of the standards that deal specifically with WBANs, ETSI SmartBAN, and IEEE 802.15.6. The first aspect that emerges is the fact that 802.15.6 provides three different security modes, including a no-security mode where no security measures are applied to communications. In ETSI there is no such thing as security is always active. A second aspect that differentiates ETSI and IEEE approach to security in WBAN is that the SmartBAN standard envisions the implementation of Physical Layer Security techniques. Finally, a third difference is that the ETSI SmartBAN standard envisions an extended view of in-body communications as it provides for the implementation of the Molecular Communications paradigm.

3.1.7 Conclusion

Security threats, solutions, and challenges in WBANs for healthcare are critical areas of concern as WBANs are often used to collect and transmit sensitive personal health information. It's important to note that security

in WBAN technology is an ongoing process, as new threats and vulnerabilities emerge over time. The presented work focuses on analyzing how current telecommunications standards applied to WBAN approach security problems. In particular the analysis carried out concerns two of the major WBAN standards, *i.e.*, ETSI SmartBAN and IEEE 802.15.6. Also in this work are analyzed other communications standards which may be involved in the design of a WBAN system. The comparison of the security features, as well as the future directions, reveals how SmartBAN is trying to answer to the challenges that current and future in- and on-body communications will have to face.

3.2 PLS threats analysis in Molecular Communications

Molecular Communications are an emerging transmission technique where, unlike traditional telecommunication systems, signals are transmitted through optical or radio-frequency waves, information is conveyed through molecules, *i.e.*, a chemical signal.

A general MCs system model is presented in Figure 3.3. The transmitted information is carried by molecules. The molecules are emitted by a bio-nanomachine (BnM), *e.g.*, a biological cell, which operates as a transmitter, and gathered by another bio-nanomachine which operates as the receiver. Both the transmission and the reception of molecules happens through chemical interactions. The communication channel consists of the environment on which the information molecules flow.

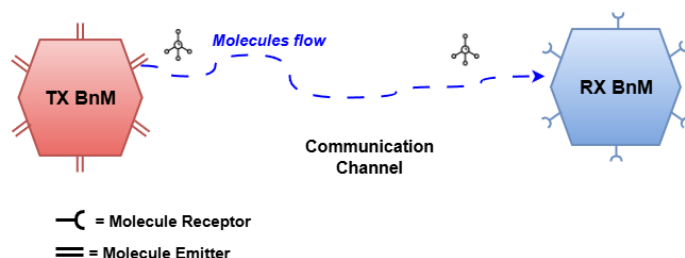


Figure 3.3: Molecular Communication system model.

To control the communication process, and thus being able to transfer information, the bio-nanomachine should have a predictable and controllable behavior. The transmitting bio-nanomachine should be capable of synthesizing, storing and emitting the information molecules of interest. Analogously, the receiving bio-nanomachine should be able to sense these information molecules. Such a bio-nanomachine can be obtained by the following methods [66]:

- *Modified biological cells* - Existing biological cells can be modified in order to let them provide additional functions, such as the emission or reception of specific molecules.
- *Artificial cells* - Another possible approach is to create a completely

artificial cell, designed to perform pre-determined tasks.

Molecular communications can be classified based on how information propagates from transmitter to receiver, that is how the molecules emitted by the transmitting bio-nanomachine reach the receiving bio-nanomachine. Three main propagation mechanisms have been identified [67]:

- *Walkway-based* - The walkway-based mechanism is the least spontaneous among the three. Also known as active transport, this propagation mechanism is carried out by exploiting chemical energy to generate movement, e.g., by using the movement of cilia or rotation of flagella. Alternatively, molecules can move along filaments or tracks, by means of molecular motors.
- *Flow-based* - In this propagation scheme, the transport of information molecules happens through the movement of the fluid mass where these molecules are located. The flow directs the molecules towards the receiver. A perfect example of this is the human circulatory system.
- *Diffusion-based* - This is the most spontaneous propagation mechanism. A transmitting bio-nanomachine emits the information molecules inside the medium, generally a fluid, where they diffuse with a random Brownian motion. In this type of propagation, the molecules flow in the opposite direction of the concentration gradient. MCs which exploits diffusion as propagation mechanism are referred to as Molecular Communication via Diffusion (MCvD).

In the 6G framework, molecular communications (MC) are envisioned as an enabling technology to realize the holistic vision of the human body, in order to integrate it as a node of the network. As introduced in Sec. 1.1, the BnM involved in MCs are expected to replace artificial nano-devices in contexts where the latter are unsuitable due to lack of bio-compatibility and health-related side-effects. The use of BnMs will provide access to a wide range of novel information gathered directly from within the human body, directly related to body signals enabling a wide range of applications, e.g., new diagnostic techniques. They will also enable new forms of interaction with human body, e.g., smart drug delivery.

Given the extreme sensitivity of the data flowing within MCs, security is a major concern. However, current literature on MCs extensively explores

various modulation schemes and channel models for different propagation mechanisms, it offers limited insight into security-related aspects.

This thesis contributes to address this gap in current literature by identifying potential threats in MC scenarios. The problem is approached by mapping the main known security threats in classical telecommunications in a MCs context. Since networking in MCs remains in an embryonic stage of development and upper ISO/OSI layers have not been defined yet, this investigation focuses exclusively on the physical layer with a PLS perspective.

Current literature is mainly focused on MCvD, therefore the presented analysis of security threats is focused on this propagation mechanism. Specifically, the analysis focuses on how these security threats may manifest within the main modulation schemes of MCvD.

Current literature mainly focuses MCvD. Accordingly, the presented analysis of security threats centers on this propagation mechanism. In particular, it examines how such threats may arise within the primary modulation schemes employed in MCvD.

3.2.1 Modulation schemes in MCvD

In a MCvD system two BnMs communicate by emitting molecules which diffuse by following a random Brownian motion. In free diffusion, the molecule flow in the opposite direction of the concentration gradient, *i.e.*, they move from a location with a higher concentration of them to a location with a lower concentration, see Figure 3.4. The molecule flow constitutes a chemical signal [68].

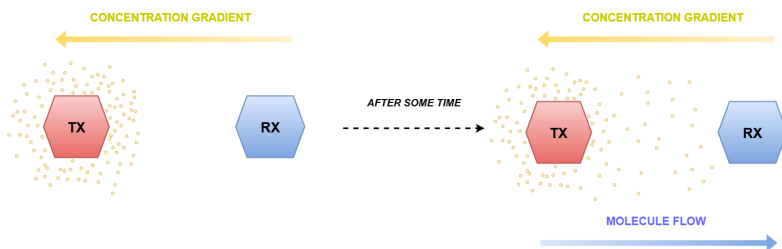


Figure 3.4: Representation of the diffusion process.

To define a communication system it's necessary to define how information are modulated inside the signal and to design the receiver to correctly

demodulate the information from the received signal. In classical telecommunication systems, there are several modulation schemes that can be used to transmit information. Analogously, in MCs the information can be transmitted by varying one or more physical properties of the chemical signal.

The physical properties that can be exploited to modulate the chemical signal can be divided in two main groups: *extensive properties*, such as emission rate, and *intensive properties*, such as molecule type. If the modulation exploits an extensive property, the demodulation phase would require an amount of molecules to correctly demodulate the signal. If the modulation exploits an intensive property, ideally a single molecule could be enough. The main modulation types developed for MCvD are listed below:

Concentration-based modulation. This modulation exploits an extensive property of information molecules: the concentration of the molecules emitted by a BnM. Both analog and digital concentration-based modulation can be considered in MCvD. Concerning analog modulation, amplitude modulation (AM) and frequency modulations (FM) have been considered [69]. Here is considered the digital domain, where concentration-based modulation rely on transmitting information bits through the concentration of released molecules over distinct time slots [68]. Different concentration levels can represent sequences of bits, thus realizing a multi-levels modulation. In [70], a particular M-ary Concentration Shift Keying (M-CSK) scheme is presented, considering multiple transmitters.

Type-based modulations. In this modulation scheme, the transmitting BnM has the ability to release various types of information molecules. The receiving BnM is equipped to detect these molecules. The emitted molecules share similar characteristics, such as size or diffusion rates, but they are uniquely identifiable because each type of molecule can be recognized by a specific receptor located on the surface of the receiving BnM. In [71] a D-MoSK (Depleted Molecular Shift Keying) modulation scheme is presented. For an M-ary MoSK, are required M different types of molecules to modulate information. D-MoSK modulation aims to lowering system complexity by reducing the required types of molecules.

Timing-based modulations. In timing-based molecular communications, the information symbol is encoded inside the emission time of a molecule.

A symbol time T_s is defined and the information symbol is encoded into the molecules emission time t_k , with $t_k < T_s$. The information molecule reaches the receiver at time $t_y = t_k + t_n$, where t_n represents random noise introduced during molecule propagation. In a 1-D environment, the distribution of t_n depends on the type of diffusion: free diffusion, *i.e.*, t_n follows a Lévy distribution, and diffusion with drift, *i.e.*, the distribution of t_n is an inverse Gaussian [72]. In [73] the authors show how high order timing-based modulation can outperform CSK and other modulation schemes.

Spatial modulations. This modulation technique exploits an intensive property of the chemical signal. Multiple spaced molecules emitting sources are involved at the transmitter side, and multiple spaced receptors are involved at the receiver side. The receiver is thus able to calculate the direction of arrival of the molecules and thus the index of the emitting source. The information symbol is encoded in the index of the emitting source. In [74] a Space Shift Keying (SSK) modulation is proposed and inter-symbol and inter-link interference is addressed.

3.2.2 Security Threats Analysis

In this section, the main physical level security threats are mapped into the molecular communications context. In other words, we aim here to detail how a known security attack can be carried out in MCs scenario.

Let's consider a system model composed of two legitimate BnM, Alice and Bob, and a malicious BnM Eve that tries to carry on an attack. The system model is showed in Figure 3.5. The diffusive channel has a strong similarity with a wireless channel: the signal can be overheard by any node in the environment, due to the broadcast nature of the channel. Alice, Bob, and Eve are positioned within the body, maintaining appropriate spatial separation to enable effective molecular communication. It is assumed that the information-carrying molecules used in the described modulation schemes and attack scenarios are entirely distinct from those naturally occurring within the biological environment. The security threats that are here considered are the followings: *jamming*, *eavesdropping*, *MITM*, and *spoofing*.

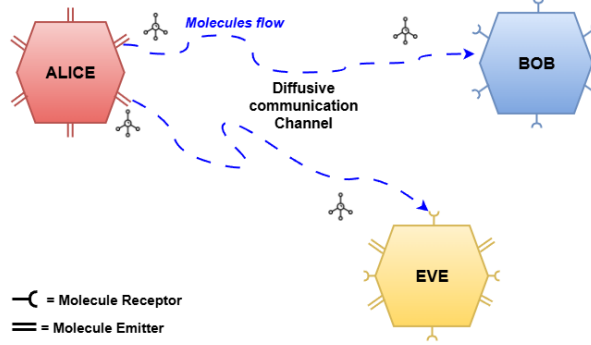


Figure 3.5: General model of a MCvD system.

Jamming

In classical telecommunications, jamming refers to the deliberate interference by Eve to a transmitted signal to prevent Bob from correctly receiving the information signal sent by Alice. This interference can be accomplished by transmitting a noisy signal into the communication channel, thereby reducing the signal-to-noise ratio (SNR) of the legitimate receiver. The frequency band of the jamming signal should include the frequency band of the legitimate signal.

Jamming is a very simple attack to carry out in diffusion-based MCs. For concentration-based, timing-based, and spatial modulations there are two possible techniques to alter (increase or decrease) the concentration of the information molecules (Figure 3.6). The first possibility is that Eve emits a large amount of information molecules in order to saturate the channel. The second possibility is that Eve emits molecules that bond with information molecules so that they form altered molecules that Bob's receptors are not capable to collect. In type-based modulations, the attack can be successful by saturating the channel with only one information molecule so that Bob receives constantly more than one symbol at the same time. More complex is the case of altering the information molecules, since it would require Eve to emit more than one type of molecule that is able to bond with the information molecules. In Figure 3.6 it is shown a jamming attack to a concentration-based modulation. Eve emits molecules that binds to the one emitted by Alice altering them so that Bob's receptor are unable to collect.

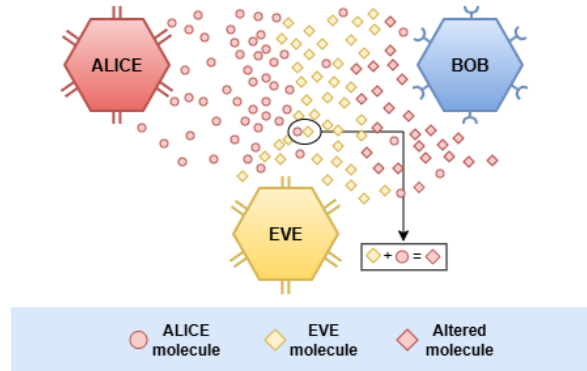


Figure 3.6: Example of jamming attack in a diffusion-based MC system.

The concentration of molecules received by Bob is reduced and corresponds to a different symbol.

In [75] is addressed the problem of a jamming attack in a concentration based modulation. The legitimate transmitter emits type A molecules with an OOK (On-Off Keying) modulation technique. The jammer is able to emit type A or B molecules to respectively increase concentration or decrease it (type B molecules react with type A molecules to form type C molecules). The authors also present a coding scheme capable of improving jamming-resistance of the system.

Eavesdropping

This type of attack is a direct consequence of the broadcast nature of the channel: an unintended user (the eavesdropper) could potentially receive the signal and try to extract the bits from the intercepted signal. This is a passive attack, as all happens without the legitimate nodes being aware of it.

In all the described diffusion-based modulation schemes Eve can eavesdrop the information signal sent by Alice to Bob, by simply standing close enough to receive the signal correctly and being equipped with receptors capable of collect information molecules.

Assuming that Eve's receptors absorb the information molecules during the eavesdropping, it would alter the concentration of the information

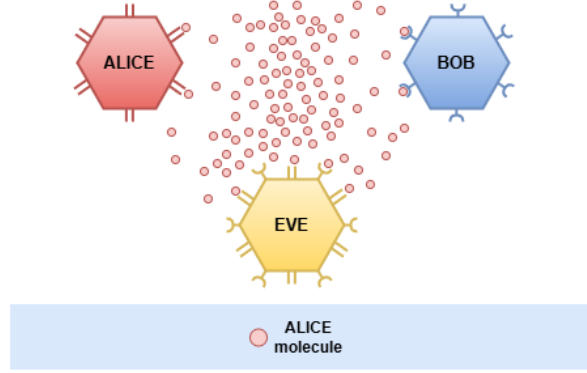


Figure 3.7: Example of eavesdropping attack in a diffusion-based MC system.

molecules, and Bob can detect the attack attempt. In order to avoid this, Eve must reintroduce all the absorbed molecules in the environment so that eavesdropping is a completely passive attack also in diffusion-based MCs.

In Figure 3.7 it is shown an example of eavesdropping attack to a concentration-based modulation. Alice emits molecules with a concentration corresponding to a specific symbol. Eve stands near to the channel and receives part of the signal emitted by Alice and demodulate it by estimating the concentration.

Authors in [65] analyze secrecy capacity for a MCvD system in an eavesdropping scenario.

Man in the Middle

In classical telecommunication systems, the Man in the Middle (MITM) attack is based on eavesdropping the signal transmitted by Alice and then sending a modified information signal to Bob.

The malicious node does not just intercept the data sent by Alice, but it also modifies the data and sends it to the intended receiver (Bob) without the receiver being aware of it. The legitimate receiver gathers corrupted information, mistakenly believing it to be legitimate. In MCs this attack can be accomplished by changing (increase or decrease) the concentration of information molecules received by Bob. The two techniques exploited to alter the concentration of information molecules are the same described for the jamming attack, but here the aim is to change the information received

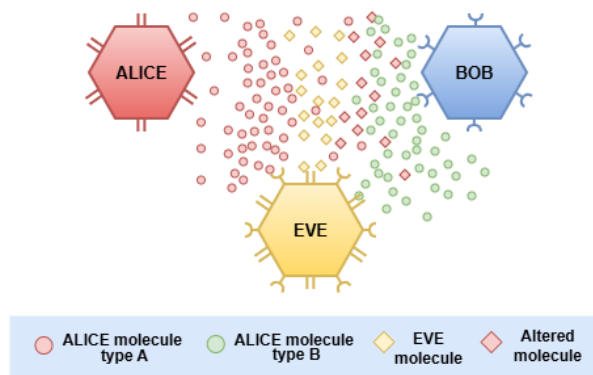


Figure 3.8: Example of man in the middle attack in diffusion-based MC systems.

by Bob, not just to destroy it.

In concentration-based modulations, to perform a MITM attack, Eve increases or decreases the molecules concentration according to the symbol it wants to send to Bob. In type-based modulations, a MITM attack can be carried out in two steps: first Eve destroys the information signal by altering the information molecules emitted by Alice; then Eve sends its information molecules to Bob. In timing-based modulations, a MITM attack involves changing the molecule concentration in order to change the emission time t_k estimated by Bob. Another possibility is that Eve destroys the information molecules sent by Alice and replaces it with molecules emitted at a different time t_k .

In spatial-based modulations, MITM attack is more difficult to carry out. Eve must first destroy the information molecules emitted by Alice and then send its own to Bob. In addition, Eve must have the same emitting sources as Alice and associate the same index to them, so they emit the same symbols.

In Figure 3.8 it is shown an MITM attack to a simple type-based modulation. Alice emits two types of molecules (A and B). If Alice is emitting molecule A, Eve could alter them emitting molecule of type B at the same time. Thus Bob receives type B molecule instead the intended ones.

To date, the MITM security threat in MCs systems has not yet been properly treated in the current literature.

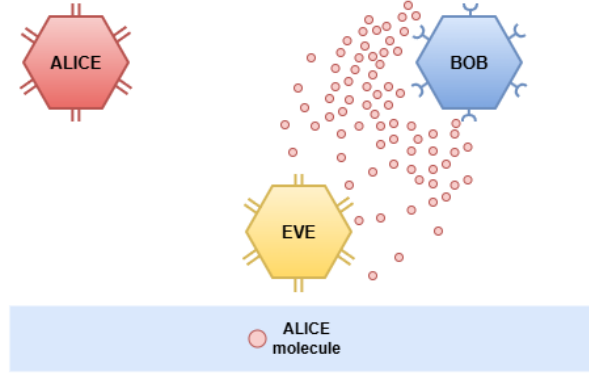


Figure 3.9: Example of spoofing attack in a diffusion-based MC system.

Spoofing

Although in classical telecommunication systems the spoofing attack does not directly involve the physical layer, it is nonetheless included in this analysis, as it does play a role at the physical layer in Molecular Communication (MC) systems.

In classical telecommunication systems, a spoofing attack happens when the attacker manages to falsify its identity, pretending to be a legitimate node. If the attack is successful, the victim authenticates the attacker as a legitimate node and thus could send it sensitive data, or receive false data from it.

This is one of the most dangerous and easiest attacks that can target an MCvD system. Since higher layers are not studied yet in MCs, the authentication process can be currently applied only at physical-layer [76].

In concentration-based, type-based and timing-based modulation techniques is sufficient that the attacker is able to emit the legitimate molecules knowing the used modulation technique. In spatial modulation techniques, requires also that Eve have the same emitting source that Alice and associate the same index to them.

In Figure 3.9 is shown a simple spoofing attack in concentration-based modulations where Eve emits the same molecules as Alice, which is not transmitting. Bob receives molecules sent by Eve, believing they are sent by Alice.

In the current literature, the problem of authentication in MCs has been addressed by very few works. In [76] a physical layer security authentication approach for diffusion-based MCs system is described. The channel impulse response (CIR) between the two legitimate bio-nanomachines is exploited to authenticate the transmitting node, where the CIR is defined as the concentration of molecules at a specific distance and time. This means that at the moment, there is almost no mechanism that allows Bob to authenticate the bio-nanomachine that is communicating with him (Figure 3.9).

3.2.3 Complexity Analysis of the Threats

In this section, the level of complexity required for EVE in order to carry on each type of attack will be analyzed, depending on the various attacks described in Sec. 3.2.2 and the types of modulation shown in Sec. 3.2.1.

The attack complexity through four *complexity parameters*, which consider aspects related to the complexity of the bio-nanomachine required to perform an attack. The four complexity parameters are:

1. *The number of different types of molecules:* Eve is required to emit or receive more than one type of molecule.
2. *Both transmitter and receiver are needed:* Eve is required to have the capability of both transmitting and receiving the molecules involved.
3. *Need for synchronization:* Eve needs to have the same timing of Alice and Bob, in order to be synchronized and thus to receive or send the symbols correctly.
4. *Capability of estimating the index of the emitting source:* Eve must have multiple receptors and must know the indices of the emitting sources to which the various symbols are associated.

Without loss of generality, in this analysis a binary signaling is considered, where only two symbols are transmitted. In particular this binary signaling is implemented in the different modulation schemes described in Sec. 3.2.1 as follows:

- **Concentration-based modulations.** The two symbols are mapped on two different concentration levels of the information molecule emitted by Alice. Bob must have receptors capable of receiving the molecule

and must be capable of estimating its concentration over a defined time slot.

- **Type-based modulations.** The two symbols are mapped on two different information molecules emitted by Alice. Bob must have receptors capable of receiving both molecules.
- **Timing-based modulations.** The two symbols are mapped in two different emission times t_{k1} and t_{k2} inside a defined symbol time T_s . Alice and Bob must be synchronized on the same time basis. Also Bob must have the capability of estimating t_{k1} and t_{k2}
- **Spatial modulations.** Alice is composed of two emitting sources. The index of the emitting source maps the transmitted symbol. Bob must be capable of estimating the index of the source that emitted the molecules and thus demodulate the transmitted symbol.

Table 3.2: Complexity level of each PLS attack for each MC modulation scheme.

PLS threat	Modulation schemes			
	Concentration based	Type based	Timing based	Spatial
Jamming	Very Low	Very Low	Very Low	Very Low
Eavesdropping	Low	High	High	High
MITM	High	High	Very High	Very High
Spoofing	Very Low	Low	Low	Very Low

For each attack type described in Sec.,3.2.2 and each modulation scheme outlined in Sec.,3.2.1, an evaluation has been conducted to determine whether the previously defined complexity parameters must be satisfied by the attacker.

A *Very Low* complexity value has been assigned if none of the complexity parameters are met, while a *Very High* complexity value has been assigned if at least three of the complexity parameters are met. The results of the analysis are reported in Table 3.2.

The discussion of the complexity analysis of the different attacks follows.

Jamming. Jamming is a very simple attack to be executed in MC diffusion-based channels. None of the four defined complexity parameters are verified for concentration-based, type-based, timing-based, and spatial modulations.

Eavesdropping. Eavesdropping attack shows a low complexity for concentration based modulation and a high complexity for type-based, timing-based and spatial modulations. The ability of both transmitting and receiving is required in all the modulation schemes because, assuming that Eve has absorbing receptors, all the absorbed molecules must be reintroduced in the environment to ensure that the attack is not detected. In a binary type-based modulation Eve is required to being able to both receive and transmit two types of molecules. In a binary timing based modulation to perform eavesdropping Eve is required to estimate T_s , t_{k1} and t_{k2} and also to share the same time base as Alice and Bob in order to synchronize correctly. In spatial modulation Eve is required to be capable of estimating the index of the emitting source to demodulate the symbol.

Man In The Middle. MITM is the most complex attack to carry out. In particular, concentration-based and type-based modulations show a high complexity, while timing-based and spatial modulations show a very high complexity. In all the considered modulation schemes, to effectively change the symbol emitted by Alice, Eve must be capable of emitting molecules that bond to Alice's molecules and alter them so that Bob's receptors are unable to collect them. In every modulation scheme, Eve must be capable of emitting at least two types of molecules. To carry on a MITM attack in a timing-based modulation, Eve is required to be synchronized with Alice and Bob in order to change the transmitted symbol. Eve must not only be capable of both transmitting and receiving the information molecule but also emitting the molecule designed to alter the information molecule in order to eliminate them. In a spatial modulation, Eve must be capable of estimating the index of Alice's emitting source. In addition, Eve should have multiple emitting sources with the same indexes to be able to both transmit and receive. In a concentration-based modulation, Eve has to be able to transmit and receive the information molecule and also to emit a molecule to alter the information molecule when it needs to lower the concentration to change the transmitted symbol. In a type-based modulation, Eve is required to be able to receive and transmit both information molecules. In addition,

Eve is required to emit molecules designed to alter both types of information molecules so that Bob’s receptors are unable to collect them.

Spoofing. Carrying out a spoofing attack in a MC diffusion-based channel has a low complexity degree. In concentration-based and spatial modulations, none of the defined complexity parameters is met. In a type-based modulation, Eve is required to be able to transmit the two types of information molecules. In a timing-based modulation, Eve is required to have the same time base of Bob so that Bob can demodulate the symbols correctly. Moreover, Eve must know the emission times t_{k1} and t_{k2} associated with the symbols.

Looking at the results presented in Table 3.2, we observe that the modulations that require the highest complexity of attacks are the timing-based and the spatial modulation. Anyway, these two modulation schemes are also the most complex to implement.

3.2.4 Conclusion

In the presented work, a comprehensive analysis of security threats in diffusion-based molecular communication (MC) systems has been conducted. Despite the promising potential of MCs, particularly in biomedical applications and as key technology for sixth-generation (6G) networks, their security vulnerabilities have received limited attention in existing literature. To address this gap, a systematic mapping of well-established physical-layer attacks, including jamming, eavesdropping, man-in-the-middle (MITM), and spoofing, has been conducted, illustrating how these threats may manifest within molecular environments.

The complexity analysis indicates that certain attacks, such as jamming and spoofing, are relatively simple to execute due to the broadcast nature of MCvD channel. In contrast, threats like MITM and eavesdropping require more advanced capabilities, including precise synchronization and molecular type differentiation. Moreover, it has been observed that specific modulation schemes inherently offer greater resilience against particular categories of attacks.

This threat mapping serves not only to enhance the security of MC systems but also contributes to the broader objective of securing future 6G infrastructures. MCs are an enabling technology for the holistic vision of human body as node of the 6G connective fabric. However, such perva-

sive connectivity must be accompanied by robust security mechanisms. By identifying vulnerabilities at the physical layer, the proposed framework supports the development of secure-by-design MCs architectures, ensuring that future bio-integrated networks provide both comprehensive physiological insight and reliable protection of sensitive biomedical data.

3.3 Confidentiality in VLC: Secure Area analysis for VLC-based Implantable Medical Devices (IMD)

In recent decades, there has been rapid progress in the development of implantable medical device (IMD) technology, driven by advancements in various scientific and engineering fields such as nano-materials, biotechnology, and microelectronics [77]. The field of IMDs has primarily been explored in the medical domain: numerous works emphasize the significance of IMDs for monitoring and treating patients. This demand has given rise to an innovative technological branch focused on creating sensors, actuators, and devices suitable for implantable applications.

Traditional IMDs operate within the RF band, which is often saturated in terms of bandwidth (BW) [78]. These devices typically use transmission power in the range of several tens of milliwatts [79]. Consequently, they face challenges in achieving the high data rates required to emulate human organs effectively [80, 81]. Moreover, conventional IMDs are susceptible to interference from other sources operating within the same RF band, leading to significant performance degradation.

OWC are a promising transmission technology capable of addressing these problems. Recent literature has explored the use of OWC for both in- to out-body communications, with studies characterizing tissue penetration of optical electromagnetic waves [82–84]. In particular, the usage of VLC (Visible Light Communications) is supported. VLC are based on the transmission light waves belonging to the spectrum of frequencies visible to the human eye (see Figure 3.10), with wavelengths between 380 *nm* (violet) and 750 *nm* (red).

One of the main enabling factors of VLC has been the widespread adoption of LEDs. LEDs emit light waves in the visible spectrum that can be modulated at much higher frequencies than traditional light bulbs, which are limited by thermal inertia and unable to support high modulation rates. The widespread adoption of LEDs for lighting purposes means that VLC can exploit an existing infrastructure: it would be sufficient to add a transmitter to the already deployed LED lamps, in order to modulate the emitting light. This is true also for IMDs, where LEDs are already installed and exploited to localize the device and to provide a visual feedback on the status of the communication or of the device.

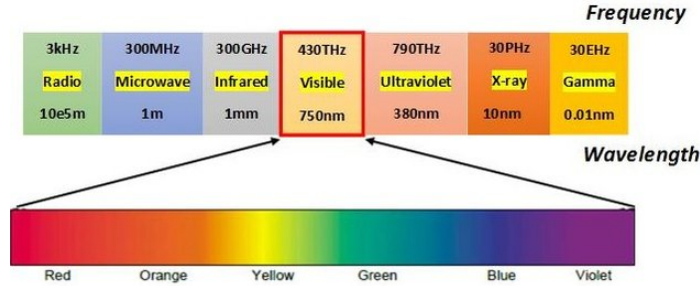


Figure 3.10: Visible Light Communication frequency spectrum.

VLC shows low absorption coefficient, which enables optical signals to penetrate human tissue effectively, making it an ideal choice for in-to-out body communications (see Figure 3.12a). Furthermore, VLC addresses the critical security concerns associated with IMD applications by ensuring enhanced data privacy and protection. Given the sensitivity and personal nature of the data transmitted within the body, VLC emerges as the preferred solution over RF technologies, as light waves are inherently confined by the environment and exhibit strong directionality. In fact, the signal can be effectively shielded by simply covering the area containing the implantable device, *e.g.*, with clothing or a bandage. This level of physical isolation is not achievable with RF antennas, whose transmitted signals readily penetrate such materials.

A compelling application field for VLC-based IMDs in the context of WBAN is that they can be used to relay in-body communication to an on-body node contributing to realize the holistic vision needed to model the human being as a node of the 6G network. Let's consider a scenario where a subcutaneous IMD is used to collect data from sensors or an other communication system, *e.g.*, molecular communication, RF or GC), inside the body. This device can transmit these signals to an on-body device located near the skin surface. The on-body device, in turn, relays this signal to the hub of the WBAN, ensuring a seamless flow of information from the inside of the body to the external networks(see Figure 3.11).

This thesis presents a MATLAB-based simulation to demonstrate the feasibility and advantages of Visible Light Communication (VLC) for in-to-out body transmission involving implantable subcutaneous devices.

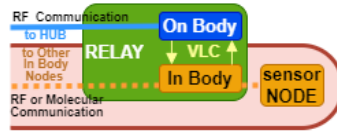


Figure 3.11: VLC Implantable Use Case

A model of light propagation through the skin layers (in-body channel) as well as a model of light propagation from the outside skin on (out-body channel) is provided. The communication source (LED) is under the skin, while the receiver (photo-detector) is outside the body. Both single LED single photo-detector (SISO) and multiple LEDs multiple photo-detectors (MIMO) communication is considered. The performance of SISO and MIMO LED-based VLC systems is characterized by means of the bit error rate (BER) as a function of the distance and angle of the receiver from the skin. Therefore, a feasible approach to Near-Field Communication (NFC) for VLC systems within IMD scenarios has been defined.

Since VLCs exhibit intrinsic security properties due to their low absorption coefficient, the system's security performance is analyzed in terms of secure area (A_{sec}), a PLS metric used to measure the region of space in which communication between a transmitter and a receiver remain confidential, in relation to the transmitter power, receiver position and channel characteristics [85].

3.3.1 System Model

Let's consider an NFC scenario in which the LED transmitter is implanted 2.17 cm beneath the skin surface, while the receiving photodiode is positioned at a maximum distance of 5 cm above the skin. In the described context, it's not feasible to approximate the light spot on the skin created by the implanted LED, as a point emission source because the distance between the receiver and the skin is comparable to the dimensions of the spot. Hence, to address this issue, a 50 mm square section of skin was subdivided into smaller square areas measuring 0.1 mm each. Each of these small skin areas was individually analyzed as a micro-LED (μ -LED) with its pattern and its intensity of emission.

The VLC system was analyzed in terms of four components for each route

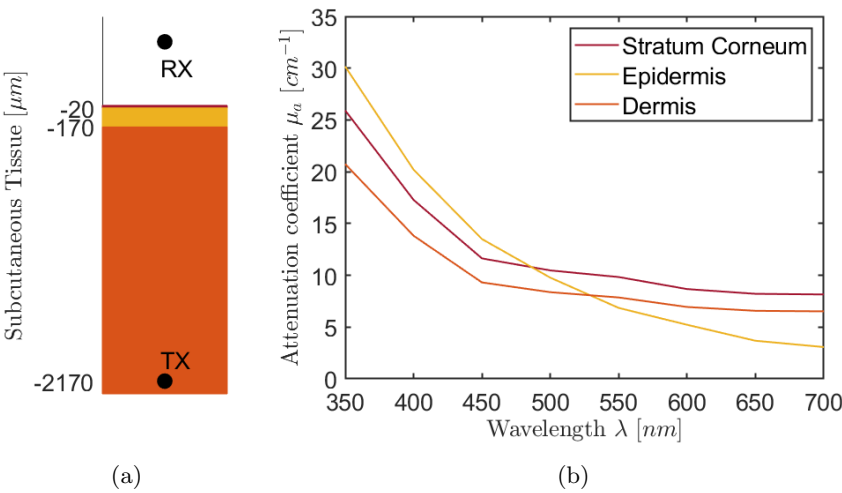


Figure 3.12: Human skin absorption for wavelength at different skin layers: 1. Stratum Corneum; 2. Epidermis; 3. Dermis.

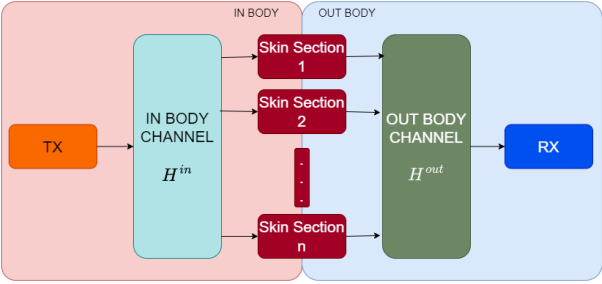


Figure 3.13: VLC System Model

from transmitter to receiver: two corresponding to the telecommunication devices, *i.e.*, the transmitting implanted LED and the receiving photodiode, and two associated with the optical channel, *i.e.*, one for characterizing light diffusion inside the body and one for external diffusion. The system model used for simulation is shown in Figure 3.13.

In-Body Channel Model

The evaluation of the in-body channel's characteristics was derived from a review of the literature, as discussed in [86]. This review includes considerations of both light power absorption (as depicted in Figure 3.12b) and scattering. Absorption and scattering collectively contribute to the formation of a light spot on the skin.

To define the impact of absorption and scattering, Stochastic Monte Carlo (MoCa) modeling is the most used solution in the literature. In fact, over the past decade, several MoCa algorithms for simulating the propagation of optical radiation in biological media have been developed for diverse applications [86, 87].

The simulation is based on modeling a large number of possible trajectories of the light from the source to the skin surface. Intensity path loss has been modeled using the Beer-Lambert law, which is represented as:

$$I_i = I_{i-1} e^{-\mu_a l} \quad (3.1)$$

where I_i is the light intensity during the i -th step of simulation, μ_a is the attenuation coefficient and l is the distance covered by the photon.

The scattering event is simulated by generating a random propagation angle denoted as θ , which determines the new direction of the photon. This angle follows a Henyey-Greenstein scattering phase function [88]:

$$p(\theta) = \frac{1}{4\pi} \frac{1 - a_f^2}{(1 + a_f^2 - 2a_f \cos \theta)^{3/2}} \quad (3.2)$$

Here, the parameter a_f represents the anisotropy factor, which is essentially the average cosine of the scattering angle θ . The value of g falls in the range $[-1, 1]$: when a_f is less than 0, it indicates back-forward scattering; when a_f equals 0, it means isotropic (Rayleigh) scattering, and when a_f is greater than 1, it denotes forward scattering (Mie scattering).

The simulation has been depicted in two dimensions because the LED's emission is axisymmetric. Consequently, the transfer function ($\mathbf{H}^{in}$) has

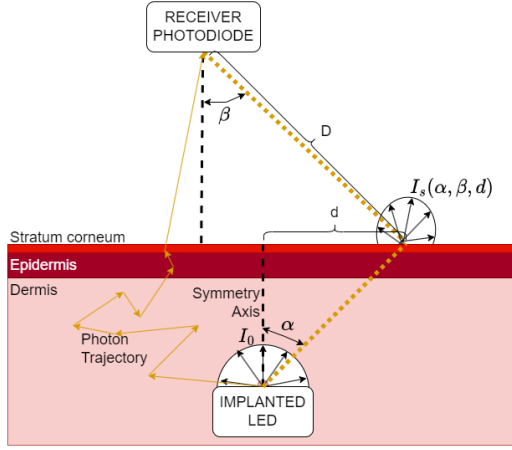


Figure 3.14: Simulated Model

been established in relation to the distance between a skin section and the symmetry axis ($\mathbf{d} = [d_1, \dots, d_N]$ with N is the number of skin section). In this simulation, the LED's emission has been presumed to exhibit uniformity concerning the angle, then the transfer function has been defined as:

$$\mathbf{H}^{in} = I_0 \mathbf{f}(\mathbf{d}) \quad (3.3)$$

where I_0 is the uniform intensity of transmitter and $\mathbf{f}(\mathbf{d})$ is the path loss that depends on light absorption and scattering (see Figure 3.14).

Out-Body Channel Model

The MoCa simulation is also used to define the emission pattern $\mathbf{g}(\alpha, \beta)$ of each small skin area that depends on the angle formed between the symmetry axis of transmission and the line joining the transmitter to the respective small emission area (α) and the angle formed between the perpendicular to the skin surface and the line joining the respective small emission area to the receiver (β). The photodiode's area, utilized as the receiver, has been intentionally kept small since it's designed for short-range communications. Consequently, the distance involved is not on the same scale as the photodiode's surface area, enabling an approximation as a point receiver.

The photodiode has been positioned in a grid arrangement oriented perpendicular to the skin. This grid is centered at the center of the light spot, with the x-axis aligned parallel to the skin and z-axis aligned perpendicular to the skin. Along the x-axis, the grid points align with the segmented areas of the 50mm skin region, while along the z-axis ranges from 5mm to 50 mm, increasing in 5mm intervals.

For each receiver position, the out-body channel path loss ($\mathbf{H}^{out}$) between specific receiver position and each small emission areas has been calculated as follows:

$$\mathbf{H}^{out} = \frac{g(\alpha, \beta) \cos \beta}{D^2} \quad (3.4)$$

where D represents the distance from a given receiver position to each of the individual small emission areas (see Figure 3.14), the cosine of β depends on the characteristic of the geometric characteristics of the system.

VLC System Model: Single Led

The prevalent modulation scheme employed in VLC applications is Amplitude Sub-carrier BPSK as suggested in standard IEEE 802.15.7 [89], chosen to maintain the LED's primary function of providing constant light signaling for human eye. For this reason, the data transmitted $s_{tx}(t)$ can assume value of 1 or -1 in base-band. We assume here a binary phase shift keying (BPSK) modulation.

The received data $s_{rx}(t)$ in base-band is characterized with the following equation, because the amplitude sub-carrier frequency is in the led modulation band, *i.e.*, the out-body channel is characterized only on path loss component:

$$s_{rx}(t) = \left(\underbrace{\sum \mathbf{H}^{in} * \mathbf{H}^{out}}_{H_{tot}} \right) s_{tx}(t) + n(t) \quad (3.5)$$

where $n(t)$ is the complex noise of Additive White Gaussian Noise (AWGN) for VLC link in amplitude sub-carrier frequency used in modulation scheme. The operator $(*)$ is a multiplication corresponding elements between two array and $\sum$ is a sum of all element of the array. The characterization of out-body communication channel for NFC application was defined as $H_{tot} = \sum \mathbf{H}^{in} * \mathbf{H}^{out}$ (see Equation (3.5)).

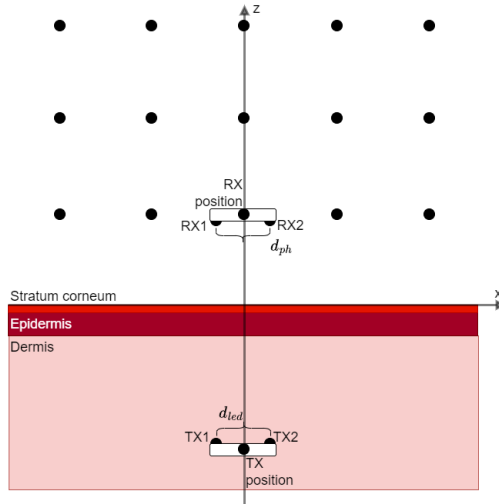


Figure 3.15: MIMO System Model

The data receiver in base-band was demodulated with the sign of real part of signal in BPSK modulation.

VLC System Model: MIMO

The MIMO used for simulation is a 2×2 , *i.e.*, the system is composed by two transmitting implanted LEDs and two photodiodes as receiver out of the body.

The mathematical treatment for MIMO application is similar to the single LED emission described above. Also in this case, as reference portion skin of 50 mm square section was subdivided into smaller square areas measuring 0.1 mm each and analyzed as a μ -LED matrix. The geometrical differences between a two system model are the position of led and photodiode in reference portion of skin and in position of receiver respectively. In the previous case, the position of led is in the center of reference portion skin, while in MIMO system model the middle point between two led is positioning in the middle of 50 mm square section of skin, *i.e.* the two leds are located at $d_{led}/2$ from the center reference portion skin along x-axis one to the right and the other to the left respect of the center (see Figure 3.15).

The same applies for position photodiode, in the previous case the po-

sition of photodiode coincided with the position of receiver, while the two photodiodes are located at $d_{ph}/2$ from receiver position along x-axis as LEDs position (see Figure 3.15).

For each transmitter, the transfer functions of in-body channel ($\mathbf{H}_1^{in}$ and $\mathbf{H}_2^{in}$) has been calculated, similar to the single LED case (as shown in equation (3.3)). The transfer functions for the out-body channel, as per formulation (3.4), are contingent on the angles α and β , and consequently, they rely on the positioning of both the transmitter and receiver. For each receiver position, four distinct transfer functions denoted as $\mathbf{H}_{i,j}^{out}$ (with $i, j = 1$ or 2) have been defined, where i and j represent the specific LED and photodiode that transmitting and receiving data, respectively.

The two LEDs transmit distinct data streams ($s_{1tx}(t)$ and $s_{2tx}(t)$). The data received from the two photodiodes ($s_{1rx}(t)$ and $s_{2rx}(t)$) at each grid position in the base-band has been characterized using linear equations, as follows:

$$\underbrace{\begin{bmatrix} s_{1rx}(t) \\ s_{2rx}(t) \end{bmatrix}}_{\mathbf{s}_{rx}(t)} = \underbrace{\begin{bmatrix} \sum \mathbf{H}_1^{in} * \mathbf{H}_{1,1}^{out} & \sum \mathbf{H}_2^{in} * \mathbf{H}_{2,1}^{out} \\ \sum \mathbf{H}_1^{in} * \mathbf{H}_{1,2}^{out} & \sum \mathbf{H}_2^{in} * \mathbf{H}_{2,2}^{out} \end{bmatrix}}_{\mathbf{H}_{tot}} \underbrace{\begin{bmatrix} s_{1tx}(t) \\ s_{2tx}(t) \end{bmatrix}}_{\mathbf{s}_{tx}(t)} + \underbrace{\begin{bmatrix} n1(t) \\ n2(t) \end{bmatrix}}_{\mathbf{N}(t)} \quad (3.6)$$

where $\mathbf{N}(t)$ is a vector of two noises ($n1(t)$ and $n2(t)$) to characterize two different additive Gaussian noise channels.

To demodulate the two different data in the simulation, the inverted function of Eq. (3.6) has been calculated as:

$$\mathbf{s}_{tx}^*(t) = \mathbf{H}_{tot}^{-1} \mathbf{s}_{rx}(t) \quad (3.7)$$

where $\mathbf{s}_{tx}^*(t)$ are the estimation of transmitted signals. Once the two transmitted signals have been separated, the received data in the base-band is demodulated using the sign of the real part of the signal in the BPSK modulation.

Secrecy Capacity and Secure Area

In order to provide foundation for the application of more advanced PLS confidentiality techniques, in this work security of VLC for IMDs is assessed. In particular, security performances of the system will be evaluated in terms

of *secure area*. The secure area is a PLS metric used in current literature to asses security of a telecommunication system [85, 90].

Without loss of generality, let's consider a discrete two-dimensional square area $\mathbf{A}$, as in the present work. The area $\mathbf{A}$ consists of $p \times p$ points, where p denotes the number of point of each side of $\mathbf{A}$. Let's also consider two legitimate devices Alice and Bob and an eavesdropper Eve. Alice transmit a message M to Bob, which tries to recover it from the received signal. Given, a fixed position of Alice and Bob, the channel capacity C_B of their link is expressed as:

$$C_B = B \log_2(1 + \frac{S_B}{N}), \quad (3.8)$$

where B is bandwidth of the channel, S_B is the power of the signal received by Bob and N is the power of the AWGN noise. The channel capacity C_E of the link between Alice and Eve is expressed as:

$$C_E = B \log_2(1 + \frac{S_E}{N}), \quad (3.9)$$

where S_E is the power of the signal received by Eve.

Given a fixed position of Eve, the secrecy capacity C_s , *i.e.*, the maximum data rate at which Eve is unable to decode any information. For simplicity, in this work it has been expressed as:

$$C_S = C_B - C_E, \quad (3.10)$$

In the worst case scenario the position of Eve is unknown, *i.e.*, it could be located in any point of the area $\mathbf{A}$ with equal probability. Given a fixed position of Alice and Bob and considering every possible position for Eve, the secure area A_{sec} is defined as the number of point of $\mathbf{A}$ where $C_S > 0$. It can be computed by defining a matrix $\mathbf{S}$, of dimensions $p \times p$, where each point corresponds to a Eve's position. To each point is assigned a value 1 if $C_S > 0$, 0 otherwise. Therefore A_{sec} can be expressed as:

$$A_{\text{sec}} = \sum_{r=1}^p \sum_{k=1}^p \mathbf{S}(r, k) \quad (3.11)$$

3.3.2 Results

This section has been divided in two parts: initially, through MoCa simulation, the functions $f(d)$ and $g(\alpha, \beta)$ were computed. Subsequently, these

Skin Parameter	Stratum Corneum	Epidermis	Dermis
Absorption coefficient (μ_a) [cm^{-1}]	8.15	3.07	6.52
Anisotropy Factor (a_f)	0.942	0.804	0.715
Layer Thickness [μm]	20	150	2000

Table 3.3: Skin Parameters used in the MC simulation

computed functions were utilized to simulate a potential transmission scenario between a single LED and a single photodiode, as well as for MIMO 2×2 transmission setups. The communication simulations were employed to calculate the Bit Error Rate (BER), which served as an assessment metric for the performance of the two prospective communication systems, and the channel capacity needed to compute the secure area (A_{sec}), in order to assess security performance of the system. The skin coefficients used in the simulation were taken from literature [91], and it has been summarized in Table 3.3.

Channel Characterization

Through MC simulation, the light intensity and direction have been established, as illustrated in Figure 3.16. Notably, the summation of light beam intensities emanating from the skin surface, in relation to the distance from the LED's symmetry axis, is depicted in Figure 3.16a. These data points have been effectively approximated with a Gaussian distribution, as indicated by the red line in Figure 3.16a. This distribution is used to define Transfer Function $\mathbf{H}^{in}$.

For each skin section used to the defined $\mathbf{H}^{in}$, the summation of light beam intensity in angle clusters has been used to define the pattern of emission of single skin section, $g(\alpha, \beta)$. The average of the distribution of light intensity in angle cluster highlights the refraction effect of the light. In fact, in Figure 3.16b the data has been fitted with a linear function with a different angular coefficient compared to α that characterized a single skin section. This angular coefficient defined similar to refractive index, n_{refr} .

The results obtained have been utilized to scale the emission pattern of a single skin section, incorporating all simulation data in the evaluation of

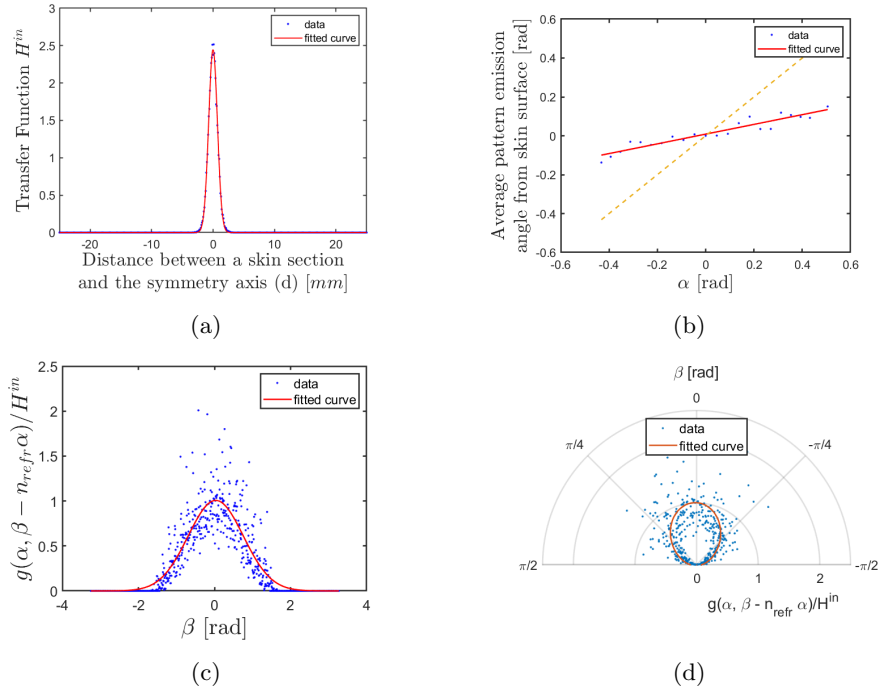


Figure 3.16: Transfer Function of In-body Channel: in (a) the distribution of light intensity along the skin surface; in (b) the refraction effect. Pattern emission of skin scaled to intensity and average direction: (c) Cartesian plot (d) Polar plot.

the emission pattern caused by scattering. The data has been modeled using a Gaussian distribution, as demonstrated in Figure 3.16c. Figure 3.16d presents the emission pattern of a single skin section. The peak of the Gaussian distribution is equal to 1 since it's scaled against $\mathbf{H}^{in}$, and it is centered at an angle of 0 rad, which is achieved by translating the angle by $n_{\text{refr}}\alpha$.

System Performance: Bit error rate

In this section the BER is reported as a function of the position of the transmitters (LEDs) and receivers (photodiodes). The signal to noise ratio (SNR) at the receiver is dependent only on the path loss as depicted in (3.5). In particular, the received power depends on the in-body and out-body path loss as depicted in (3.3) and (3.4), respectively. The path loss depends on the position of the receiver, given the location of the transmitter, as depicted in Fig. 3.14. The performance of the simulated implanted VLC system is assessed by analyzing the BER. A total of 1×10^5 symbols with Manchester coding were transmitted from the transmitter to the receiver.

In the case of single LED communication, the BER remains below 1×10^{-5} up to a distance of 35 mm, as illustrated in Figure 3.17a.

In the MIMO case, the alignment between LEDs and photodiodes becomes more critical. In this simulation, both transmitters and receivers are positioned along the same axis, which represents the optimal configuration. Moreover, the system is not axisymmetric, in contrast to the previous case.

Consequently, the position of the receiver along the x-axis (as shown in Figure 3.17b) differs from its position along the y-axis (as depicted in Figure 3.17c). In both cases, the BER remains below 1×10^{-5} up to a distance of 30 mm. However, the width of the maximum BER region is narrower compared to the previous case when the position of the receiver is in the plane (xz) including the segment joining the two transmitters/receivers (as shown in Figure 3.17b), while the maximum BER region is comparable to the previous case when the position of the receiver is in the plane (yz) orthogonal to the segment joining the two transmitters/receivers (as shown in Figure 3.17c).

It is worth noting that the maximum BER limit is 30 mm for the MIMO case, while it is 35 mm for the single LED case. This is due to the self-interference of the LEDs in the MIMO case.

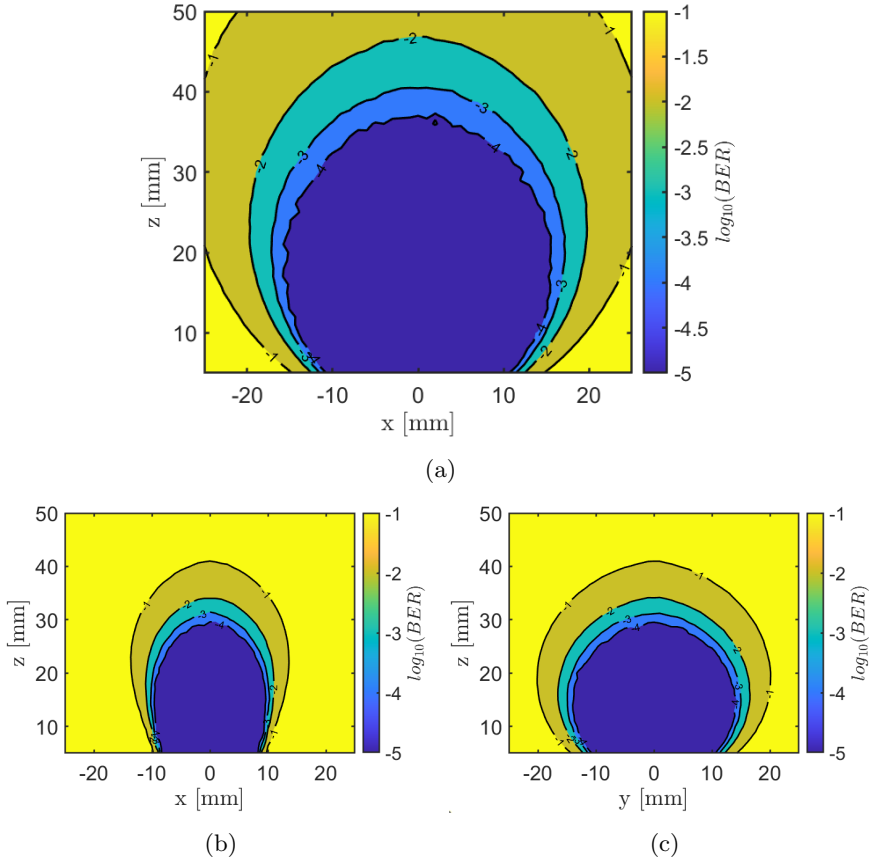


Figure 3.17: Bit error rate (BER) over the vertical plane for the single LED (a) and multi-LED (b)–(c) systems.

System Performance: Secure Area

In the same configuration of the BER analysis conducted in the previous section, security performance of the system have been evaluated analyzing the evolution of the secure area A_{sec} . In particular, considering a fixed position for the LEDs, the position of the receiver has been varied along the vertical and horizontal axes.

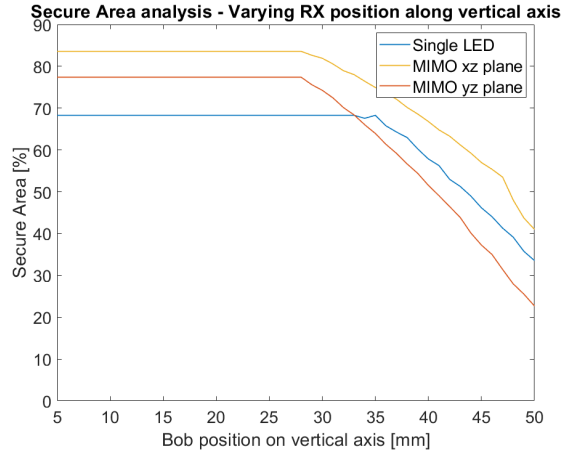
In the single LED case, only the position of the receiver on the xz vertical plane was considered, as positioning it on the yz plane yields equivalent results. Therefore, z and x are the vertical and horizontal axis, respectively. Two cases were considered. First, with the receiver position fixed at $x = 0$ mm along the horizontal axis, its position along the vertical axis was varied within the range $z = [5, 50]$ mm. For each position of the receiver, A_{sec} was computed using (3.11). Subsequently, the receiver was fixed at $z = 35$ mm, while its horizontal position x was swept in the range $x = [-25, 25]$ mm. Also here, A_{sec} was computed for each position of the receiver.

In the MIMO case, both xz and yz vertical planes were separately considered. For both planes the same kind of tests described for the single LED case were conducted. In particular, with the vertical position fixed at $z = 35$ mm, the receiver position along the horizontal axes was varied within the ranges $x = [-25, 25]$ mm and $y = [-25, 25]$ mm when considering the xz and yz planes, respectively. In the other case, the position of the receiver on the horizontal axis was fixed to $x = 0$ mm when considering the xz plane and to $y = 0$ mm when considering the yz plane. In both cases, the position of the receiver on the vertical axis z was varied in the range $z = [5, 50]$ mm.

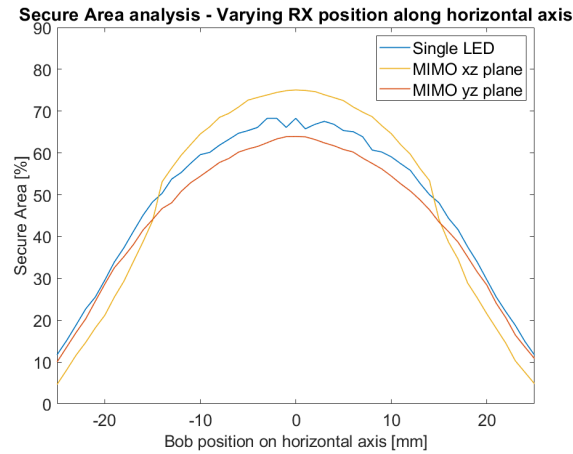
The results of the analysis are showed in Figure 3.18. In particular, Figure 3.18a shows the security performance while varying the RX position along the vertical axis on xz planes for single LED case and on xz and yz planes for MIMO case. Similarly, Figure 3.18b shows the security performance while varying the RX position along the horizontal axis, for the same cases.

In Figure 3.18a, the single LED case exhibits the worst security performance in terms of secure area. This outcome is directly linked to the higher reliability of the configuration, which ensures lower BER values at greater distances from the transmitter, consistently with the results shown in Figure 3.17a.

Conversely, the MIMO configuration demonstrates a pronounced sensitivity to receiver placement. Specifically, positioning the receiver on the xz plane yields superior security performance compared to placement on the yz



(a)



(b)

Figure 3.18: Evolution of the Secure Area A_{sec} moving RX position along the (a) vertical axis (b) horizontal axis.

plane, as corroborated by Figures 3.17b and 3.17c.

Figure 3.18b further reinforces the advantage of the MIMO configuration with the receiver located on the xz plane, which consistently delivers the best security performance. However, when the receiver is positioned on the yz axis, the secure area is significantly reduced. This degradation is due to the improved BER at extended horizontal distances, which, while enhances communication reliability, diminishes the area at which secure communications can be carried on.

In summary, the MIMO configuration provides the best results in terms of security, specially when the receiver is positioned on the xz vertical plane.

3.3.3 Conclusion

The presented work reported a comprehensive study on the use of optical wireless communication (OWC), in particular visible light communication (VLC), for implantable medical devices (IMDs). The presented work has explored the potential application of single LED and multiple LEDs multiple-input multiple-output (MIMO) techniques in IMD scenarios to manage and control the receiver region. Through MATLAB simulations, the feasibility and advantages of using VLC for subcutaneous IMDs was demonstrated. The impact of the receiver's movement on the bit error rate (BER) of the VLC system and identified the optimal receiver placement for VLC systems in IMD scenarios was evaluated. The intrinsic security properties of the system were also investigated by analyzing the evolution of the secure area caused by receiver's movements for both single LED and MIMO configuration.

Finally, we proposed a novel approach for OWC-based communication in IMDs, which can potentially improve the reliability, efficiency and security of IMDs. These contributions can pave the way for the development of more advanced and effective OWC-based IMDs in the future.

3.4 PLA in GC Communications: GC-based Biometric Authentication in WBAN

Galvanic coupling is a new non-RF transmission technology applied in the context of communications between devices placed on or inside human body. A general system model for Galvanic Coupling communication between two on body devices is showed in Figure 3.19. Two communicating devices are equipped with a pair of electrodes applied to human skin. The transmitting device apply a weak current to human body through the electrodes: the primary current flows directly between the transmitter's electrodes while a secondary weak current that carries the information is detected by the receiver's electrodes. The currents usually emitted by galvanic coupling devices are in the order of 0.5 mA, which is within the limit of 1 mA defined by the International Commission on non-ionizing radiation Protection (IC-NIRP) [92].

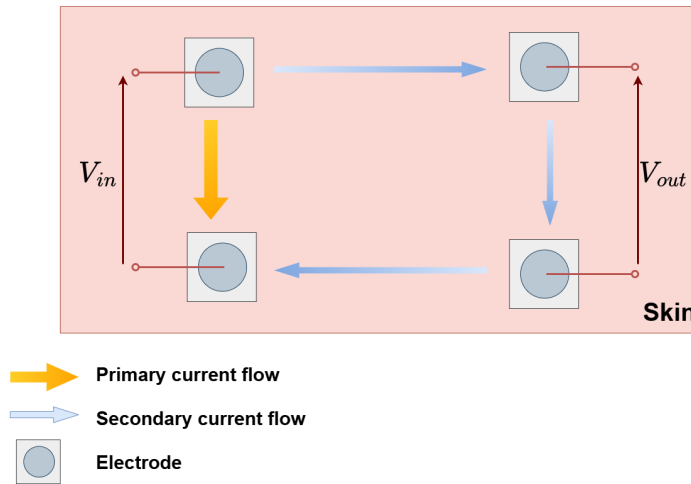


Figure 3.19: Galvanic Coupling system model.

It has been experimentally demonstrated that the secondary weak current can be detected at distances of up to 20-30 cm from the transmitting electrodes [92]. The operating frequency range of GC communications is from 1 kHz up to 100 MHz, which represents a trade-off between tissue attenuation and avoiding interference with other natural electrical signal, for

example, the one emitted by the nervous system.

Finally, one of the main advantages that makes GC communications a valid transmission technology for in and on body communications is the low power consumption required for communicating. GC consumes 0.24 nJ for every received bit while other RF technologies, like Zigbee and Ultra Wide Band, consumes from 2.5 nJ to 106 nJ per received bit [93].

In the context of PLA, a compelling feature of GC communication channel is that it presents subject-specific characteristics, which means that it is linked to the individual's physiological and anatomical features and varies from person to person. These features can be exploited to enable secure user identification.

In this thesis, a physical layer biometric authentication framework based on GC communications is introduced, exploiting a Long Short-Term Memory (LSTM) neural network to authenticate nodes of the GC communication system. Classic biometric authentication methods are based on the extraction of biometric features and exploits Artificial Intelligence (AI) and Machine Learning (ML) to enhance the reliability of the authentication.

The main advantage of using a LSTM neural network, is that it doesn't extract any feature and operates directly on the raw data. Therefore, it has been used to develop a lightweight authentication protocol, suitable for the low-resourced devices placed on body that communicates through GC. The developed protocol incorporates biometric verification at the physical layer and has been validated through experiments involving human subjects.

Related works. GC has been extensively studied as a secure and energy-efficient communication method for WBANs. Its ability to propagate signals primarily through biological tissues has led to various applications in medical and biometric systems. Prior research has addressed physical layer design for authentication [94], transmission optimization for implantable devices [95], and experimental channel characterization [96], highlighting the subject-specific and stable nature of GC channels.

Building on these foundations, further studies have investigated the use of GC for cryptographic key extraction. Methods have been proposed to generate secret keys by exploiting the reciprocity of GC channels, enabling secure communication between on-body devices without the need for key exchange [12]. Earlier approaches have also leveraged channel randomness and movement-induced variations to enhance key entropy and authentication

robustness [97,98]. The integration of authentication into the key extraction process has been shown to be effective in preventing impersonation attacks [99].

In parallel, biometric authentication using physiological signals, particularly EEG, has been extensively studied. Multi-class classifiers such as neural networks [100], multilayer perceptrons [101], and convolutional architectures [102–104] have demonstrated high accuracy in user identification. Other machine learning models, including decision trees, support vector machines, and probabilistic classifiers, have also yielded promising results [105]. One-class classifiers have been employed to detect intrusions and authenticate users based on specific EEG patterns [106–108], while anomaly detection techniques like Isolation Forest and Local Outlier Factor have been applied to physiological data, although not always in the context of authentication [109].

3.4.1 Methodology

In this section the overall adopted methodology is introduced. Both the considered threat models and the design of the authentication framework are presented.

In wearable medical systems, secure communication between on-body devices is essential to protect sensitive biometric data. Traditional wireless methods are vulnerable to interception, particularly in scenarios where adversaries attempt to compromise the system through physical proximity or direct interaction with the body. Two primary threat models are considered:

- Over-the-Air (OTA) attacks: this attack is a consequence of the fact that GC signals are not entirely confined to the skin: part of the transmitted power is irradiated to the outside [110]. An adversary can place external antennas in the vicinity of the subject and attempt to intercept, inject, or modify the irradiated GC signal.
- Direct contact attacks: the adversary physically connects a pair of electrodes to the subject's body and transmits falsified data to the receiver [110].

To counter these threats, the proposed system integrates biometric authentication directly into the physical communication layer. The approach leverages the subject-specific nature of the GC propagation channel, which

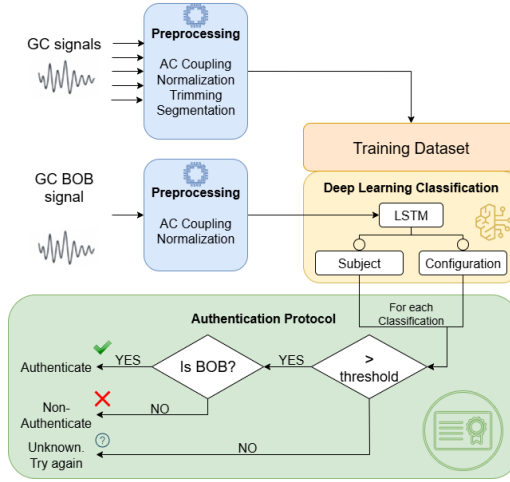


Figure 3.20: Architecture of proposed authentication system.

is influenced by physiological factors such as tissue composition, body geometry, and electrode placement. This results in a transmission path that is unique and stable over time, making it difficult to replicate or spoof.

The authentication framework is designed to operate in on-body configurations, such as communication between a wearable or implantable sensor and a GC-enabled smartwatch receiver. It consists of two neural network classifiers:

- A subject classifier, which identifies the individual transmitting the signal, ensuring protection against OTA attacks by verifying the identity regardless of electrode placement.
- A configuration classifier, which determines the distance between electrodes, enabling detection of anomalies in electrode setup and mitigating direct contact attacks.

The following subsections describe the signal processing pipeline, the architecture of the deep learning models for classification, and the authentication protocol, which is designed to operate efficiently in low-resource medical environments, as shown in Fig. 3.20.

GC Signal Dataset for Neural Network

The signals used in this work were pseudo-random noise (PrN) sequences, which were used as a representative surrogate for arbitrary data exchanged between wearable devices. A PrN sequence is generated through a linear feedback shift register (LFSR) and a generator polynomial of degree m , which defines which stages of the register are used as feedback of the register. In particular, the stages defined by the generator polynomial are combined through a modulo-2 addition (XOR) and then fed back to the input [111,112]. In this work, a generator polynomial of degree $m = 15$ was used to generate the PrN sequence which were then transmitted in base band. The generator polynomial was expressed as follows:

$$p(z) = z^{15} + z^{14} + 1 \quad (3.12)$$

To enable biometric authentication based on GC signals, the raw data acquired during the experimental campaign is preprocessed and structured into a dataset suitable for training and evaluating a neural network classifier. The pre-processing pipeline is designed to preserve subject-specific signal characteristics while removing irrelevant or misleading components. The first step involves removing the DC component from the acquired signal. This is achieved by applying AC coupling to the entire signal packet, effectively centering the signal around zero and eliminating any baseline offset introduced by the acquisition system. This ensures that only the dynamic variations of the signal are retained, which are more informative for classification.

Next, the signal is normalized with respect to its maximum absolute value. This operation scales the signal to the range $[-1, 1]$, removing amplitude variations caused by path loss and electrode placement. While path loss is a prominent and easily distinguishable feature, it is also the most vulnerable to estimation by an eavesdropper. By eliminating this dependency, the classifier is forced to rely on more subtle and robust features embedded in the signal structure.

The third step consists of trimming the signal to exclude portions where no transmission occurred. These segments contain only noise and do not contribute meaningful information for authentication. Removing them improves the signal-to-noise ratio and ensures the dataset contains only relevant samples.

The cleaned signal is then segmented into fixed-length portions of 100 samples, without overlap between segments. This segmentation allows the

neural network to process consistent input sizes and facilitates batch training. Each segment is treated as an independent sample, capturing a snapshot of the GC signal under specific conditions.

Finally, each signal segment is annotated with two labels: one indicating the electrode configuration used during acquisition, and the other identifying the subject. This dual-labeling enables the network to learn both configuration-specific and subject-specific patterns, supporting evaluation of classification performance across different experimental setups.

Deep Learning Classifier

Recurrent Neural Networks (RNNs) are a class of artificial neural networks designed to process sequential data. Unlike feedforward networks, RNNs include internal memory states that allow them to retain information across time steps. This temporal structure enables RNNs to model dynamic behaviors and dependencies in time-series data, making them particularly suitable for tasks such as speech recognition, handwriting analysis, and physiological signal processing.

LSTM networks are a specialized type of RNN designed to overcome the limitations of standard RNNs, such as vanishing or exploding gradients. LSTMs introduce memory cells and gating mechanisms that regulate the flow of information, enabling the network to learn long-term dependencies more effectively. This makes LSTMs particularly well-suited for analyzing biomedical signals like GC, where subtle temporal patterns may carry subject-specific information useful for authentication. The LSTM network used in this work is designed to classify segments of GC signals into subject identities. The input to the network consists of signal segments of 100 samples, preprocessed as described in 3.4.1. The architecture includes a single LSTM layer with 120 hidden units, followed by fully connected layers and a final softmax output layer. The softmax layer outputs a probability distribution over the possible classes, corresponding to the subjects in the dataset.

The dataset used for training and evaluation consists of 72,840 signal segments. A random 90% of the dataset (70,416 segments) is used for training, while the remaining 10% (7,824 segments) is reserved for testing. This split ensures that the model is evaluated on unseen data, providing a reliable estimate of its generalization performance.

The network is trained using the Adam optimizer, which combines the

advantages of adaptive learning rates and momentum. The learning rate is set to 0.002, and the model is trained for 200 epochs. The primary evaluation metric is classification accuracy, which measures the proportion of correctly identified signal segments. The training process is performed in batches, and early stopping is applied to prevent overfitting.

This architecture provides a balance between model complexity and computational efficiency, making it suitable for deployment on resource-constrained wearable devices. By learning directly from raw GC signals, the network avoids the need for handcrafted features and leverages the temporal dynamics of the signal for robust user authentication.

Authentication Protocol and Performance Evaluation

The authentication protocol is based on the output of two independent neural network classifiers. One classifier identifies the subject transmitting the signal, while the other determines the electrode configuration used during signal acquisition. Both networks operate on 100-sample GC signal segments and produce a probability distribution over their respective classes using a softmax activation function. The class associated with the highest probability is considered the most likely prediction.

However, in biometric authentication systems, a classification result alone may not be sufficient. A misclassification could be interpreted as a successful impersonation attempt. To mitigate this risk, the protocol introduces a confidence score threshold: if the highest predicted probability, which is the confidence score, does not exceed a predefined value θ , the system rejects the authentication attempt. This mechanism allows the system to respond with “uncertain” rather than forcing a potentially incorrect classification.

This threshold-based decision introduces a trade-off between responsiveness and security. Increasing θ improves the reliability of accepted predictions but leads to a higher rejection rate. Conversely, lowering θ increases the number of accepted segments but may reduce robustness against adversarial attacks. Given that each 100-sample segment corresponds to approximately 2 milliseconds of transmission time, the protocol can afford to discard uncertain segments and request retransmission without significantly impacting system performance.

To evaluate the effectiveness of the protocol, two key performance indicators are defined:

- Reliability: the proportion of signal segments that are accepted by the

system, *i.e.*, those for which the confidence score exceeds the threshold θ . It is defined as:

$$Reliability = \frac{N_{\text{accepted}}}{N_{\text{total}}},$$

where N_{accepted} is the number of segments with confidence score above θ , and N_{total} is the total number of transmitted segments.

- Security: the proportion of correctly classified segments among those that are accepted. It reflects the system's robustness against impersonation and is defined as:

$$Security = \frac{N_{\text{correct}}}{N_{\text{accepted}}},$$

where N_{correct} is the number of accepted segments that are correctly classified.

In the results section, the threshold θ is varied from 0 (inclusive) to 1 (exclusive) to analyse the trade-off between reliability and security, and to identify optimal operating points for different application scenarios.

3.4.2 Experimental Campaign

In previous work [12], the GC channel was characterized through estimation of the Channel Impulse Response (CIR) for cryptographic key extraction.

In contrast, the present study focuses on biometric authentication and directly uses the received signal segment as input to a classification model. This approach eliminates the need for channel estimation, reducing computational complexity and enabling faster processing suitable for real-time applications.

Measurement Setup and Experimental Protocol

The experimental platform (see Fig. 3.21) consists of two battery-powered laptops acting as transmitter (TX) and receiver (RX), connected via LINE OUT and LINE IN ports, respectively. Signal generation and acquisition are performed using MATLAB. The analog signal is transmitted through the TX sound card and injected into the body via two Ag/AgCl surface electrodes (PG10C, FIAB, CE-certified). The received signal is captured by two electrodes and digitized at 48 kHz with 16-bit resolution.

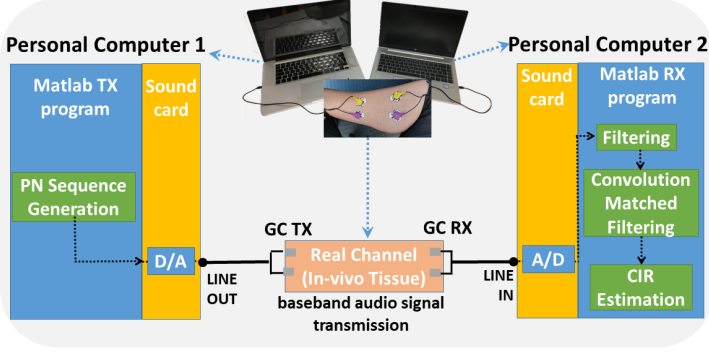


Figure 3.21: Experimental setup for GC signal acquisition, showing the connection between transmitter and receiver, and the placement of electrodes on the forearm.

To ensure galvanic isolation and avoid shared ground paths, both laptops operate on battery power. A 50 Hz notch filter is applied to suppress power line interference.

Experiments were conducted on eight healthy volunteers (ages 24-30). Electrodes were placed on the inner forearm, with multiple configurations tested to evaluate signal variability. As illustrated in Figure 3.22, the longitudinal distance (d_l) between TX and RX pairs was set to 3, 5, 10, and 15 cm, while the transverse distance (d_t) within each pair was 3 or 5 cm. Each configuration was repeated five times to ensure statistical consistency.

Before each test, the system was configured to ensure optimal signal transmission and acquisition. The following steps were followed:

- Cable connection: TX connected to LINE OUT, RX to LINE IN.
- Audio settings: LINE IN and LINE OUT volumes set to 80% of maximum; all other channels disabled.
- Format configuration: Both laptops set to “16-bit, 48000 Hz”; audio enhancements disabled.
- Power isolation: Laptops disconnected from power supply to avoid ground loops.

- Electrode placement: Electrodes applied to the skin with the distances specified above.
- Transmission: RX software started approximately 2 seconds before TX to ensure synchronization.

If the signal was not properly acquired, the test was repeated. This procedure ensures consistent and reliable acquisition of GC signals for biometric analysis under controlled conditions.

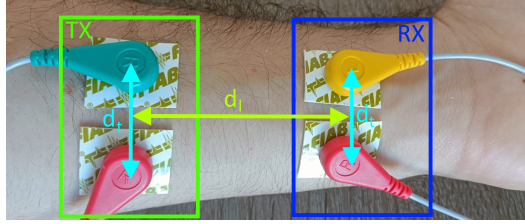


Figure 3.22: Electrode placement scheme illustrating longitudinal distance (d_l) between TX and RX pairs, and transverse distance (d_t) within each pair.

3.4.3 Numerical Results and Discussion

This section presents the performance evaluation of the proposed biometric authentication system based on GC signals. The analysis focuses on the system's ability to withstand the two threat models introduced: direct contact attacks, which are addressed through classification of the electrode configuration, and OTA attacks, mitigated by identifying the subject transmitting the signal.

For each scenario, the system's behavior is studied by varying the confidence score threshold θ in the range $[0, 1)$, and observing how this affects the two key performance indicators: reliability and security. As defined earlier, reliability measures the proportion of signal segments accepted by the system, while security quantifies the accuracy of classification among those accepted segments.

To complement the analysis of these metrics, we also present confusion matrices for two representative operating conditions. The first corresponds to $\theta = 0$, where the confidence-based rejection mechanism is disabled and all segments are accepted. The second refers to the case where θ is set such that

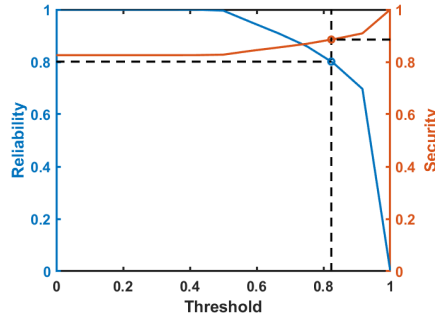


Figure 3.23: Reliability and Security Performance.

the system achieves a reliability of 0.8, meaning that 2 out of 10 segments are rejected.

Reliability and Security Performance for Direct Contact Attack

The evaluation of the system under the direct contact attack scenario focuses on its ability to correctly classify the electrode configuration used during signal acquisition. This classification is crucial to detect unauthorized physical connections to the body, where an adversary attempts to transmit falsified data by mimicking the legitimate electrode setup.

Figure 3.23 illustrates the behaviour of reliability and security as the confidence score threshold θ is varied. When the threshold is set to zero, meaning that all segments are accepted regardless of confidence, the system achieves a security of 82%. This indicates that even without any filtering, the classifier is able to correctly identify the configuration in the majority of cases. By increasing θ to a value that yields a reliability of 0.8, the security improves to 88%, showing that discarding low-confidence segments enhances the robustness of the classification. However, achieving perfect security (100%) requires setting θ so high that only 0.02% of the segments are accepted. Given that each segment corresponds to approximately 2 milliseconds of transmission time, this implies that authenticating with absolute certainty would require around 10 seconds of continuous transmission, which is impractical for real-time applications.

Figure 3.24 provides further insight into the classification performance by comparing the confusion matrices for two operating points: $\theta = 0$ and the

True Class	3;10	1065	48			198	
	3;15	46	1212				
	3;3			1232	9		118
	3;5			32	981	23	338
	5;10	128			95	1038	9
	5;5			183	135	1	933
	Predicted Class	3;10	3;15	3;3	3;5	5;10	5;5

(a) Reliability = 100%.

True Class	3;10	963	33			124	
	3;15	37	1180				
	3;3			1077	2		42
	3;5			12	756	13	156
	5;10	98			70	935	7
	5;5			66	58		639
	Predicted Class	3;10	3;15	3;3	3;5	5;10	5;5

(b) Reliability = 80%.

Figure 3.24: Confusion Matrix.

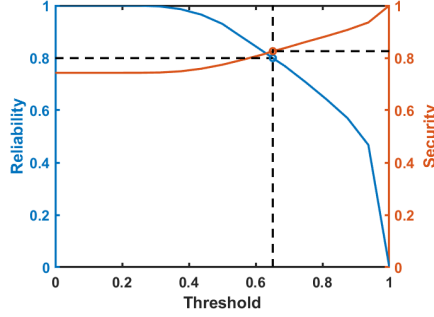


Figure 3.25: Reliability and Security Performance.

threshold corresponding to 80% reliability. In both cases, the system demonstrates high overall accuracy, but struggles to distinguish between configurations with similar electrode placements. Misclassification events are more frequent when the longitudinal and transverse distances between electrodes are close, suggesting that the signal features become less discriminative in these conditions.

Despite these limitations, the authentication system remains effective in detecting adversarial configurations, provided that the attacker is not positioned too close to the legitimate transmitter. The results confirm that the GC signal retains sufficient spatial sensitivity to support configuration-based authentication, especially when low-confidence segments are filtered out.

Reliability and Security Performance for OTA Attack

In the context of OTA attacks, the authentication system is evaluated based on its ability to correctly identify the subject transmitting the GC signal. This classification is essential to prevent adversaries from impersonating legitimate users by injecting signals wirelessly through their own GC transmission channel.

Figure 3.25 shows the variation of reliability and security as the confidence score threshold θ increases. When the threshold is set to zero, meaning that all segments are accepted, the system achieves a security of 74%. This baseline performance indicates that the classifier is able to distinguish users with reasonable accuracy even without filtering. By increasing θ to a value

that yields a reliability of 0.8, the security improves to 82%, confirming that discarding low-confidence segments enhances the system's resistance to impersonation. As in the previous scenario, achieving perfect security (100%) requires a threshold so strict that only 0.02% of the segments are accepted, resulting in an impractical authentication time of approximately 10 seconds.

The confusion matrices in Figure 3.26 provide additional insight into the classification behavior. Unlike the direct contact attack scenario, where errors are concentrated among similar configurations, the misclassifications in the OTA case are more evenly distributed across the matrix. As the threshold increases from 0 to the value corresponding to 80% reliability, the number of errors gradually decreases across all classes. This suggests that the classifier is able to refine its predictions uniformly, without being disproportionately affected by specific subject pairs.

Although some subjects exhibit signal characteristics that are more similar to others, leading to occasional misclassifications, the differences are less pronounced than in the configuration classification task. Overall, the system demonstrates strong discriminative capability in identifying users, and the use of a confidence-based rejection mechanism significantly improves its robustness against OTA attacks.

3.4.4 Conclusion

This work presents a physical-layer biometric authentication framework for WBANs based on GC signals. By leveraging a deep LSTM network trained on raw GC data, the system successfully exploits the subject-specific features of the propagation channel, providing robust protection against both direct contact and over-the-air attacks. Experimental results confirm that the confidence-based rejection mechanism significantly improves security while maintaining acceptable reliability, making the approach suitable for real-time applications in medical environments.

Future work will focus on expanding the experimental dataset to include a larger and more diverse subject population, improving generalization and reducing bias. Additional research will explore the integration of multimodal biometric features and adaptive thresholding techniques to dynamically balance security and usability. Moreover, reliability assessment through repeated measurements and cross-session analysis will be conducted to further validate system consistency over time.

True Class	AAR	748	1	7	40	43	28	63	41
	ABG	4	908	16	19	4	20	21	10
	DS	17	103	704	17	12	59	8	16
	FF	52	8	7	619	122	41	102	16
	MM	41	4	2	144	691	77	65	9
	MZ	29	21	50	30	100	655	29	29
	SM	42	12	9	83	61	43	714	41
	ZX	75	16	15	14	5	14	53	775
		AAR	ABG	DS	FF	MM	MZ	SM	ZX
		Predicted Class							

(a) Reliability = 100%.

True Class	AAR	691	1	3	15	19	19	31	23
	ABG	1	871	11	9	1	11	13	8
	DS	11	65	646	9	2	30	3	10
	FF	36	5	2	472	67	15	46	7
	MM	23	3	1	81	580	45	33	6
	MZ	22	12	34	12	66	577	17	14
	SM	28	8	5	34	23	26	605	22
	ZX	56	9	5	4	4	3	21	722
		AAR	ABG	DS	FF	MM	MZ	SM	ZX
		Predicted Class							

(b) Reliability = 80%.

Figure 3.26: Confusion Matrix.

3.5 Authentication of Outputs of Human Body: Measuring Aesthetic Information in AI- Generated and Human Artistic Images

The aim of the work presented in this section is to define a new approach to the problem of detecting AI generated signals and data to emulate outputs of the human body in the context of WBANs. In this context, an artwork is modeled as a special kind of signal used by an artist to communicate their psycho-physical status, such as their thoughts or emotions.

Here a novel and interpretable framework for authorship detection in visual art, grounded in information theory, is presented.

The proposed theoretical approach applies information theory to quantify aesthetic properties of images and interprets the creative process as an information transmission channel between the artist and the viewer. Within this framework, the existence of an aesthetic transmission gap is hypothesized: a measurable discrepancy in how emotional and aesthetic intent is conveyed, that AI models cannot fully bridge.

Inspired by the Informational Aesthetics (InfAes) metrics introduced by [24, 113], in this thesis the InfAes metrics are extended, adapted and applied to the problem of distinguishing human-made artworks from those generated by AI. Specifically, the contributions of the presented work can be summarized as:

- The artistic process has been formalized as a communication system, where the artist acts as the transmitter, the artwork as the channel, and the viewer as the receiver. Generative AI models are hypothesized to be unable of fully replicating the expressive signal transmitted by human artists.
- The six aesthetic metrics defined within the InfAes framework have been reinterpreted, adapted and applied to both human and AI-generated images to evaluate their discriminative power. These metrics assess images across multiple levels, quantifying both structural and aesthetic properties, quantifying its structural and aesthetic properties.
- A curated dataset comprising 994 human artworks and 994 AI images generated using Stable Diffusion v2.1 [114] was constructed. The number of paintings was established by selecting, from the WikiArt

database [115], those human artworks whose dimensions closely matched the native resolution of the Stable Diffusion model employed for generating the synthetic images. Details regarding dataset construction and pre-processing are provided in Section,3.5.3.

- An empirical evaluation using clustering and classification techniques has been conducted. Specifically, the following techniques have been applied:
 - Gaussian fitting and K-means clustering (Section 3.5.4) to assess separability of human and AI images based on individual and combined metrics.
 - A Fine Gaussian Support Vector Machine (SVM) classifier trained on selected InfAes metrics (Section 3.5.4), achieving superior performance compared to a state-of-the-art deep learning model (Swin Transformer [116]), with an accuracy of 84.26%.
- Metric complexity and trade-offs are analyzed (Section,3.5.1) by evaluating the computational cost associated with each metric and identifying optimal subsets that balance performance and efficiency of the classification.
- A comparison with existing methods is presented (Section,3.5.4), demonstrating that the interpretable, training-light framework proposed herein outperforms black-box CNN-based classifiers in terms of robustness and generalizability.

Related works. Several studies have explored entropy and complexity metrics to distinguish AI-generated images from human-made ones. Papia et al. [117] use Shannon entropy and multi-scale complexity across artistic genres, while Yuan et al. [118] propose multi-granularity entropy patterns for detection. Cozzolino et al. [119] introduce a zero-shot entropy-based detector using a lossless encoder, and Zhang et al. [120] apply prediction entropy and Kullback-Leibler divergence to detect adversarial perturbations.

While these methods rely on statistical image properties, the approach presented in this thesis builds on the information theory to interpret the creative process as an information channel. Handcrafted metrics that quantify aesthetic structure and intent were used, offering interpretability and a principled framework for authorship detection.

Other works focus on local pixel relationships and texture analysis. Zhong et al. [121] detect synthetic images by segmenting them into patches and analyzing pixel correlations, while Tan et al. [122] examine neighboring pixel distributions during upsampling. In contrast, the method proposed in this thesis operates at a higher semantic level, analyzing global and compositional self-similarity and mutual information to capture the structural and aesthetic coherence of artworks.

A large body of research relies on convolutional neural networks (CNNs) to learn features from raw image data. Chinta et al. [123] and Khan et al. [124] use CNNs to extract hierarchical and semantic features, while Li and Stamp [125] compare CNNs with traditional models across genres. Bird et al. [126], Harsanto et al. [127], Oo et al. [128], and Ohal et al. [129] propose CNN-based classifiers trained on curated datasets. Unlike these data-driven, black-box approaches, the method proposed in this thesis is training-free and interpretable. We use handcrafted metrics rooted in information theory, enabling transparent analysis of aesthetic and structural properties without requiring large labeled datasets or deep feature learning.

Finally, Ha et al. [116] investigate the performance of both automated detectors and human experts in distinguishing AI-generated art across styles and models. Their findings show that while automated methods perform well under standard conditions, they struggle with adversarial perturbations, and human experts are prone to false positives. The approach proposed in this thesis offers robustness and interpretability without relying on human annotation or adversarial training. By modeling the artistic process as an information-theoretic transmission, we provide a scalable and theoretically grounded framework for authorship detection.

3.5.1 System Model and InfAes Metrics

The artistic creation process can be conceptualized as a telecommunication system (Fig. 3.27), inspired by information theory. This analogy frames the transmission of aesthetic and emotional intent from the artist to the viewer as a structured flow of information:

- **Transmitter (TX):** The artist, who seeks to express their psycho-physical condition (emotions, thoughts, and intentions) through their painting technique.
- **Channel:** The artwork itself, which carries the expressive signal. This

includes elements like color palette, composition, and structural patterns.

- Receiver (RX): The human viewer, who interprets the artwork based on their own empathic abilities and perceptual faculties.

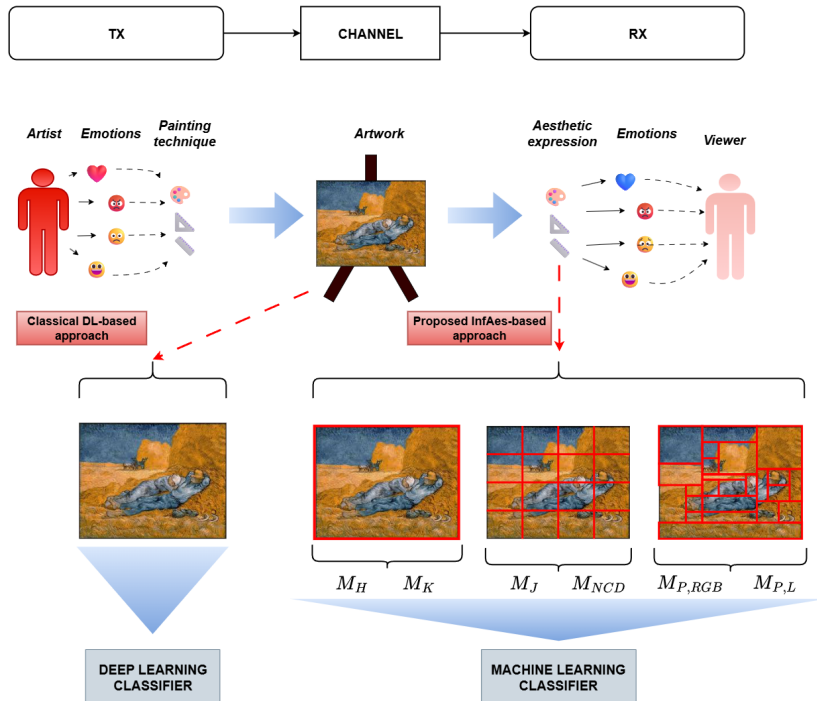


Figure 3.27: The system model. The different InfAes metrics are illustrated at the bottom, depending on the parts of the image they act on.

This work focuses exclusively on the transmission side, analyzing the expressive signal encoded by the artist into the artwork without modeling the viewer’s perceptual or cognitive processes. This approach enables the isolation and quantification of the structural and aesthetic properties of the artwork, avoiding the variability and subjectivity inherent in human reception.

In this conceptual model, the artist expresses an intrinsic “human factor”, linked to their psycho-physical condition, within the artwork, which

generates a certain degree of aesthetic expression.

Let's now consider an AI image-generation model trained to replicate both the artwork (the channel) and the artist's painting technique (part of the TX in Fig. 3.27). In this study, it's hypothesized that the AI model is unable to fully capture and reproduce the human factor conveyed by the artist. This limitation gives rise to an "aesthetic transmission gap": a measurable discrepancy between the artist's expressive intent and the viewer's reception, which may serve to distinguish AI-generated artworks.

To rigorously investigate the presence of a measurable aesthetic transmission gap, hence the existence of an intrinsic human factor, in artistic expression, the framework of Informational Aesthetics (InfAes) is adopted. InfAes provides a principled, interpretable approach to quantify aesthetic and structural properties of images using information-theoretic metrics. Within the presented system model, InfAes metrics allow to formalize and measure the aesthetic transmission gap.

The bottom of Fig. 3.27 illustrates how different InfAes metrics act on various parts of the image:

- Global Aesthetics (GA): Analyze the image as a whole, focusing on entropy and complexity.
- Compositional Aesthetics (CA): Examine self-similarity across image regions.
- Mutual Information Partitioning (MIP): Measure coherence between color/luminance distributions and spatial regions of the image.

These metrics serve as tools to quantify how effectively the artwork conveys aesthetic information, and to detect whether the expressive signal is more likely to originate from a human or an AI.

Informational Aesthetics Metrics

The concept of aesthetic measure was originally formalized by George D. Kirkhoff in [113]. He defined the aesthetic measure M as the ratio between order O and complexity C ,

$$M = \frac{O}{C}. \quad (3.13)$$

The authors in [24] formalized the aesthetic measure concept in [113] using an information theoretic perspective. In particular, they defined three

sets of metrics that represent an aesthetic measure: the *Global Aesthetics* (GA) metrics, the *Compositional Aesthetics* (CA) metrics, and the *Mutual Information-based Partitioning* (MIP) metrics.

GA metrics evaluate the overall visual characteristics of the image, treating it as a single, unified entity. They do not consider how the image is spatially divided or arranged, but rather focus on global properties such as color palette and arrangement complexity.

CA metrics assess the structural arrangement of the image, focusing on how different regions or parts relate to one another in terms of similarity, variation, or pattern. They evaluate the spatial distribution of the visual elements in the image.

MIP metrics measure how coherently the distribution of the colors in an image is reflected across its spatial regions. In other words, it quantifies how much knowing one region of the image tells you about the colors distribution inside it, using the concept of mutual information from information theory.

GA and CA metrics has been computed considering an RGB and a Y_{709} luminance representation of the images, respectively. MIP metric is applied to both the RGB and the Y_{709} luminance representation of the images. Y_{709} luminance of a pixel x is the weighted sum of linear RGB values that approximates perceived brightness

$$Y_{709}(x) = 0.2126R(x) + 0.7152G(x) + 0.0722B(x), \quad (3.14)$$

where $R(x)$, $G(x)$ and $B(x)$ are the levels of red, green and blue of pixel x , respectively [130]. It refers to the luminance component defined in the ITU-R BT.709 standard [131], which is used for HDTV color encoding, and it represents how bright a pixel appears to the human eye based on its RGB values.

Let's consider a generic image with N pixels. The histogram $\mathbf{h}$ of the image can be defined as the probability distribution of its color or luminance levels. These levels are known as bins of the histogram. The vector $\mathbf{h}$ can be expressed as:

$$\mathbf{h} = [h_1, h_2, \dots, h_b, \dots, h_{N_{bin,*}}] \quad (3.15)$$

where $N_{bin,*}$ is the number of bins of the histogram and is equal to $N_{bin,L} = 256$ for a luminance and to $N_{bin,RGB} = 256^3$ for an RGB representation of the image. The generic element of the vector h_b represents the probability of having the b -th color or luminance level in the image and it can be expressed

as $h_b = \frac{N_b}{N}$, where N_b is the number of pixels which have the b -th level of color or luminance.

Global Aesthetics (GA) metrics The first global aesthetics metric is the *Global Aesthetic Entropy Distance*, M_H . It is defined as

$$M_H = \frac{H_{max} - H_I}{H_{max}}, \quad (3.16)$$

where $H_I = H(\mathbf{h})$ is the entropy of the color histogram of the image and $H_{max} = H(\text{RGB})$ represents the maximum entropy of an RGB image.

For an RGB digital image with a depth of 8 bits per color channel, H_{max} is equal to 24 bit [24]. M_H quantifies the reduction of uncertainty derived from the artist's choice of a specific color palette, represented by H_I . In the presented work, H_I is defined as

$$H_I = - \sum_{b=1}^{N_{bin,RGB}} \frac{N_b}{N} \log_2 \left(\frac{N_b}{N} \right), \quad (3.17)$$

where $N_{bin,RGB}$ is the number of bins of the image histogram, *i.e.*, the number of all possible RGB colors, N is the number of pixels of the image, N_b is the number of pixels in the b -th bin, *i.e.*, the number of pixels of a specific color. By applying (3.16) to a set of digital paintings with same size and color depth, a lower value of M_H is expected for those with a high variety of colors, since this would result in a higher entropy H_I . Conversely, for images with low variety of colors, H_I is lower, resulting in a higher value of M_H .

The second global aesthetic metric is the *Global Aesthetic Complexity*, M_K . It is defined as

$$M_K = \frac{N \cdot H_{max} - K_I}{N \cdot H_{max}}, \quad (3.18)$$

where K_I is the Kolmogorov complexity of the image, and $N \cdot H_{max}$ represents the size (in bit) of the uncompressed raw image. In general, the Kolmogorov complexity $K(\xi)$ of a string ξ is defined as the shortest program to compute ξ on a universal computer. For an image, the Kolmogorov complexity represents the minimal amount of information needed to generate that image. It can be proved that $K(\xi)$ is not computable, but it can be approximated using a standard compression algorithm [24, 132]. In this work, JPEG compressor has been applied [133] and thus K_I is the size in

bit of the compressed image. The level of compression achieved by JPEG compression, which corresponds to the magnitude of K_I , is correlated to the level of complexity of the image. A complex image, with a high level of details, will be less compressed by the JPEG algorithm, and a lower value of M_K will be obtained. Vice-versa if the image is simple and homogeneous.

Compositional Aesthetics (CA) metrics Compositional aesthetics metrics evaluate the internal consistency of an image by measuring how its regions are similar to one another, a property we refer to as *self-similarity*. An image is self-similar if its regions are statistically or structurally similar, regardless of their size or position in the image. Two regions of an image are *statistically similar* if the probability distributions of their colors are similar. Two regions are *structurally similar* if they show a similar level of complexity, in other words, a similar level of details. The regions are defined here by superimposing an $r \times r$ regular grid over the image, resulting in a total of $R_G = r^2$ regions.

The first compositional aesthetics metric is the *Compositional Aesthetic Statistical Self-similarity*, M_J , and it is defined as

$$M_J = 1 - \frac{J(N^{(1)}, \dots, N^{(R_G)}; \mathbf{h}^{(1)}, \dots, \mathbf{h}^{(R_G)})}{H_I} = \frac{\sum_{i=1}^{R_G} N^{(i)} H_I^{(i)}}{H_I}, \quad (3.19)$$

where $J(N^{(1)}, \dots, N^{(R_G)}; \mathbf{h}^{(1)}, \dots, \mathbf{h}^{(R_G)})$, is the Jensen-Shannon divergence [134], $N^{(i)}$ with $i = 1, \dots, R_G$ is the number of pixels of the i -th region of the grid, and $H_I^{(i)}$ is the entropy of the luminance of that region. Since the superimposed grid is regular, the regions have all the same number of pixels $N^{(i)} = N_G, \forall i$. The vector $\mathbf{h}^{(i)}$ is the luminance histogram of the i -th region, as in (3.15). The generic element $h_b^{(i)}$ of the vector is here defined as $h_b^{(i)} = \frac{N_b^{(i)}}{N_G}$, where $N_b^{(i)}$ is the number of pixels which have the b -th luminance level in the i -th region, and N_G is the number of pixels in each of the R_G (equally sized) sub-regions of the image.

The entropy of the i -th region is thus computed as

$$H_I^{(i)} = - \sum_{b=1}^{N_{bin,L}} \frac{N_b^{(i)}}{N_G} \log_2 \left(\frac{N_b^{(i)}}{N_G} \right), \quad (3.20)$$

The Jensen-Shannon divergence measures the “distance” between two probability distributions, *i.e.*, their degree of dissimilarity. The complementary to one is used as a *similarity* measure. In our case, the Jensen-Shannon divergence can be expressed as

$$\begin{aligned} J(N^{(1)}, \dots, N^{(R_G)}; \mathbf{h}^{(1)}, \dots, \mathbf{h}^{(R_G)}) = \\ = H_I - \sum_{i=1}^{R_G} N_G H_I^{(i)} \geq 0. \end{aligned} \quad (3.21)$$

Eq. (3.21) equals zero only if all $\mathbf{h}^{(i)}$ are equal, *i.e.*, if all image regions show the same luminance probability distribution.

Thus, M_J quantifies the statistical self-similarity of an image by computing the Jensen-Shannon divergence across regional luminance distributions, normalized by the global image entropy. A higher M_J indicates greater similarity among regions, reflecting a more coherent luminance structure of the whole image.

The second compositional aesthetic metric is the *Compositional Aesthetics Structural Self-similarity* (CAstrS), M_{NCD} . It is defined as

$$M_{NCD} = 1 - \text{avg}_{1 \leq i < j \leq r} \{NCD^{(i,j)}\}. \quad (3.22)$$

where $NCD^{(i,j)}$ is the Normalized Compression Distance between the i -th and the j -th region of the image [24, 132]. $NCD^{(i,j)}$ is defined as

$$NCD^{(i,j)} = \frac{K_I^{(i,j)} - \min\{K_I^{(i)}, K_I^{(j)}\}}{\max\{K_I^{(i)}, K_I^{(j)}\}}, \quad (3.23)$$

where $K_I^{(i)}$ and $K_I^{(j)}$ are the Kolmogorov complexity of the JPEG compressed i -th and the j -th image region, respectively, and $K_I^{(i,j)}$ is the Kolmogorov complexity of the compressed pair of regions. $K_I^{(i)}$ and $K_I^{(j)}$ are obtained by cropping the i -th and j -th region from the original image and subsequently compressing them into JPEG format. Similarly, $K_I^{(i,j)}$ is obtained by cropping the i -th and j -th regions, placing them side by side, and then compressing the resulting composite image into JPEG format.

If two regions are very similar, compressing them together will not add much more data than compressing one of them alone (low NCD). If two regions are not similar, the joint compression will be significantly larger (high NCD). If an image shows a higher structural self-similarity, then its regions are on average structurally similar. Therefore, the average NCD in (3.22) will be lower, resulting on a higher M_{NCD} value.

Mutual Information-based Partitioning Inspired by the work in [24], in this thesis is proposed a method to analyze the creative process by modeling it as an information channel between the color palette and the regions of the image. The analysis is carried out by applying a partitioning algorithm to the images. Depending on whether the algorithm is applied to an RGB image or to its Y_{709} luminance representation, two metrics are defined: the first one is the *Mutual-Information based Partitioning of Color Structure*, $M_{P,RGB}$, and the second is the *Mutual-Information based Partitioning of Luminance Structure*, $M_{P,L}$.

Without loss of generality, let us focus on an RGB image with N pixels and a color histogram $\mathbf{h}$ defined as in (3.15) with $N_{bin,RGB}$ color bins and N_b pixels in the b -th color bin. Let us also define the vector containing the set of image regions obtained through the MIP algorithm 1 as $\mathbf{r} = [r_1, r_2, \dots, r_i, \dots, r_{R_P}]$, where R_P is the number of spatial regions obtained at the end of the algorithm. In this case the regions are not obtained by simply superimposing a regular grid as in M_J and M_{NCD} .

The MIP metric M_P can be defined as

$$M_P = \frac{I(\mathbf{h}; \mathbf{r})}{H_I} \quad (3.24)$$

where $I(\mathbf{h}; \mathbf{r})$ is the mutual information between the colors distribution and the set of R_P regions of the image, normalized to the image entropy H_I . The specific R_P regions of the image are obtained as an output of the MIP algorithm. The MIP algorithm divides an image into regions by recursively splitting it in a way that maximizes mutual information between colors distribution (or luminance) and spatial location (which region a pixel belongs to). This means that the algorithm tries to find partitions where color and space are statistically dependent, in other words, where certain colors tend to appear in specific parts of the image.

The mutual information between the color distribution and the spatial partitioning is given by:

$$I(\mathbf{h}; \mathbf{r}) = H(\mathbf{h}) - H(\mathbf{h} | \mathbf{r}), \quad (3.25)$$

where $H(\mathbf{h})$ is the entropy of the colors distribution and $H(\mathbf{h}|\mathbf{r})$ is the conditional entropy of the colors distribution given the spatial regions. This mutual information quantifies the degree of dependence between color and spatial structure in the image, and reflects how much information about the

color distribution is preserved across the spatial regions identified by the MIP algorithm at convergence.

The conditional entropy $H(\mathbf{h}|\mathbf{r})$ in (3.25) can be written as

$$H(\mathbf{h} | \mathbf{r}) = \sum_{i=1}^{R_P} \frac{N^{(i)}}{N} H(\mathbf{h}|r_i) \quad (3.26)$$

where $N^{(i)}$ is the number of pixel of the i -th region and $H(\mathbf{h}|r_i)$ is the entropy of $\mathbf{h}$ conditioned to the i -th region of the vector $\mathbf{r}$. Since conditioning the entropy on the i -th region is equivalent to computing the entropy of the region itself, $H(\mathbf{h}|r_i)$ can be expressed as

$$H(\mathbf{h} | r_i) = - \sum_{b=1}^{N_{bin,*}} \frac{N_b^{(i)}}{N^{(i)}} \log_2 \frac{N_b^{(i)}}{N^{(i)}} \quad (3.27)$$

where $N_b^{(i)}$ is the number of pixels of the i -th region in the b -th bin of the histogram and $N^{(i)}$ is the number of pixel of the i -th region. The algorithm stops when the number of regions reaches a selected threshold R_P .

Let us now describe the partitioning procedure (Algorithm 1) to compute the vector of regions $\mathbf{r}$:

- The image is iteratively split into regions of decreasing size (analogously to a decision tree). At each iterative step s , the elements of the vector of the regions $\mathbf{r}_s = [r_1, \dots, r_i, \dots, r_s]$ are sorted in descending order according to $H(\mathbf{h}|r_i)$, $\forall r_i \in \mathbf{r}_s$ (see Eq. (3.27)).
- The region r_m that shows the maximum conditional entropy is selected, $r_m = \arg \max_{r_i \in \mathbf{r}_s} H(\mathbf{h}|r_i)$. This step is performed to select the region which contains the highest information, instead of selecting it randomly.
- Then, the algorithm evaluates all possible binary splits (horizontal or vertical cuts) of the region r_m . Let $r_m \subset \mathbb{R}^2$ be a rectangular image region. We define the set of all possible binary splits of r_m as:

$$\mathcal{S}_m = \{(r_\alpha, r_\beta) | r_\alpha \cup r_\beta = r_m, r_\alpha \cap r_\beta = \emptyset\}$$

where r_α, r_β are axis-aligned subregions. Each split $(r_\alpha, r_\beta) \in \mathcal{S}_m$ is generated by a function:

$$(r_\alpha, r_\beta) = f_{\text{split}}(r_m, d, \tau)$$

where:

- $d \in \{\text{horizontal}, \text{vertical}\}$ is the direction of the cut,
- $\tau \in \mathbb{N}$ is the thickness of the subregion (in pixels),
- r_α is the subregion of thickness τ in direction d ,
- $r_\beta = r_m \setminus r_\alpha$ is the complementary region.

The function f_{split} generates axis-aligned subregions that preserve the rectangular structure of r_m , and it is used to enumerate candidate partitions for mutual information evaluation. Its behavior is the following:

- If $d = \text{horizontal}$, the region is split along the y-axis:
 - * r_α contains the top τ rows of r_m (in pixels),
 - * r_β contains the remaining rows.
- If $d = \text{vertical}$, the region is split along the x-axis:
 - * r_α contains the leftmost τ columns of r_m (in pixels),
 - * r_β contains the remaining columns.

For every couple of candidate sub-regions $(r_\alpha, r_\beta) \in \mathcal{S}_m$, the conditional entropies $H(\mathbf{h} \mid r_\alpha)$ and $H(\mathbf{h} \mid r_\beta)$ are computed with (3.27). Then, the conditional entropy of the split is computed as

$$H(\mathbf{h} \mid [r_\alpha, r_\beta]) = \frac{N^{(\alpha)}}{N} H(\mathbf{h} \mid r_\alpha) + \frac{N^{(\beta)}}{N} H(\mathbf{h} \mid r_\beta), \quad (3.28)$$

where $N^{(\alpha)}$ and $N^{(\beta)}$ represent the number of pixels of r_α and r_β , respectively. Thus, at the s -th step of the algorithm, two new regions are defined by selecting the split $(\bar{r}_\alpha, \bar{r}_\beta)$ which provides the lower conditional entropy

$$(\bar{r}_\alpha, \bar{r}_\beta) = \arg \min_{r_\alpha, r_\beta} \{H(\mathbf{h} \mid [r_\alpha, r_\beta])\}, \quad \forall r_\alpha, r_\beta \in \mathcal{S}_m \quad (3.29)$$

- The vector $\mathbf{r}_s$ is then updated by replacing r_m with $\bar{r}_\alpha$ and $\bar{r}_\beta$. The vector $\mathbf{r}_s$ thus increases its size at each step.
- This process continues until a predefined stopping criterion is met, *e.g.*, when the number of regions reaches a specified threshold R_P , or when further splits no longer yield a significant gain in mutual information. At the end of the process the vector $\mathbf{r}_s$ becomes $\mathbf{r}$ which is then used to compute the total mutual information $I(\mathbf{h}; \mathbf{r})$ (see Eq. (3.25)).

- The metric M_P is finally computed as the total mutual information achieved after partitioning $I(\mathbf{h}; \mathbf{r})$, normalized by the image's entropy H_I (see Eq. (3.24)).

A highly structured image, where specific colors are spatially localized in distinct regions, will result in a higher M_P value, indicating strong mutual dependence between color and spatial layout. Conversely, if the image is less structured, with colors uniformly or randomly distributed across regions, the mutual information will be lower, and so will M_P .

Algorithm 1 Mutual Information Partitioning (MIP)

Require: Image I with N pixels, histogram bins $\mathbf{h} = \{h_1, \dots, h_{N_{\text{bin}}}\}$, maximum number of regions R_P

Ensure: Final partitioning $\mathbf{r}$, normalized mutual information metric M_P

```

1: Initialize region set  $\mathbf{r} \leftarrow \{I\}$ 
2: Compute global entropy  $H_I$  of image using histogram  $\mathbf{h}$ 
3: while  $||\mathbf{r}|| < R_P$  do
4:   Sort regions  $\mathbf{r}$  in descending order of  $H(\mathbf{h} \mid r_i)$ 
5:   Select region  $r_m = \arg \max_{r_i \in \mathbf{r}} H(\mathbf{h} \mid r_i)$ 
6:   Initialize best conditional entropy  $H_{\min} \leftarrow \infty$ 
7:   for all valid binary splits  $(r_\alpha, r_\beta)$  of  $r_m$  do
8:     Compute  $H(\mathbf{h} \mid r_\alpha)$  and  $H(\mathbf{h} \mid r_\beta)$  with (3.27)
9:     Compute  $H_{\text{split}} = H(\mathbf{h} \mid [r_\alpha, r_\beta])$  with (3.28)
10:    if  $H_{\text{split}} < H_{\min}$  then
11:       $H_{\min} \leftarrow H_{\text{split}}$ 
12:       $(\bar{r}_\alpha, \bar{r}_\beta) \leftarrow (r_\alpha, r_\beta)$ 
13:    end if
14:  end for
15:  Replace  $r_m$  in  $\mathbf{r}$  with  $\bar{r}_\alpha$  and  $\bar{r}_\beta$ 
16: end while
17: Compute total mutual information  $I(\mathbf{h}; \mathbf{r})$ 
18: Compute  $M_P = \frac{I(\mathbf{h}; \mathbf{r})}{H_I}$ 
19: return  $\mathbf{r}, M_P$ 

```

Metrics Complexity Analysis

The complexity of the InfoAes metrics has been evaluated in terms of the number of multiplications required to compute them.

Let's first define the following notation: n_χ is the color depth of each color channel, R_G is the number of regions of the regular grid used to compute M_J and M_{NCD} (each region have the same number of pixels N_G), R_P is the threshold in terms of number of regions considered as the stop criterion for the mutual-information based partitioning algorithm, and $C(M)$ is the complexity of the metric M in terms of number of multiplication required to compute it. Thus, the expression of the complexity of each metric is reported in the following:

- $C(M_H) = 3(2^{3n_\chi}) + 1$
- $C(M_K) = 3 + 11N$
- $C(M_J) = R_G(1 + 2^{n_\chi}) + 2^{n_\chi}$
- $C(M_{NCD}) = 1 + \frac{R_G(R_G-1)}{2}(22N_G + 1) + 11R_GN_G$
- $C(M_{P,RGB}) = 1 + 2^{3n_\chi} + R_P(2 + 2^{3n_\chi})$
- $C(M_{P,L}) = 1 + 2^{n_\chi} + R_P(2 + 2^{n_\chi})$.

In this work, the following values have been assigned to the variables: $n_\chi = 8$, $R_G = 16$, $R_P = 80$, $N = 768 \times 768$ and $N_G = 192 \times 192$. The order of magnitude (OOM) of the complexity of InfoAes metrics is reported in Table 3.4. The OOM is shown on a base-10 logarithmic scale.

By analyzing the complexity of the metrics, the following comments can be derived:

- M_H has a quite high complexity because the entropy H_I is computed for the RGB representation of the image, *i.e.*, taking all three color channels into account.
- M_K has an intermediate complexity, mainly due to the complexity of the JPEG compression algorithm applied to the images, which depends on the number of pixel N of the image.
- M_J has the lowest complexity since it is applied to the Y_{709} luminance representation of the image, thus all the entropies required to compute the Jensen-Shannon divergence are computed only for one channel.
- M_{NCD} has the second-highest complexity, since to compute the metric the JPEG compression algorithm has to be applied to all image regions in the regular grid and to all possible pairs of regions.

- $M_{P,RGB}$ is the metric with the highest complexity, since the partitioning algorithm is applied to an RGB image and thus all the entropies are computed for three color channels.
- $M_{P,L}$ has a quite low complexity since the partitioning algorithm is applied only to the Y_{709} luminance channel.

Table 3.4: Order of magnitude (OOM) of the complexity of each InfoAes metric. The complexity is reported in base-10 logarithmic scale.

Metric	Complexity OOM ($\log_{10} scale$)
M_H	7
M_K	6
M_J	3
M_{NCD}	8
$M_{P,RGB}$	10
$M_{P,L}$	4

3.5.2 Analysis Methods to Evaluate Informational Aesthetics Metrics

In order to evaluate the effectiveness of the described metrics in distinguishing AI and human artworks, two different approaches were adopted:

- A statistical analysis was conducted on the output values (raw data) produced by all the metrics across a dataset comprising both AI-generated and human-created images. The objective of this analysis was to assess whether the metrics, considered individually or grouped into sets, reveal the presence of two distinct clusters, potentially corresponding to the two image sources and indicating systematic differences in visual structure or aesthetic properties.
- A model based on a trained classifier was developed to classify AI-generated and human-created images.

Raw Data Analysis

Gaussian fitting of the raw data Each metric was applied to all images in the dataset, resulting in two distinct vectors of values: one correspond-

ing to human-created images and the other to AI-generated images. These vectors were used to fit two Gaussian curves, one per class, by optimizing the parameters (mean, standard deviation, and amplitude) to best approximate the empirical distributions. Once fitted, these Gaussian models were used to classify all data points: for each image, we evaluated the likelihood under both curves and assigned it to the class whose Gaussian yielded the higher probability density at that point. This procedure allowed us to assess whether the metric could effectively separate the two image types based on their statistical profiles.

K-means clustering K-means clustering is a data partitioning algorithm that assigns each element from a set of o data points to one of χ predefined clusters. The algorithm aims to minimize the sum of distances between data points and the centroids of their assigned clusters, typically using the Euclidean distance. Clusters are defined by their centroids, and the number of clusters χ must be specified before the algorithm begins [135]. In general, the initial positions of the χ centroids are chosen randomly. In our work, we employed the k-means++ algorithm to initialize the centroids, as it improves convergence efficiency and reduces the likelihood of poor cluster selection [136]. Once the initial centroids are selected, each data point is assigned to the nearest centroid. The centroid of each cluster is then updated by computing the mean of all data points assigned to it. This process is repeated iteratively until the centroids converge or a predefined number of iterations is reached.

Trained Classifier

While statistical analysis of raw metric distributions can reveal trends and separability between human and AI-generated artworks, it does not provide a predictive model. In contrast, classification algorithms learn decision boundaries from labeled data, enabling automated and scalable authorship detection.

To identify the classifier that yielded the best performance on our dataset, we compared several widely used machine learning algorithms. The following groups of classifiers were considered:

- Decision Trees
- Logistic Regression

- Discriminant Analysis
- K-Nearest Neighbors (KNN)
- Naive Bayes classifiers
- Support Vector Machines (SVM)

The comparison revealed that the Fine Gaussian SVM achieved the highest classification performance. Detailed results of this comparison are reported in Sec.3.5.4.

SVM is a supervised machine learning algorithm commonly used for binary classification tasks. It operates by identifying an optimal hyperplane that maximizes the margin between two classes in an Ψ -dimensional feature space. The data points closest to the hyperplane are referred to as support vectors, and they play a critical role in defining the decision boundary. The margin is the distance between the hyperplane and the nearest support vectors from each class. The optimal hyperplane is the one that maximizes this margin [137,138].

If the data are linearly separable, a linear SVM is employed. In cases where the data are not linearly separable, a non-linear SVM is used. Non-linear SVMs leverage kernel functions to project the data into a higher-dimensional space where linear separation becomes feasible [139].

There are three main kernel functions categories:

- **Polynomial Kernel:** exploit a polynomial equation to map data points. The expression of polynomial kernel is:

$$\kappa(\mathbf{u}, \mathbf{v}) = (\mathbf{u}\mathbf{v} + c)^p, \quad (3.30)$$

where c is a constant and p is the polynomial degree.

- **Sigmoid Kernel:** it is based on hyperbolic tangent function. The kernel expression is:

$$\kappa(\mathbf{u}, \mathbf{v}) = \tanh(\gamma \mathbf{u}^T \mathbf{v} + r), \quad (3.31)$$

where r is a constant and γ is a scaling parameter.

- **Radial Basis Function (RBF)** Identifies data patterns by assessing the distance between points, making it particularly effective for circular or spherical relationships. The kernel expression is:

$$\kappa(\mathbf{u}, \mathbf{v}) = e^{-\gamma \|\mathbf{u} - \mathbf{v}\|^2}, \quad (3.32)$$

where γ is a parameter that controls the influence of each training example.

The *Gaussian SVM* used in this work exploits a Gaussian kernel which is a special case of RBF kernel. The expression of a Gaussian kernel is:

$$\kappa(\mathbf{u}, \mathbf{v}) = e^{-\frac{\|\mathbf{u}-\mathbf{v}\|^2}{2\sigma^2}}, \quad (3.33)$$

where σ is a parameter that controls the spread of the kernel function. In particular we used a *Fine* Gaussian SVM since the value of σ is small, so the resulting model is more sensitive to local variations in the data.

3.5.3 Experimental setup and tests

The used experimental setup consists of a digital image dataset composed of paintings created by human artists and corresponding paintings generated by an image-generation AI. In particular, our dataset includes 994 human paintings and 994 AI generated paintings.

The painting dataset was carefully composed in order to obtain a fair comparison between AI and human images. The procedure and the criteria adopted to select the human paintings and to generate the AI paintings are described in this section.

The metrics described in Section 3.5.1 have been implemented in MATLAB R2024b and computed for each painting in the considered dataset. To compute M_J and M_{NCD} , a regular grid of $R_G = 4 \times 4$ regions has been considered. To compute $M_{P,RGB}$ and $M_{P,L}$ a threshold $R_P = 80$ number of regions was set as a stop criterion for the partitioning algorithm.

After assembling the dataset, the metrics detailed in Sec. 3.5.1 for each image have been computed. Then created a new dataset in which each painting is represented by its associated set of InfoAes metrics has been created. In addition to these metrics, several categorical variables were considered: the artist's name, the genre of the painting, the artistic style, and a label indicating whether the image originated from a human or an AI source.

The following tests on the described dataset were carried on:

- An initial analysis of the raw data associated with each metric was conducted. Based on the Gaussian fitting curves derived from the human and AI values of each metric, an attempt was made to estimate the corresponding label (AI or Human).

- A second analysis of the raw data was performed by applying the K-means clustering algorithm in two distinct experiments. In the first experiment, the metrics were sorted in descending order according to their performance in the previous analysis. Starting from the top two metrics, five metric sets were defined by progressively adding one additional metric to each subsequent set. The K-means clustering algorithm was then applied to each set. In the second experiment, the metrics were sorted in ascending order based on their complexity. Again, starting from the top two metrics, five metric sets were constructed by incrementally including one additional metric per set, and the K-means clustering algorithm was applied to each configuration.
- The Fine Gaussian SVM classifier was applied to the same metric sets defined in the K-means clustering experiments. For each set, the classifier was trained using 90% of the dataset, selected via the HoldOut algorithm, [140], while the remaining 10% was reserved for testing. A K-fold algorithm was employed on the training data to validate the model and record training performance. This procedure was repeated for each metric across 100 randomly selected training-test intervals. The results were evaluated in terms of the mean and variance of precision, recall, F1-score, and accuracy, computed over all intervals.

Image Generation AI Model

In the presented work, employed Stable Diffusion (SD) was employed as the image generation AI model, as it is among the most widely adopted models in the literature.

Stable Diffusion is an image-generation AI model based on Latent Diffusion Models (LDM) [114]. Essentially, the training process of an LDM involves converting an image into completely random noise by incrementally adding gaussian noise through multiple steps. During training, a *noise predictor* is trained to estimate the noise added at each step. Specifically, the noise predictor consists of a fully connected convolutional neural network called *U-Net*. Image generation reverses the described process. To generate an image, the first step is to create an image consisting only of noise. The noise is sampled starting from a seed and it is then progressively removed by U-Net until the final image is formed. This process occurs in a predefined number of steps. Since diffusion is a demanding process in terms of computational resources, LDM does not operate in pixel space but in the *latent*

space, which is smaller than pixel space. To operate in the latent space, LDM uses a Variational Autoencoder (VAE) [141]. The VAE is a neural network composed of two parts: an encoder trained to convert images into latent space and a decoder trained to reconstruct the original image.

In particular, the version 2.1 of the main model was used, *i.e.*, *v2-1_768-ema-pruned.ckpt* (checkpoint). SD uses OpenCLIP (Contrastive Language-Image Pre-training), an open-source version of CLIP developed by OpenAI [142], as a text encoder to guide the generation through textual prompts. We used SD through the Web UI provided by Automatic1111 on GitHub [143].

SD 2.1 is designed to natively generate images with a size of 768×768 pixels. Although it is possible to specify a custom image size, setting a size that deviate from the native size may cause noticeable variations in the InfAes metrics values. This aspect was verified by generating images of different sizes, ranging from 256×256 to 1024×1024 .

For each size, 100 images was generated using the same prompt but with different seeds, which enables the generation of distinct images despite using identical input. The experiment to verify that different image sizes affect the InfAes metrics consisted in two distinct prompts, each one inspired by a different Van Gogh painting: a well-known piece, “The Starry Night” (TSN), and a lesser-known one, “Red Vineyards at Arles” (RVA). For each batch of 100 images of a given size, each InfAes metric was computed for all images, followed by the calculation of the mean value for each metric. The detailed results of the described preliminary experiments are shown in Appendix A.1.

An example of SD generated images with different sizes is shown in Fig. 3.28, where are reported SD generated images taken from the 256×256 , 768×768 and 2048×2048 batch of paintings for both the TSN and the RVA. Fig.s 3.28g and 3.28h show the original Van Gogh paintings. In the case of images generated with a size lower than the native 768×768 , the difference from the original image is particularly evident, since the AI generated image basically consists of regions of color that appear almost randomly distributed. This behavior decreases as the image size increases. For images larger than the native, SD fills the extra pixels either by duplicating elements already present in the painting or by introducing additional patterns.

Considering all the aforementioned issues, image generation was restricted to the native resolution of Stable Diffusion.

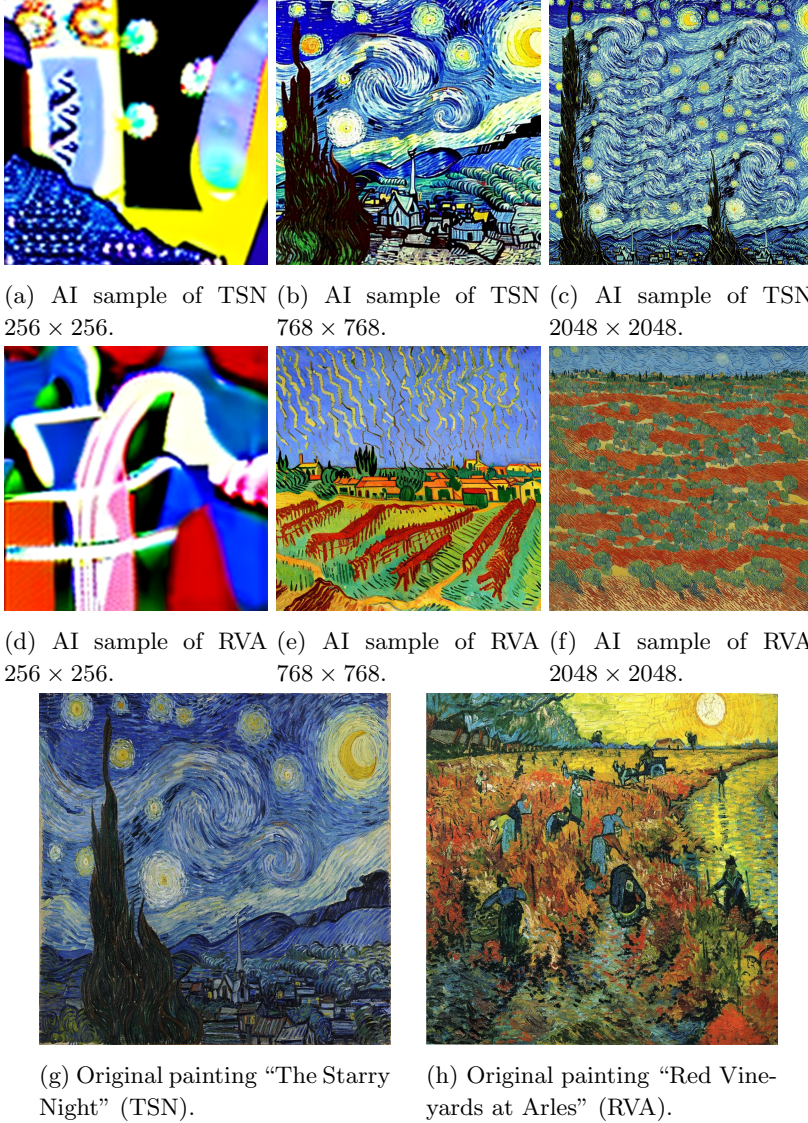


Figure 3.28: Effects of generating images at various sizes with AI Stable Diffusion model 2.1 and comparison with the original paintings.

Dataset Generation

The human paintings dataset was constructed starting from the WikiArt dataset, as provided in the work presented in [144]. The WikiArt dataset provided by the authors is composed of 81.444 images of paintings gathered from the WikiArt website [145]. The paintings in the dataset are organized by artistic style, which ranges from Early Renaissance art (XIV-XV century) to Contemporary Realism (XXI century). This work focuses on artistic styles that emerged from the eighteenth century onward, as these styles are characterized by a higher degree of personal creativity expressed by the artists [146].

The template of the prompt is designed to maximize the amount of information provided about the painting, without explicitly describing it. These are the information used by WikiArt to catalog the paintings. The prompt is composed of the following fields in the following order:

1. ***Artist's Name*** (*e.g* Eugene Boudin)
2. ***Painting Title*** (*e.g* A squall from North-west)
3. ***Artistic Style*** (*e.g* Impressionism)
4. ***Painting Genre*** (*e.g* Landscape)

The final prompt appears as follows:

Prompt Template
Artist Name / Painting Title / Artistic Style / Painting Genre

The information associated to each painting inserted in the prompt have been gathered from another dataset provided by Artificio on Hugging face [147]. This dataset is a labeled version of wikiart dataset, where the paintings are associated with labels that report: artist's name, painting title, artistic style, genre of the painting and date of the painting.

Some post-processing steps was applied to the AI-generated images to align them with their human counterparts. The first step involved resizing the AI-generated images to match the dimensions of the human-made ones. The images were resized to avoid that the lack in the number of pixels, at most equal to $\pm 20\%$ of 768×768 , could affect the InfAes metrics. The potential impact of image resizing on the metrics was also evaluated. Following the procedure adopted for image generation size, 100 images were

generated using the same prompt but different noise seeds, all initially sized at 768×768 . Each image was then resized to three different resolutions (256×256 , 512×512 , and 1024×1024). For each batch of 100 images, the InfAes metrics were computed for all images, and the mean value of each metric was subsequently calculated. The results of this pre-test are presented in Appendix A.1.

The second step involved the compression of the images. SD 2.1 outputs images in a PNG format but WikiArt images are all saved as JPEG, with different Quality Factors (Q_F). We first verified how compressing an image to JPEG format with different Q_F can affect the InfoAes metrics as follows. 100 images of size 768×768 was generated, using the same prompt but different noise seeds. All images were compressed using the JPEG format, with the compression quality factor (Q_F) progressively decreased from 100 to 5. For each resulting batch of 100 compressed images corresponding to a specific Q_F value, the InfAes metrics were computed for all images, followed by the calculation of the mean value for each metric. The results of this pre-test are presented in Appendix A.1.

In order to have a fair comparison we compressed the AI images to a JPEG format with the same Q_F as the human counterpart. We calculated the Q_F of the human images with the software ImageMagik [148]. The dataset containing both the AI-generated and human images used in this work is available at [149].

3.5.4 Results

This section describes and discusses the results of the experiments detailed in Sec. 3.5.3.

Results: Gaussian Fitting Of The Raw Data

Figures 3.29, 3.30 and 3.31 display the histograms and the corresponding Gaussian fitting curves for the GA metrics (M_H and M_K), the CA metrics (M_J and M_{NCD}), and the MIP metrics ($M_{P,RGB}$ and $M_{P,L}$), respectively.

Mean and standard deviation of each InfoAes metric is reported in Table 3.5 for both the AI and the human images dataset. The mean value of M_K for AI-generated artworks is slightly higher than that for human ones. This suggests that AI paintings, on average, demonstrate a slightly higher level of detail and global complexity. Additionally, the mean value of M_{NCD}

Table 3.5: Human vs. AI comparison in terms of mean and standard deviation of InfAes metrics.

	M_H		M_K		M_J		M_{NCD}		$M_{P,RGB}$		$M_{P,L}$	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Human	0.434	0.120	0.910	0.065	0.901	0.044	0.451	0.095	0.252	0.045	0.183	0.067
AI	0.421	0.063	0.934	0.053	0.880	0.042	0.537	0.102	0.280	0.047	0.226	0.069

is significantly higher for AI paintings compared to human ones, indicating that AI artworks tend to display a greater degree of structural self-similarity, *i.e.*, the regions of the AI images,located by the regular grid, show a more similar level of details and more similar patterns.

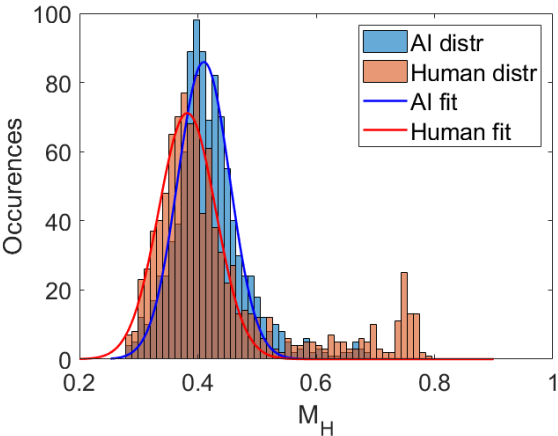
Table 3.6 shows the results of the estimation process, which leverages the relative amplitudes of the fitted Gaussian curves in order to classify the image: human or AI-generated. Although precision, recall, and F1-score are typically used in supervised classification, they are here employed as heuristic measures to evaluate how well the fitted Gaussian distributions can separate AI and human images. The results indicate that the two best-performing metrics are M_{NCD} and M_K , whereas the lowest-performing metric is $M_{P,RGB}$. This means that the AI and human paintings considered in our dataset exhibit more pronounced differences in terms of level of detail and global complexity (represented by M_K), as well as in structural self-similarity measured by M_{NCD} .

No significant differences emerges analyzing how much structured are the colors of the images (represented by $M_{P,RGB}$). However, a slightly larger difference appears analyzing the structure of the luminance distribution of the images (represented by $M_{P,L}$), which is the third best metric. The performances of M_H and M_J indicate that AI and Human paintings show, respectively, no significant differences in terms of richness of the color palette and of statistical self-similarity, *i.e.*, the AI images regions, located by the regular grid, show no differences in the distribution of the Y_{709} luminance compared to human images.

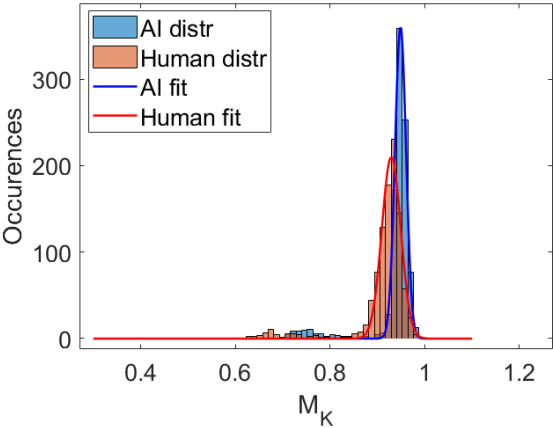
Results: K-means Clustering Of The Raw Data

Table 3.7 shows the results of the first K-means clustering experiment described in Sec. 3.5.3, while Table 3.8 shows the results of the second K-means clustering experiment described in Sec. 3.5.3.

An analysis of the results in Table 3.7 can lead us to the insight that pro-

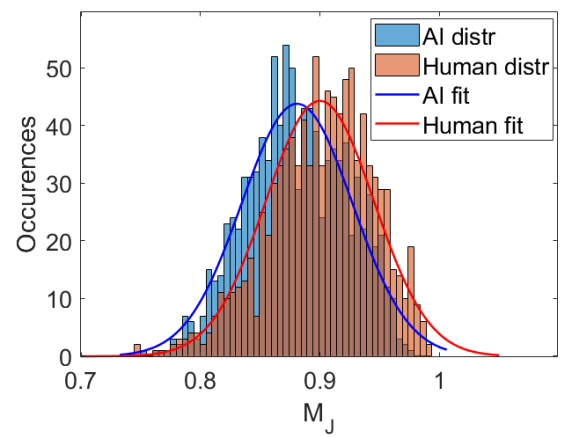


(a) M_H

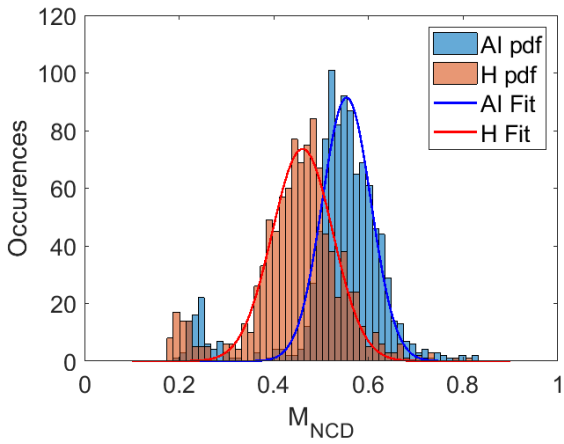


(b) M_K

Figure 3.29: Global Aesthetics metrics distribution and Gaussian fitting curves: AI vs. Human comparison.



(a) M_J



(b) M_{NCD}

Figure 3.30: Compositional Aesthetics metrics distribution and Gaussian fitting curves: AI vs. Human comparison.

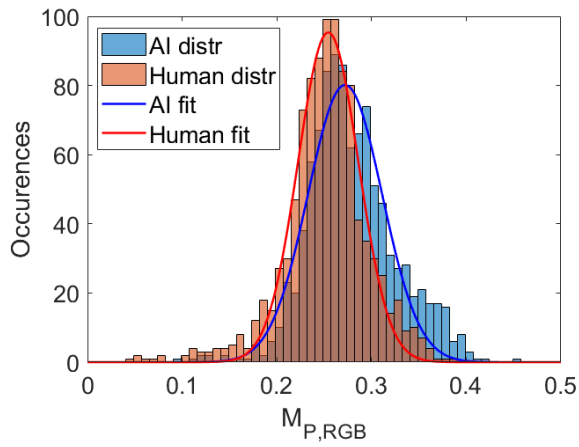
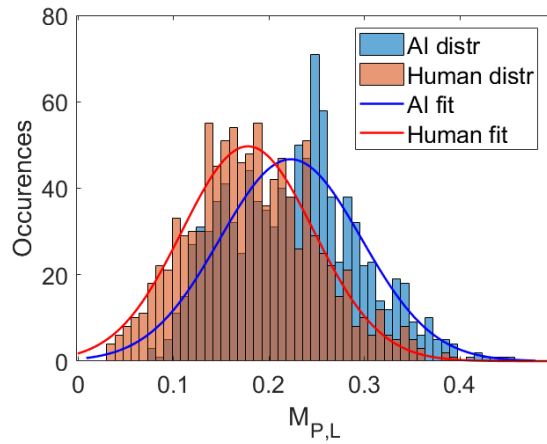
(a) $M_{P,RGB}$ (b) $M_{P,L}$

Figure 3.31: Mutual Information-based Partitioning metrics distribution and gaussian fitting: AI vs. Human comparison.

Table 3.6: Results of label estimation exploiting a Gaussian fitting of the histograms of the metrics.

Metric	Precision	Recall	F1-score	Accuracy
M_H	0.591	0.758	0.665	0.617
M_K	0.715	0.794	0.752	0.739
M_J	0.608	0.580	0.594	0.603
M_{NCD}	0.773	0.812	0.792	0.787
$M_{P,RGB}$	0.610	0.512	0.557	0.592
$M_{P,L}$	0.628	0.606	0.617	0.624

Table 3.7: Results of K-means clustering considering metrics set composed with increasing histogram-based label estimation accuracy metrics.

Metrics set	Precision	Recall	F1-score	Accuracy	Complexity
M_K, M_{NCD}	0.835	0.611	0.704	0.745	14
$M_K, M_{NCD}, M_{P,L}$	0.711	0.684	0.697	0.703	18
$M_K, M_{NCD}, M_{P,L}, M_H$	0.698	0.685	0.692	0.695	25
$M_K, M_{NCD}, M_{P,L}, M_H, M_J$	0.687	0.688	0.688	0.688	28
$M_K, M_{NCD}, M_{P,L}, M_H, M_J, M_{P,RGB}$	0.657	0.728	0.691	0.674	38

gressively adding metrics that performed more poorly (see Table 3.6) tends to deteriorate K-means clustering performance. The first row of the table confirms that M_K and M_{NCD} are the best metrics and together provide the best performance for the first experiment.

An analysis of the results in Table 3.8 reveals that progressively adding metrics with higher complexity (see Table 3.6) does not directly affect K-means clustering performance. Notably, the greatest improvements are observed when M_K and M_{NCD} are included in the metric set (corresponding to the second and fifth rows of the table). Conversely, the largest increase in complexity, caused by the inclusion of $M_{P,RGB}$ in the dataset (see last row of the table) results in a deterioration of clustering performance.

Results: Comparison of Classifiers

The classifiers listed in Sec. 3.5.2 have been trained using all the InfoAes metrics. A training set composed by 90% of the entire dataset has been considered, while reserving the remaining 10% of the data as test set. The

Table 3.8: Results of K-means clustering considering metrics set composed with increasing complexity metrics.

Metrics set	Precision	Recall	F1-score	Accuracy	Complexity
$M_J, M_{P,L}$	0.609	0.597	0.603	0.607	7
$M_J, M_{P,L}, M_K$	0.626	0.614	0.620	0.623	13
$M_J, M_{P,L}, M_K, M_H$	0.627	0.619	0.623	0.625	20
$M_J, M_{P,L}, M_K, M_H, M_{NCD}$	0.687	0.688	0.688	0.688	28
$M_J, M_{P,L}, M_K, M_H, M_{NCD}, M_{P,RGB}$	0.657	0.728	0.691	0.674	38

Table 3.9: Performance of the classifiers.

Classifier	Precision		Recall		F1-score		Accuracy	
	μ	σ	μ	σ	μ	σ	μ	σ
Fine Tree	0.832	0.027	0.766	0.038	0.797	0.030	0.807	0.026
Binary GLM Logistic Regression	0.775	0.027	0.748	0.037	0.761	0.032	0.768	0.027
Linear discriminant analysis	0.774	0.040	0.750	0.038	0.766	0.033	0.771	0.028
Weighted KNN	0.888	0.034	0.755	0.041	0.815	0.030	0.831	0.025
Kernel Naive Bayes	0.777	0.041	0.716	0.043	0.744	0.035	0.756	0.029
Fine Gaussian SVM	0.875	0.036	0.797	0.039	0.834	0.029	0.843	0.025

described process has been repeated across 100 different training-test different intervals over the dataset and we evaluated the results in terms of mean (μ) and standard deviation (σ) of the precision, recall, F1-score and accuracy.

Table 3.9 shows the performance of the classifiers. The Fine Gaussian SVM classifier [137] is the most effective model for distinguishing AI-generated from human-made artworks using the InfAes metrics. It achieves the highest accuracy at 84.3%, with a strong balance between precision (0.875) and recall (0.797), resulting in a robust F1-score of 0.834. These values indicate that the classifier not only makes highly reliable predictions (high precision) but also successfully identifies a large proportion of human-created images (high recall). Its low standard deviations further confirm its consistency and stability across different training-test splits, making it the most reliable choice among the tested models.

Fig. 3.32 shows the confusion matrix obtained by the Fine Gaussian SVM in the testing phase. Each value in the confusion matrix represents the rounded average across 100 randomized training-test splits, providing a stable estimate of classifier behavior.



Figure 3.32: Classification of Human/AI artworks: confusion matrix obtained with the Fine Gaussian SVM.

Results: Fine Gaussian SVM Classifier Performance

Table 3.10 shows the results of the first experiment described in Sec. 3.5.3, while Table 3.11 shows the results of the second experiment described in Sec. 3.5.3.

In Table 3.10, the combination of M_K , M_{NCD} , and $M_{P,L}$ yields a strong F1-score of 0.813, with precision at 0.860 and recall at 0.773. This suggests the classifier is highly confident in its predictions (high precision) while still capturing a large portion of actual human images (good recall). As more metrics are added, the F1-score improves slightly, reaching 0.834 with all six metrics, indicating a well-balanced model. In Table 3.11, the initial metric set (M_J , $M_{P,L}$) performs poorly, with an F1-score of 0.586, showing that these metrics alone are insufficient. However, adding M_K and M_{NCD} significantly boosts performance, confirming their central role.

Further analyzing Table 3.10 and 3.11 we can notice that, respectively:

- Progressively adding metrics that performed poorly (see Table 3.6) tends to improve the performance of the classification. However, since the two best performing metrics, M_K and M_{NCD} , are included since

the beginning, the best improvement is given by the inclusion of $M_{P,L}$ (see third row of the table) which is the third best metric.

- Progressively adding metrics with a higher complexity tends to improve the performances of the classifier. However also here emerges how M_K and M_{NCD} are the most relevant metrics to the performance of the classifier. Their inclusion on the metrics set produces the highest accuracy increase (see second and fourth row of the table).

In both experiments, the set that includes all the metrics yields the best result (see the last row of Tables 3.10 and 3.11), achieving an accuracy of 84.3 %.

Table 3.10: Results of Fine Gaussian SVM classifier considering metrics set composed with increasing histogram-based label estimation accuracy metrics

Metrics set	Precision		Recall		F1-score		Accuracy		Compl.
	μ	σ	μ	σ	μ	σ	μ	σ	
M_K, M_{NCD}	0.817	0.042	0.781	0.040	0.798	0.029	0.804	0.026	14
$M_K, M_{NCD}, M_{P,L}$	0.860	0.035	0.773	0.040	0.813	0.028	0.824	0.025	18
$M_K, M_{NCD}, M_{P,L}, M_H$	0.851	0.041	0.792	0.039	0.820	0.031	0.827	0.028	25
$M_K, M_{NCD}, M_{P,L}, M_H, M_J$	0.862	0.040	0.787	0.038	0.822	0.030	0.832	0.026	28
$M_K, M_{NCD}, M_{P,L}, M_H, M_J, M_{P,RGB}$	0.875	0.036	0.797	0.039	0.834	0.029	0.843	0.025	38

Table 3.11: Results of Fine Gaussian SVM classifier considering metrics set composed with increasing complexity metrics

Metrics set	Precision		Recall		F1-score		Accuracy		Compl.
	μ	σ	μ	σ	μ	σ	μ	σ	
$M_J, M_{P,L}$	0.596	0.053	0.580	0.053	0.586	0.036	0.593	0.034	7
$M_J, M_{P,L}, M_K$	0.786	0.041	0.752	0.044	0.768	0.030	0.775	0.026	13
$M_J, M_{P,L}, M_K, M_H$	0.807	0.046	0.770	0.042	0.787	0.030	0.794	0.027	20
$M_J, M_{P,L}, M_K, M_H, M_{NCD}$	0.862	0.040	0.787	0.038	0.822	0.030	0.832	0.026	28
$M_J, M_{P,L}, M_K, M_H, M_{NCD}, M_{P,RGB}$	0.875	0.036	0.797	0.039	0.834	0.029	0.843	0.025	38

To summarize the results discussed above, Figs. 3.33a and 3.33b illustrate the accuracy of the Fine Gaussian SVM classifier as a function of the metric sets used in the first and second experiments, respectively. Fig. 3.33a shows that the overall variation in accuracy when adding new metrics is relatively small, as the two most effective metrics, M_K and M_{NCD} , are already

Table 3.12: Comparison of the best classifier and K-mean clustering with literature image classifier. In round brackets we reported the metric set used for our K-means clustering and Fine Gaussian SVM classifier.

Classifier	Precision	Recall	F1-score	Accuracy
K-means clustering	0.835	0.611	0.704	0.745
Fine Gaussian SVM	0.875	0.797	0.834	0.843
Swin Transformer Image Classifier	0.606	0.961	0.743	0.668

included in the initial set. This confirms that accuracy is strongly influenced by the quality of the selected metrics. Fig. 3.33b demonstrates that increasing complexity by adding additional metrics consistently improves classifier performance, although the improvement is not always proportional to the increase in complexity.

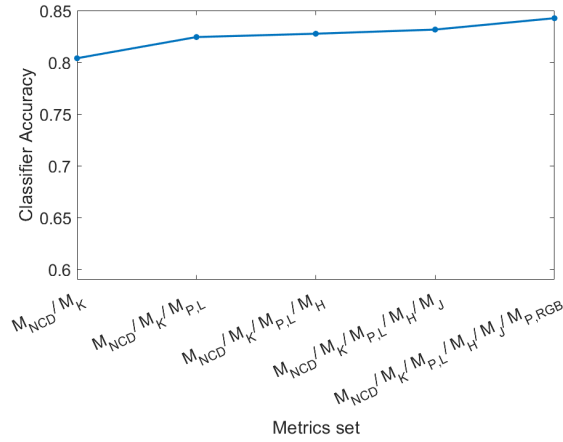
3.5.5 Comparison with Classifiers in Literature

To contextualize the performance of the proposed framework, we compare it against a state-of-the-art deep learning model commonly used for image classification: the Swin Transformer [116]. We used the Swin Transformer to classify all the AI and human images in our curated dataset and we compared the results with those produced with our best performing K-mean clustering and Fine Gaussian SVM classifiers. Table 3.12 show the results of such a comparison, while Fig. 3.34 shows the confusion matrix obtained by applying the Swin Transformer image classifier to our dataset.

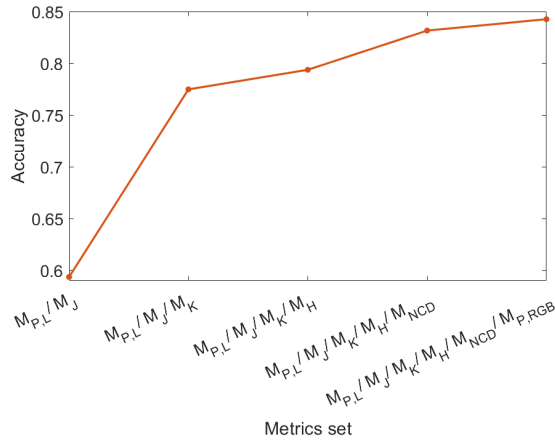
An analysis of Table 3.12 shows how both our K-means clustering and Fine Gaussian SVM classifier outperform the Swin Transformer Classifier. The main problem of the Swin Transformer image classifier is that it tends to classify the majority of images as human. So it almost always correctly classify human images and on the other hand it misclassify the majority of AI images.

3.5.6 Discussion

The results of the conducted experiments highlight distinct behaviors between clustering and classification approaches when applied to the task of distinguishing AI-generated and human-made artworks using Informational



(a) Accuracy of Fine Gaussian SVM classifier in the first experiment



(b) Accuracy of Fine Gaussian SVM classifier in the second experiment

Figure 3.33: Fine Gaussian SVM accuracy across the metric sets defined for the two experiments.

Aesthetics (InfAes) metrics.

K-means clustering, while conceptually simple and unsupervised, does

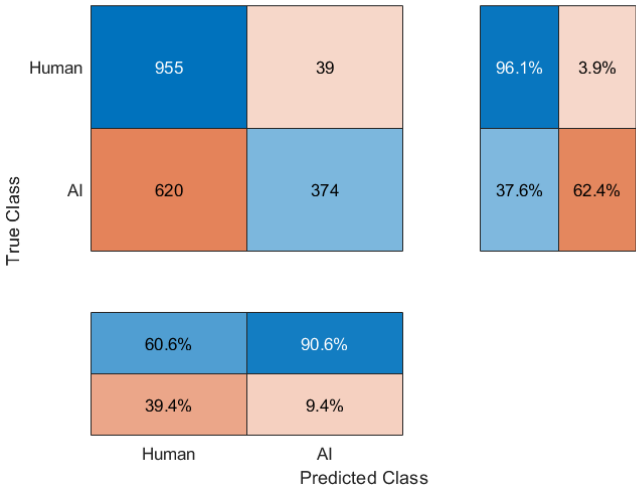


Figure 3.34: Classification of Human/AI artworks: confusion matrix obtained with the literature image classifier.

not yield particularly strong performance. The clustering accuracy peaks at 74.5% when using only the two most discriminative metrics, M_K and M_{NCD} , and tends to degrade as more metrics are added. This suggests that metric quality is more important than quantity, and that including less informative metrics introduces noise that hinders cluster separation. Moreover, the increased computational complexity associated with larger metric sets does not yield better clustering results. In fact, the best clustering outcome is achieved with the smallest and most effective metric subset. Given this, a practical alternative could be to directly use the histogram of M_{NCD} for authorship estimation, bypassing clustering altogether.

In contrast, the Fine Gaussian SVM classifier demonstrates a consistent improvement in performance as more metrics are included. Each new metric contributes complementary information that the classifier can exploit, leading to a steady increase in accuracy. The highest classification performance (84.3%) is achieved using all six InfAes metrics, despite this set has the highest computational complexity.

When considering both performance and computational cost, the metric set composed by M_K , M_{NCD} , and $M_{P,L}$ offers the best trade-off, achiev-

ing high classification accuracy (82.43%) while maintaining moderate complexity. This set of metrics captures the global complexity, the structural self-similarity, and the luminance-based spatial organization, three complementary aspects of visual information that are particularly effective in distinguishing human artistic intent from algorithmic generation.

3.5.7 Conclusion

This study presented an interpretable and training-light framework for distinguishing AI-generated artworks from human-made ones, based on the theory of Informational Aesthetics. By modeling the creative process as an information channel, an “aesthetic gap” that current generative models fail to bridge was identified. The analysis of 994 human and AI-generated images using six handcrafted InfAes metrics reveals that global complexity, structural self-similarity, and spatial coherence provide robust signals for authorship detection. The Fine Gaussian SVM classifier implemented in this work, achieves superior performance compared to a state-of-the-art deep learning model, confirming the effectiveness of interpretable metrics over black-box approaches. These findings underscore the limitations of current generative models in replicating human artistic intent and provide a principled path toward understanding the informational uniqueness of human creativity. Future work will explore the generalizability of this framework across different generative architectures and investigate more direct measures of emotional content in visual art.

Chapter 4

Conclusion

This chapter summarizes the contributions of the thesis and discusses avenues for future research.

4.1 Summary of contribution

The vision outlined in the IMT-2030 framework for sixth-generation (6G) mobile networks calls for a transformative shift: from traditional connectivity to a pervasive digital fabric in which the human body becomes an active node. This paradigm introduces two fundamental challenges, as discussed in the Introduction: achieving hyper-reliable low-latency communications (HRLLC) and securely integrating human-centric data from resource-constrained devices. Throughout this thesis, these challenges have been addressed through the strategic application of Physical Layer Security (PLS) techniques.

To meet the stringent latency requirements of HRLLC, the thesis proposed and evaluated Noise Loop (NL) modulation as a lightweight confidentiality mechanism. The technique was adapted to two key transmission technologies:

- All-Optical Processing (AOP)-based Optical Communication via Fiber (OCvF): A novel Optical Noise Loop (ONL) modulation was developed and theoretically validated, demonstrating full secrecy rate and resilience against eavesdropping, even under worst-case scenarios.
- Terahertz (THz) Communications: The effects of THz-specific noise

sources (thermal, molecular, and phase noise) were analyzed, revealing how phase noise impacts reliability and how loop gain can mitigate its effects.

To support the integration of the human body as a network node, the thesis explored Wireless Body Area Networks (WBANs) as a foundational architectural layer. Within this framework, the thesis investigated three emerging transmission paradigms, each offering distinct advantages for secure, low-power, and context-aware data exchange in human-centric environments:

- **Molecular Communications (MCs):** A threat modeling framework was developed for MC via Diffusion (MCvD), mapping classical attacks (jamming, eavesdropping, MITM, spoofing) and analyzing their complexity across modulation schemes.
- **Galvanic Coupling (GC):** A biometric authentication protocol was designed using raw GC signals and a Long Short-Term Memory (LSTM) neural network. The protocol effectively counters both Over-the-Air and Direct Contact attacks, with experimental validation showing high reliability and security.
- **Visible Light Communications (VLC):** A channel model for in-to-out body transmission was developed for both single LED and 2×2 MIMO configurations. The MIMO setup demonstrated tradeoff between Bit Error Rate (BER) and secure area, confirming its suitability for secure IMD applications.

Finally, the thesis addressed a novel challenge introduced by generative AI: distinguishing synthetic outputs from genuine human outputs in WBANs. Treating artworks as expressive signals of the human body, a framework based on Informational Aesthetics (InfAes) was developed. Six handcrafted metrics were used to quantify structural and aesthetic properties of images. A Fine Gaussian SVM classifier trained on these metrics outperformed a state-of-the-art deep learning model, achieving 84.26% accuracy in authorship detection. This result highlights the limitations of AI in replicating human intent and supports the need for authentication mechanisms in human-as-a-node systems.

In summary, the thesis demonstrates that PLS offers a coherent and effective response to the dual challenge posed by 6G: securing ultra-fast

communications and safeguarding human-centric data. By operating directly at the physical layer, PLS techniques provide lightweight, adaptive, and robust security solutions that align with the architectural and operational constraints of future networks.

4.2 Directions for future work

Several avenues for future research emerge from the present study, offering opportunities to extend and refine the proposed solutions. In particular, the following directions are suggested for advancing Physical Layer Security (PLS) strategies in the context of 6G novel transmission technologies.

Given the potential impact of security vulnerabilities in MCs, future research should focus on developing countermeasures tailored to the molecular domain. This includes the design of physical layer authentication techniques to mitigate spoofing attacks and physical layer confidentiality techniques in order to effectively counter eavesdropping threats.

Building on the theoretical analysis and simulation results, the proposed hardware implementation for ONL should be realized in practice in order to collect experimental data to corroborate the conclusions drawn. Additionally, the adoption of the AOP-based NL modulation should be considered also for OWC systems.

Future work on GC-based biometric PLA should aim to expand the experimental dataset to include a larger and more diverse subject population. This would improve generalization and reduce bias. Additional research should investigate the integration of multi-modal biometric features and the use of adaptive thresholding techniques to dynamically balance security and usability.

Concerning the research conducted about the authentication of the outputs of the human body, a potential future direction for the research is the evaluation of alternative models for Stable Diffusion alongside different generative AI algorithms beyond Stable Diffusion, in order to assess the generalizability of the proposed framework across different generative architectures and training paradigms. Also, the informational aesthetic metrics exploited in this paper to quantify the “aesthetic gap” between AI and human artworks represent a quite indirect measure of the investigated human factor. Therefore, a more direct measure of the emotional content expressed by an artwork is needed. Finally, a compelling research direction is the extension

of the developed framework to other types of output signals produced by human body in WBANs.

The physical layer solutions proposed in thesis, supported by theoretical analysis and simulation results, lay the groundwork for future experimental validation and practical deployment. However, the rapid evolution of communication paradigms and the increasing relevance of human-centric networking call for continuous innovation and interdisciplinary collaboration.

These future directions aim to consolidate the role of PLS as a foundational element in 6G architectures. By continuing to explore lightweight, interpretable, and biologically compatible security mechanisms, this research contributes to building networks that are not only faster and smarter, but also trustworthy, inclusive, and human-aware.

Appendix A

A.1 Results of the pre-tests described in Sec. 3.5

In this appendix the results of the three pre-tests in 3.5.3. In summary, this appendix shows the impact on the InfAes metrics of image size generation, resizing, and compression.

A.1.1 Impact of generating the image with a size different from the native one

In this appendix the results of the pre-tests regarding the generation of images with SD at different sizes are presented and discussed. The results are showed in Fig. A.1. In particular, Fig. A.1a shows that M_H tends to decrease as the size of the generated image increases. This indicates that SD 2.1 enriches the color palette as the generated image size increases. Fig. A.1b shows that M_K rapidly decreases as long as the size of the generated batch of images is lower than 512×768 , after which the decrease becomes much slower. This indicates that images generated with a size lower than 512×768 are compressed more by the JPEG algorithm. Fig. A.1c shows that M_J rapidly increases as long as the size of the generated image set is lower than 512×768 , after which the increase becomes much slower. This indicates that images generated with size lower than 512×768 presents a higher statistical self-similarity, *i.e.*, the image regions presents similar luminance probability distributions between each others. Fig. A.1d shows that M_{NCD} , similarly to M_K , rapidly decreases as long as the size of the generated batch of images is lower than 512×768 , after which the increase becomes much slower. This indicates that images generated with a size higher than 512×768 show a

lower structural self-similarity. Fig. A.1e and Fig. A.1f shows that $M_{P,RGB}$ and $M_{P,L}$ decreases as the size of the generated image increases, respectively. This indicates that the images becomes less structured and more complex as the size they are generated at increases, for both RGB and Y_{709} luminance representation of the image.

The graphs in Fig.,A.1 clearly illustrate that all metrics are influenced by the size at which the images are generated, for both the starryNight and redVineyards painting sets.

A.1.2 Impact of resizing the image

In this section the results of the pre-tests regarding resizing images of size 768×768 to different sizes, described in Sec.3.5.3 are reported. The results are showed in Fig. A.2.

Fig. A.2a shows that M_H decreases when the 768×768 images are up-scaled to a bigger size and increases when the images are down-scaled to a smaller size. This indicates that down-scaling an image reduces its color palette, whereas up-scaling tends to enrich it. Fig. A.2b shows that M_K increases when the 768×768 images are up-scaled to a bigger size and decreases when down-scaled to a lower size. This indicates that down-scaling an image reduces the compression that can be achieved by the JPEG compression algorithm while up-scaling it improves the compression that can be achieved by the JPEG compression algorithm. Fig. A.2c shows that M_J is not noticeably affected when the 768×768 images are up-scaled to a bigger size while it decreases when the images are down-scaled to a smaller size. This indicates that up-scaling an image doesn't affect its statistical self-similarity, *i.e.*, the similarity of the luminance distribution of the image regions remains almost unchanged. Instead, down-scaling the images tends to reduce the statistical self-similarity of the images, *i.e.*, the similarity of the luminance distribution of the image regions is reduced. Fig. A.2d shows that M_{NCD} rapidly increases when the 768×768 images are up-scaled to a bigger size. Down-scaling the images to a smaller size shows mixed effects on M_{NCD} . Down-scaling to 512×512 reduces M_{NCD} , *i.e.* the structural self-similarity of the images. However, down-scaling the size of the images to 256×256 increases M_{NCD} , *i.e.*, the structural self-similarity of the images. Fig. A.2e show that $M_{P,RGB}$ decreases when the 768×768 images are up-scaled to a bigger size and increases when they are down-scaled to a smaller size. This indicates that down-scaling images tends to decrease heterogene-

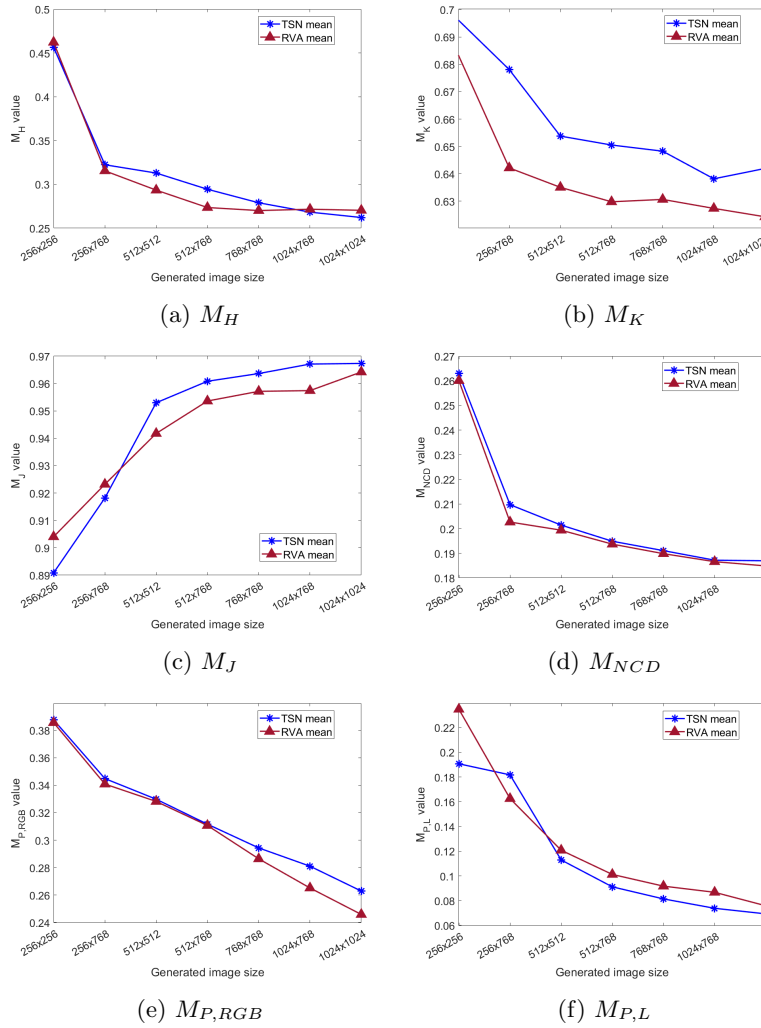


Figure A.1: Effects of generating images at various sizes on the InfAes metrics.

ity of the color distribution of the images making them more structured while up-scaling increases the heterogeneity of the color distribution of the images. Fig. A.2f show that $M_{P,L}$ increases when the 768×768 images are down-scaled to smaller sizes and has very little variations when they are up-scaled to bigger sizes. This indicates that, similarly to $M_{P,RGB}$ down-scaling tends to reduce the heterogeneity of the Y_{709} luminance distribution of the image and up-scaling doesn't affect it.

The graphs in Fig.,A.1 clearly demonstrate that all metrics are influenced by the resizing operation. In general, the extent of the variations depends on the magnitude of the up-scaling or down-scaling applied. However, when the resizing magnitude remains sufficiently limited, the resulting variations are nearly negligible.

A.1.3 Impact of JPEG compression with different Q_F

In this section the results of the pre-tests described in Sec. 3.5.3 regarding the compression of images generated at size 768×768 and JPEG compressed with different quality factor Q_F are reported.

Fig. A.3a shows that M_H remains unchanged as long as $Q_F \geq 30$, after which starts decreasing. This indicates that the color palette of the image remains unchanged, as long as $Q_F \geq 30$. Fig. A.3b show that M_K increases as the Q_F of the JPEG compression decreases. This indicates that when the Q_F decreases the images are more compressed by the JPEG algorithm. Fig. A.3c shows that M_J remains unchanged as long as $Q_F \geq 30$, after which it starts decreasing. This indicates that the statistical self-similarity of the image remains unaltered, as long as $Q_F \geq 30$. Fig. A.3d shows that M_{NCD} increases as the Q_F of the JPEG compression decreases. This indicates that when the Q_F decreases the regions and couples of regions of the image are more compressed by the JPEG algorithm.

Fig. A.3e shows that $M_{P,RGB}$ tends to slowly increase as Q_F decreases. When $Q_F \geq 30$ the artifacts introduced by the JPEG compression produce bigger variations on the values of $M_{P,RGB}$ and magnitude of the variation depends on the painting. This indicates that decreasing the Q_F of the JPEG compression tends to reduce the heterogeneity of the color distribution of the image which tends to become more structured, until $Q_F \geq 30$. Fig. A.3f shows that $M_{P,L}$ remains almost unchanged as long as $Q_F \geq 30$, after which it start to rapidly increase. This indicates that the Y_{709} luminance distribution of the images remains almost unaltered by the JPEG compression, until

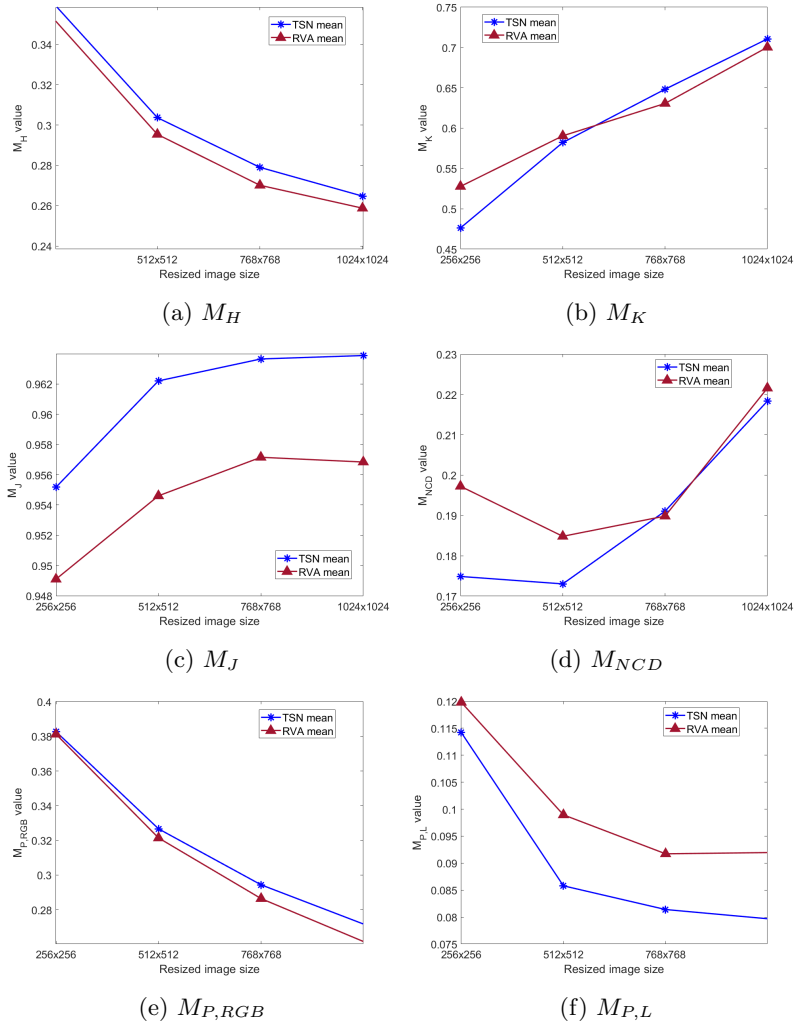


Figure A.2: Effects of resizing 768x768 images to different sizes on the InfAes metrics.

$Q_F \geq 30$. When $Q_F < 30$, the artifacts introduced by the JPEG compression produce bigger variations on the values of $M_{P,RGB}$ and the magnitude of the variation depends on the painting. This indicates that decreasing the Q_F of the JPEG compression tends to reduce the heterogeneity of the color distribution of the image which tends to become more structured, until $Q_F \geq 30$. When $Q_F < 30$ the compression artifacts introduced by the JPEG compression start to alter the heterogeneity of the Y_{709} luminance distribution of the image increasing it.

The graphs in Fig.,A.3 clearly indicate that the metrics most affected by variations in the compression quality factor (Q_F) are M_K and M_{NCD} , due to their strong dependence on this parameter. The remaining metrics exhibit negligible or limited sensitivity, provided that $Q_F \geq 30$.

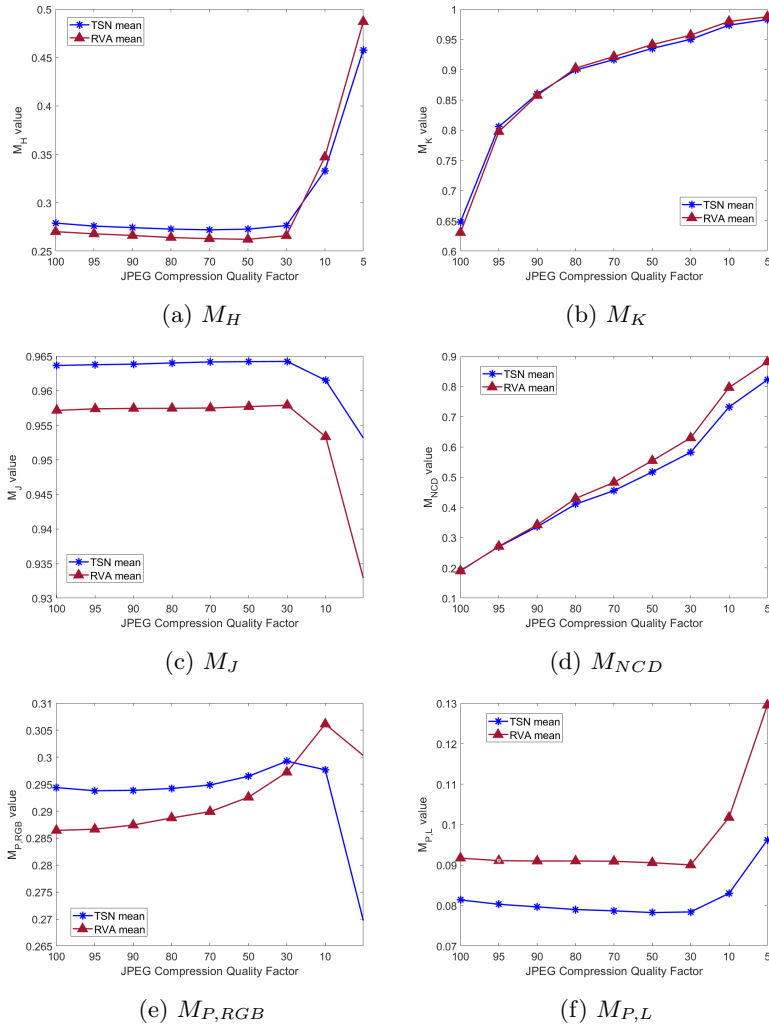


Figure A.3: Effects of compressing images to JPEG with different compression quality factors (Q_F) on the InfAes metrics.

Bibliography

- [1] S. Caputo, S. Viciani, S. Gherardini, G. Borghini, F. Cataliotti, and L. Mucchi, “Noise level modulation for secure optical communications,” *IEEE Access*, vol. 12, pp. 148 693–148 705, 2024. [Online]. Available: <http://dx.doi.org/10.1109/ACCESS.2024.3476520>
- [2] S. Caputo, G. Borghini, G. Pasolini, N. Longhi, G. Cuzzo, and L. Mucchi, “Physical layer security at thz communications: the effects of phase and molecular noise,” in *2024 IEEE International Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0 and IoT)*. IEEE, May 2024, pp. 304–309. [Online]. Available: <http://dx.doi.org/10.1109/MetroInd4.0IoT61288.2024.10584207>
- [3] S. Caputo, G. Borghini, S. Jayousi, and L. Mucchi, *Visible Light Communications for Implantable Medical Devices*. Springer Nature Switzerland, 2024, pp. 75–87. [Online]. Available: http://dx.doi.org/10.1007/978-3-031-72524-1_7
- [4] S. Caputo, A. Viziello, A. Coviello, M. Magarini, G. Borghini, P. Savazzi, S. Jayousi, and L. Mucchi, “Biometric authentication in galvanic coupling communications: A physical layer approach for secure wearable medical systems,” in *Proceedings of the 2025 EAI International Conference on Body Area Networks (BodyNets)*, September 2025, accepted.
- [5] G. Borghini, S. Caputo, S. Jayousi, M. Magarini, M. Pierobon, and L. Mucchi, “Security threats in diffusion-based molecular communication systems,” in *2025 19th International Symposium on Medical Information and Communication Technology (ISMICT)*. IEEE, May 2025, pp. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/ISMICT64722.2025.11059392>
- [6] G. Borghini, A. Enersen, S. Caputo, M. Pierobon, and L. Mucchi, “Measuring aesthetic information in ai-generated artistic images: A tool for authorship detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, October 2025, submitted.

- [7] D. N. R. ITU-R, "Framework and overall objectives of the future development of imt for 2030 and beyond," *Int. Telecommu. Union (ITU) Recommendation (ITU-R)*, 2023.
- [8] M. Series, "Imt vision–framework and overall objectives of the future development of imt for 2020 and beyond," *Recommendation ITU*, vol. 2083, no. 0, pp. 1–21, 2015.
- [9] R. Singh, A. Kaushik, W. Shin, M. D. Renzo, V. Sciancalepore, D. Lee, H. Sasaki, A. Shojaeifard, and O. A. Dobre, "Towards 6g evolution: Three enhancements, three innovations, and three major challenges," *IEEE Network*, pp. 1–1, 2025. [Online]. Available: <http://dx.doi.org/10.1109/MNET.2025.3574717>
- [10] D. Uzunidis, K. Moschopoulos, C. Papapavlou, K. Paximadis, D. M. Marom, M. Nazarathy, R. Munoz, and I. Tomkos, "A vision of 6th generation of fixed networks (f6g): Challenges and proposed directions," *Telecom*, vol. 4, no. 4, pp. 758–815, Nov. 2023. [Online]. Available: <http://dx.doi.org/10.3390/telecom4040035>
- [11] G. Borghini, S. Caputo, L. Mucchi, A. Rashid, S. Jayousi, M. Hämäläinen, T. Paso, and M. Hernandez, "Security of wireless body area networks for healthcare applications: Comparison between etsi and ieee approaches," in *2023 IEEE 17th International Symposium on Medical Information and Communication Technology (ISMICT)*. IEEE, May 2023, pp. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/ISMICT58261.2023.10152140>
- [12] S. Caputo, A. Vizziello, A. Coviello, M. Magarini, S. Jayousi, P. Savazzi, and L. Mucchi, "Secret key extraction using galvanic coupling in wireless body area networks," in *2025 19th International Symposium on Medical Information and Communication Technology (ISMICT)*. IEEE, May 2025, pp. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/ISMICT64722.2025.11059426>
- [13] M. Mitev, T. M. Pham, A. Chorti, A. N. Barreto, and G. Fettweis, "Physical layer security-from theory to practice," *IEEE BITS the Information Theory Magazine*, vol. 3, no. 2, pp. 67–79, Jun. 2023. [Online]. Available: <http://dx.doi.org/10.1109/MBITS.2023.3338569>
- [14] L. Mucchi, S. Caputo, P. Marcocci, G. Chisci, L. Ronga, and E. Panayirci, "Security and reliability performance of noise-loop modulation: Theoretical analysis and experimentation," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6335–6350, Jun. 2022. [Online]. Available: <http://dx.doi.org/10.1109/TVT.2022.3160094>
- [15] J. M. Hamamreh and H. Arslan, "Secure orthogonal transform division multiplexing (otdm) waveform for 5g and beyond," *IEEE Communications Letters*, vol. 21, no. 5, pp. 1191–1194, May 2017.

- [16] A. K. Junejo, F. Benkhelifa, B. Wong, and J. A. Mccann, "Lora-lisk: A lightweight shared secret key generation scheme for lora networks," *IEEE Internet of Things Journal*, vol. 9, no. 6, pp. 4110–4124, Mar. 2022. [Online]. Available: <http://dx.doi.org/10.1109/JIOT.2021.3103009>
- [17] H. Salman and H. Arslan, "Pls-iot enhancement against eavesdropping via spatially distributed constellation obfuscation," *IEEE Wireless Communications Letters*, vol. 12, no. 9, pp. 1508–1512, Sep. 2023. [Online]. Available: <http://dx.doi.org/10.1109/LWC.2023.3279989>
- [18] A. Yasar and R. T. Yazicigil, "Physical-layer security for energy-constrained integrated systems: Challenges and design perspectives," *IEEE Open Journal of the Solid-State Circuits Society*, vol. 3, pp. 262–273, 2023. [Online]. Available: <http://dx.doi.org/10.1109/OJSSCS.2023.3327326>
- [19] H.-L. Huang, D. Wu, D. Fan, and X. Zhu, "Superconducting quantum computing: a review," *Science China Information Sciences*, vol. 63, pp. 1–32, 2020.
- [20] B.-M. Zhou and Z. Yuan, "Breaking symmetric cryptosystems using the offline distributed grover-meets-simon algorithm," *Quantum Information Processing*, vol. 22, no. 9, Sep. 2023. [Online]. Available: <http://dx.doi.org/10.1007/s11128-023-04089-9>
- [21] M. Bloch and J. Barros, *Physical-layer security: from information theory to security engineering*. Cambridge University Press, 2011.
- [22] J. He, S. Huang, Y. Chen, S. Chang, Y. Zhang, and Z. Feng, "Radio frequency fingerprint identification for ofdm system considering unknown multipath fading channel," *IEEE Transactions on Cognitive Communications and Networking*, vol. 10, no. 6, pp. 2076–2087, Dec. 2024. [Online]. Available: <http://dx.doi.org/10.1109/TCCN.2024.3405481>
- [23] N. Xie and S. Zhang, "Blind authentication at the physical layer under time-varying fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 7, pp. 1465–1479, Jul. 2018. [Online]. Available: <http://dx.doi.org/10.1109/JSAC.2018.2824583>
- [24] J. Rigau, M. Feixas, and M. Sbert, "Informational aesthetics measures," *IEEE Computer Graphics and Applications*, vol. 28, no. 2, pp. 24–34, Mar 2008. [Online]. Available: <http://dx.doi.org/10.1109/MCG.2008.34>
- [25] S. Rothe, N. Koukourakis, H. Radner, A. Lonnstrom, E. Jorswieck, and J. W. Czarske, "Physical layer security in multimode fiber optical networks," *Scientific Reports*, vol. 10, no. 1, Feb 2020. [Online]. Available: <https://doi.org/10.1038/s41598-020-59625-9>
- [26] K. Guan, A. M. Tulino, P. J. Winzer, and E. Soljanin, "Secrecy capacities in space-division multiplexed fiber optic communication systems," *IEEE*

- Transactions on Information Forensics and Security*, vol. 10, no. 7, pp. 1325–1335, Jul 2015. [Online]. Available: <https://doi.org/10.1109/tifs.2015.2405897>
- [27] M. Baldi, F. Chiaraluce, L. Incipini, and M. Ruffini, “Code-based physical layer secret key generation in passive optical networks,” *Ad Hoc Networks*, vol. 89, pp. 1–8, Jun. 2019. [Online]. Available: <http://dx.doi.org/10.1016/j.adhoc.2019.02.003>
- [28] O. Buskila, A. Eyal, and M. Shtaif, “Secure communication in fiber optic systems via transmission of broad-band optical noise,” *Optics Express*, vol. 16, no. 5, p. 3383, 2008. [Online]. Available: <https://doi.org/10.1364/oe.16.003383>
- [29] Y. Wan, J. Ren, B. Liu, Y. Mao, S. Chen, X. Wu, Y. Li, Y. Wu, L. Zhao, T. Sun, and R. Ullah, “Secure ofdm transmission scheme based on chaotic encryption and noise-masking key distribution,” *Optics Letters*, vol. 47, no. 11, p. 2903, May 2022. [Online]. Available: <https://doi.org/10.1364/ol.460052>
- [30] A. W. Abdulwahhab, A. K. Abass, M. A. Saleh, and F. F. Rashid, “Enhancing performance of optical chaotic-based secure fiber-optic communication system,” *Optical and Quantum Electronics*, vol. 55, no. 5, Apr 2023. [Online]. Available: <https://doi.org/10.1007/s11082-023-04757-1>
- [31] E. Jorswieck, A. Lonnstrom, K.-L. Besser, S. Rothe, and J. W. Czarske, “Achievable physical-layer secrecy in multi-mode fiber channels using artificial noise,” in *2021 17th International Symposium on Wireless Communication Systems (ISWCS)*. IEEE, Sep 2021. [Online]. Available: <https://doi.org/10.1109/iswcs49558.2021.9562176>
- [32] N. Skorin-Kapov, M. Furdek, S. Zsigmond, and L. Wosinska, “Physical-layer security in evolving optical networks,” *IEEE Communications Magazine*, vol. 54, no. 8, pp. 110–117, 2016.
- [33] M. Obeed, A. M. Salhab, M.-S. Alouini, and S. A. Zummo, “Survey on physical layer security in optical wireless communication systems,” in *2018 Seventh International Conference on Communications and Networking (ComNet)*. IEEE, 2018, pp. 1–5.
- [34] F. J. Lopez-Martinez, G. Gomez, and J. M. Garrido-Balsells, “Physical-layer security in free-space optical communications,” *IEEE Photonics Journal*, vol. 7, no. 2, pp. 1–14, 2015.
- [35] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, “Secure quantum key distribution with realistic devices,” *Reviews of Modern Physics*, vol. 92, no. 2, p. 025002, 2020.

- [36] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.
- [37] M. Conti, N. Dragoni, and V. Lesyk, “A survey of man in the middle attacks,” *IEEE Communications Surveys e amp; Tutorials*, vol. 18, no. 3, pp. 2027–2051, 2016. [Online]. Available: <http://dx.doi.org/10.1109/COMST.2016.2548426>
- [38] X. Liu and H. C. Luong, “A fully integrated 0.27-thz injection-locked frequency synthesizer with frequency-tracking loop in 65-nm cmos,” *IEEE Journal of Solid-State Circuits*, vol. 55, no. 4, pp. 1051–1063, 2020.
- [39] *IEEE Standard for High Data Rate Wireless Multi-Media Networks-Amendment 2: 100 Gb/s Wireless Switched Point-to-Point Physical Layer*, IEEE Std. IEEE Std 802.15.3d-2017, 2017, amendment to IEEE Std 802.15.3-2016 as amended by IEEE Std 802.15.3e-2017.
- [40] J. Woo, M. I. W. Khan, M. I. Ibrahim, R. Han, A. P. Chandrakasan, and R. T. Yazicigil, “Physical-layer security for thz communications via orbital angular momentum waves,” in *2022 IEEE Workshop on Signal Processing Systems (SiPS)*. IEEE, 2022, pp. 1–6.
- [41] B. Ning, Z. Chen, W. Chen, and L. Li, “Improving security of thz communication with intelligent reflecting surface,” in *2019 IEEE Globecom Workshops (GC Wkshps)*. IEEE, Dec. 2019, pp. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/GCWkshps45667.2019.9024636>
- [42] A. B. Carlson and P. B. Crilly, *Communication Systems: An Introduction to Signals and Noise in Electrical Communication*, 5th ed. McGraw-Hill, 2010.
- [43] R. Zhang, K. Yang, Q. H. Abbasi, K. A. Qaraqe, and A. Alomainy, “Analytical modelling of the effect of noise on the terahertz in-vivo communication channel for body-centric nano-networks,” *Nano Communication Networks*, vol. 15, pp. 59–68, 2018.
- [44] U. R. Pfeiffer, R. Jain, J. Grzyb, S. Malz, P. Hillger, and P. Rodríguez-Vázquez, “Current status of terahertz integrated circuits-from components to systems,” in *IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*. IEEE, 2018, pp. 1–7.
- [45] R. Han, Z. Hu, C. Wang, J. Holloway, X. Yi, M. Kim, and J. Mawdsley, “Filling the gap: Silicon terahertz integrated circuits offer our best bet,” *IEEE Microwave Magazine*, vol. 20, no. 4, pp. 80–93, 2019.
- [46] Z. Hossain, Q. Xia, and J. M. Jornet, “Terasim: An ns-3 extension to simulate terahertz-band communication networks,” *Nano Communication Networks*, vol. 17, pp. 36–44, 2018.

- [47] M. Seo, M. Urteaga, M. Rodwell, and M.-J. Choe, "A 300 ghz pll in an inphbt technology," in *Proceedings of the IEEE MTT-S International Microwave Symposium*. IEEE, 2011, pp. 1–4.
- [48] Z. Chen, W. Deng, H. Jia, Y. Liu, J. Wu, P. Guan, J. Zhu, L. Mao, Z. Wang, and B. Chi, "A u-band pll using implicit distributed resonators for sub-thz wireless transceivers in 40 nm cmos," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 9, pp. 1574–1578, 2020.
- [49] A. Demir, "Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and $1/f$ noise," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 9, pp. 1869–1884, 2006.
- [50] H. Mehrpouyan, M. R. Khanzadi, M. Matthaiou, A. M. Sayeed, R. Schober, and Y. Hua, "Improving bandwidth efficiency in e-band communication systems," *IEEE Communications Magazine*, vol. 52, no. 3, pp. 121–128, 2014.
- [51] S. Bicaïs and J.-B. Dore, "Phase noise model selection for sub-thz communications," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.
- [52] G. Cuozzo, N. Longhi, and G. Pasolini, "Characterization of orthogonal chirp division multiplexing and performance evaluation at thz frequencies in the presence of phase noise," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 238–255, 2024.
- [53] L. Mucchi, L. S. Ronga, and L. Cipriani, "A new modulation for intrinsically secure radio channel in wireless systems," *Wireless Personal Communications*, vol. 51, pp. 67–80, 2009.
- [54] "Ieee standard for information technology– local and metropolitan area networks– specific requirements– part 15.1a: Wireless medium access control (mac) and physical layer (phy) specifications for wireless personal area networks (wpan)," *IEEE Std 802.15.1-2005 (Revision of IEEE Std 802.15.1-2002)*, pp. 1–700, 2005.
- [55] "Ieee standard for low-rate wireless networks," *IEEE Std 802.15.4-2020 (Revision of IEEE Std 802.15.4-2015)*, pp. 1–800, 2020.
- [56] "Iso/iec/ieee international standard - information technology – telecommunications and information exchange between systems – local and metropolitan area networks – specific requirements – part 15-6: Wireless body area network," *ISO/IEC/IEEE 8802-15-6:2017(E)*, pp. 1–274, 2018.
- [57] "Ieee standard for information technology-telecommunications and information exchange between systems local and metropolitan area networks-specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications," *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pp. 1–3534, 2016.

- [58] “Iso/ieee standard for health informatics - point-of-care medical device communication - transport profile - cable connected,” *ISO/IEEE 11073-30200:2004(E)*, pp. 1–77, 2004.
- [59] M. Hämäläinen, L. Mucchi, M. Girod-Genet, T. Paso, J. Farserotu, H. Tanaka, D. Anzai, L. Pierucci, R. Khan, M. M. Alam, and P. Dalmagne, “Etsi smartban architecture: The global vision for smart body area networks,” *IEEE Access*, vol. 8, pp. 150 611–150 625, 2020.
- [60] “Ieee standard for transport of key management protocol (kmp) datagrams,” *IEEE Std 802.15.9-2021 (Revision of IEEE Std 802.15.9-2016)*, pp. 1–76, 2022.
- [61] “Ieee standard for low-rate wireless networks amendment 3: Advanced encryption standard (aes)-256 encryption and security extensions,” *IEEE Std 802.15.4y-2021 (Amendment to IEEE Std 802.15.4-2020 as amended by IEEE Std 802.15.4z-2020 and IEEE Std 802.15.4w-2020)*, pp. 1–23, 2021.
- [62] M. Toorani, “Security analysis of the ieee 802.15.6 standard,” *International Journal of Communication Systems*, vol. 29, no. 17, pp. 2471–2489, Feb. 2016. [Online]. Available: <https://doi.org/10.1002/dac.3120>
- [63] “IEEE standard for local and metropolitan area networks - part 15.6: Wireless body area networks,” *IEEE Std 802.15.6-2012*, pp. 1–271, 2012.
- [64] A. Rashid, T. Pecorella, and F. Chiti, “Toward resilient wireless sensor networks: A virtualized perspective,” *Sensors*, vol. 20, no. 14, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/14/3902>
- [65] L. Mucchi, S. Jayousi, A. Martinelli, S. Caputo, and P. Marcocci, “An overview of security threats, solutions and challenges in wbans for healthcare,” in *2019 13th International Symposium on Medical Information and Communication Technology (ISMICT)*. IEEE, May 2019, pp. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/ISMICT.2019.8743798>
- [66] T. Nakano, M. J. Moore, F. Wei, A. V. Vasilakos, and J. Shuai, “Molecular communication and networking: Opportunities and challenges,” *IEEE Transactions on NanoBioscience*, vol. 11, no. 2, pp. 135–148, June 2012. [Online]. Available: <https://doi.org/10.1109/TNB.2012.2191570>
- [67] M. Pierobon and I. F. Akyildiz, “A physical end-to-end model for molecular communication in nanonetworks,” *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 4, pp. 602–611, May 2010.
- [68] M. S. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad, and A. Goldsmith, “A survey on modulation techniques in molecular communication via diffusion,” *IEEE Communications Surveys and Tutorials*, vol. 23, no. 1, pp. 7–28, 2021. [Online]. Available: <http://dx.doi.org/10.1109/COMST.2020.3048099>

- [69] E. B. Pehlivanoglu, B. D. Unluturk, and O. B. Akan, *Modulation in Molecular Communications: A Look on Methodologies*. Springer International Publishing, 2017, pp. 79–97. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-50688-3_4
- [70] E. Shitiri and H.-S. Cho, “An m-ary concentration shift keying with common detection thresholds for multitransmitter molecular communication,” *IEEE Internet of Things Journal*, pp. 1–1, 2024. [Online]. Available: <http://dx.doi.org/10.1109/JIOT.2024.3359999>
- [71] M. H. Kabir, S. M. Riazul Islam, and K. S. Kwak, “D-mosk modulation in molecular communications,” *IEEE Transactions on NanoBioscience*, vol. 14, no. 6, pp. 680–683, Sep. 2015. [Online]. Available: <http://dx.doi.org/10.1109/TNB.2015.2436409>
- [72] Y. Murin, N. Farsad, M. Chowdhury, and A. Goldsmith, “Communication over diffusion-based molecular timing channels,” in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, Dec. 2016. [Online]. Available: <http://dx.doi.org/10.1109/GLOCOM.2016.7841672>
- [73] B. C. Akdeniz, A. E. Pusane, and T. Tugcu, “Position-based modulation in molecular communications,” *Nano Communication Networks*, vol. 16, pp. 60–68, Jun. 2018. [Online]. Available: <http://dx.doi.org/10.1016/j.nancom.2018.01.004>
- [74] Y. Huang, M. Wen, L.-L. Yang, C.-B. Chae, X. Chen, and Y. Tang, “Space shift keying for molecular communication: Theory and experiment,” in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, Dec. 2019. [Online]. Available: <http://dx.doi.org/10.1109/GLOBECOM38437.2019.9013607>
- [75] S. Shahbaz, M. Mirmohseni, and M. Nasiri-Kenari, “A jamming resistant molecular communication scheme,” in *2022 10th Iran Workshop on Communication and Information Theory (IWCIT)*. IEEE, May 2022. [Online]. Available: <http://dx.doi.org/10.1109/IWCIT57101.2022.10206797>
- [76] S. Zafar, W. Aman, M. M. U. Rahman, A. Alomainy, and Q. H. Abbasi, “Channel impulse response-based physical layer authentication in a diffusion-based molecular communication system,” in *2019 UK/ China Emerging Technologies (UCET)*. IEEE, Aug. 2019. [Online]. Available: <http://dx.doi.org/10.1109/UCET.2019.8881878>
- [77] Y.-H. Joung, “Development of implantable medical devices: From an engineering perspective,” *International Neurourology Journal*, vol. 17, no. 3, p. 98, 2013.

- [78] N. Chi, H. Haas, M. Kavehrad, T. D. Little, and X.-L. Huang, "Visible light communications: demand factors, benefits and opportunities [guest editorial]," *IEEE Wireless Communications*, vol. 22, pp. 5–7, Apr. 2015.
- [79] K. Agarwal, R. Jegadeesan, Y.-X. Guo, and N. V. Thakor, "Wireless power transfer strategies for implantable bioelectronics," *IEEE Reviews in Biomedical Engineering*, vol. 10, pp. 136–161, 2017.
- [80] H.-J. Kim, H. Hirayama, S. Kim, K. J. Han, R. Zhang, and J.-W. Choi, "Review of near-field wireless power and communication for biomedical applications," *IEEE Access*, vol. 5, pp. 21 264–21 285, 2017.
- [81] A. C. Thompson, S. A. Wade, N. C. Pawsey, and P. R. Stoddart, "Infrared neural stimulation: Influence of stimulation site spacing and repetition rates on heating," *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 3534–3541, Dec. 2013.
- [82] I. Sohn, Y. H. Jang, and S. H. Lee, "Ultra-low-power implantable medical devices: Optical wireless communication approach," *IEEE Communications Magazine*, vol. 58, pp. 77–83, May 2020.
- [83] I. Ahmed, S. Halder, A. Bykov, A. Popov, I. V. Meglinski, and M. Katz, "In-body communications exploiting light: A proof-of-concept study using ex vivo tissue samples," *IEEE Access*, vol. 8, pp. 190 378–190 389, 2020.
- [84] A. Minotto, P. A. Haigh, L. G. Lukasiewicz, E. Lunedei, D. T. Gryko, I. Darwazeh, and F. Cacialli, "Visible light communication with efficient far-red/near-infrared polymer light-emitting diodes," *Light: Science Applications*, vol. 9, Apr. 2020.
- [85] D. Marabissi, A. Abrardo, and L. Mucchi, "A new framework for physical layer security in hetnets based on radio resource allocation and reinforcement learning," *Mobile Networks and Applications*, vol. 28, no. 4, pp. 1473–1481, Aug. 2023. [Online]. Available: <http://dx.doi.org/10.1007/s11036-023-02149-z>
- [86] V. V. Tuchin, "Light scattering study of tissues," *Physics-Uspekhi*, vol. 40, pp. 495–515, May 1997.
- [87] I. V. Meglinski and S. J. Matcher, "Quantitative assessment of skin layers absorption and skin reflectance spectra simulation in the visible and near-infrared spectral regions," *Physiological Measurement*, vol. 23, pp. 741–753, Oct. 2002.
- [88] L. C. Henyey and J. L. Greenstein, "Diffuse radiation in the galaxy," *The Astrophysical Journal*, vol. 93, p. 70, Jan. 1941.
- [89] "Ieee standard for local and metropolitan area networks-part 15.7: Short-range optical wireless communications," IEEE, 2011.

- [90] Q. Cheng, J. Zhu, T. Xie, J. Luo, and Z. Xu, "Time-invariant angle-range dependent directional modulation based on time-modulated frequency diverse arrays," *IEEE Access*, vol. 5, pp. 26 279–26 290, 2017. [Online]. Available: <http://dx.doi.org/10.1109/ACCESS.2017.2772246>
- [91] A. N. Bashkatov, E. A. Genina, and V. V. Tuchin, "Optical properties of skin, subcutaneous, and muscle tissues: a review," *Journal of Innovative Optical Health Sciences*, vol. 4, pp. 9–38, Jan. 2011.
- [92] A. Vizziello, P. Savazzi, G. Magenes, and P. Gamba, "Phy design and implementation of a galvanic coupling testbed for intra-body communication links," *IEEE Access*, vol. 8, pp. 184 585–184 597, 2020. [Online]. Available: <http://dx.doi.org/10.1109/ACCESS.2020.3029862>
- [93] M. Seyedi, B. Kibret, D. T. H. Lai, and M. Faulkner, "A survey on intrabody communications for body area network applications," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2067–2079, Aug. 2013. [Online]. Available: <http://dx.doi.org/10.1109/TBME.2013.2254714>
- [94] W. J. Tomlinson, "Physical layer design and implementation of a biometric authentication system using galvanic coupling intra-body communication," PhD thesis, Northeastern University, 2018. [Online]. Available: <https://repository.library.northeastern.edu/files/neu:m044c525c/fulltext.pdf>
- [95] M. Swaminathan, A. Vizziello, D. Duong, P. Savazzi, and K. R. Chowdhury, "Beamforming in the body: Energy-efficient and collision-free communication for implants," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*. IEEE, 2017, pp. 1–9.
- [96] A. Vizziello, P. Savazzi, R. R. Guerra, and F. Dell'Acqua, "Experimental channel characterization of human body communication based on measured impulse response," *IEEE Transactions on Communications*, vol. 72, no. 7, pp. 3970–3984, 2024.
- [97] L. Shi, J. Yuan, S. Yu, and M. Li, "Ask-ban: Authenticated secret key extraction utilizing channel characteristics for body area networks," in *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks*. ACM, 2013, pp. 155–166.
- [98] —, "Mask-ban: Movement-aided authenticated secret key extraction utilizing channel characteristics in body area networks," *IEEE Internet of Things Journal*, vol. 2, no. 1, pp. 52–62, 2015.
- [99] J. Yuan, L. Shi, S. Yu, and M. Li, "Authenticated secret key extraction using channel characteristics for body area networks," in *Proceedings of the 2012 ACM Conference on Computer and Communications Security*. ACM, 2012, pp. 1030–1032.

- [100] Q. Gui, Z. Jin, and W. Xu, "Exploring eeg-based biometrics for user identification and authentication," in *2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2014, pp. 1–6.
- [101] D. Kaur, D. Singh, and P. Roy, "A novel framework of eeg-based user identification by analyzing music-listening behavior," *Multimedia Tools and Applications*, vol. 76, pp. 25 581–25 602, 2017.
- [102] Q. Wu, Y. Zeng, C. Zhang, L. Tong, and B. Yan, "An eeg-based person authentication system with open-set capability combining eye blinking signals," *Sensors*, vol. 18, no. 11, p. 335, 2018.
- [103] T. Schons, G. Moreira, P. Silva, V. Coelho, and E. Luz, "Convolutional network for eeg-based biometric," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Springer, 2018, pp. 601–608.
- [104] A. Bidgoly, H. Bidgoly, and Z. Arezoumand, "Towards a universal and privacy preserving eeg-based authentication system," *Scientific Reports*, vol. 21, pp. 1–12, 2022.
- [105] A. Valsaraj, I. Madala, N. Garg, M. Patil, and V. Baths, "Motor imagery based multimodal biometric user authentication system using eeg," in *2020 International Conference on Cyberworlds (CW)*. IEEE, 2020, pp. 272–279.
- [106] A. Zuquete, B. Quintela, and J. Cunha, "Biometric authentication using brain responses to visual stimuli," in *Third International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS)*, 2010, pp. 103–112.
- [107] E. Gupta, M. Agarwal, and R. Sivakumar, "Blink to get in: Biometric authentication for mobile devices using eeg signals," in *2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.
- [108] L. Moctezuma and M. Molinas, "Multi-objective optimization for eeg channel selection and accurate intruder detection in an eeg-based subject identification system," *Scientific Reports*, vol. 10, p. 5850, 2020.
- [109] S. Chakravarthy and H. Rajaguru, "Effective breast tumor classification using k-strongest strength with local outlier factor algorithm," *International Journal of Aquatic Science*, vol. 12, pp. 1596–1603, 2021.
- [110] W. J. Tomlinson, S. Banou, S. Blechinger-Slocum, C. Yu, and K. R. Chowdhury, "Body-guided galvanic coupling communication for secure biometric data," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4143–4156, Aug. 2019. [Online]. Available: <http://dx.doi.org/10.1109/TWC.2019.2921964>
- [111] L. C. Nguyen, N. V. Duc, N. V. Huong, and N. V. Tuan, "On the desired properties of linear feedback shift register (lfsr) based high-speed pn-sequence generator," in *Lecture Notes in Networks and Systems*, vol. 115. Springer, 2020, pp. 563–573.

- [112] A. Ahmad, S. S. Al-Busaidi, and M. J. Al-Musharafi, "On properties of pn sequences generated by lfsr - a generalized study and simulation modeling," *Indian Journal of Science and Technology*, vol. 9, no. 30, 2016.
- [113] G. D. Birkhoff, *Aesthetic Measure*. Harvard University Press, Dec 1933. [Online]. Available: <http://dx.doi.org/10.4159/harvard.9780674734470>
- [114] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jun 2022, pp. 10 674–10 685. [Online]. Available: <http://dx.doi.org/10.1109/CVPR52688.2022.01042>
- [115] "Wikiart - visual art encyclopedia," <https://www.wikiart.org>, 2025.
- [116] A. Y. J. Ha, J. Passananti, R. Bhaskar, S. Shan, R. Southen, H. Zheng, and B. Y. Zhao, "Organic or diffused: Can we distinguish human art from ai-generated images?" in *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security (CCS '24)*. ACM, Dec 2024, pp. 4822–4836. [Online]. Available: <http://dx.doi.org/10.1145/3658644.3670306>
- [117] E.-M. Papia, A. Kondi, and V. Constantoudis, "Entropy and complexity analysis of ai-generated and human-made paintings," *SSRN Electronic Journal*, 2023. [Online]. Available: <http://dx.doi.org/10.2139/ssrn.4350021>
- [118] L. Yuan, X. Li, Y. Zhang, J. Zhang, H. Li, and X. Gao, "Mlep: Multi-granularity local entropy patterns for universal ai-generated image detection," <https://arxiv.org/abs/2504.13726>, 2025.
- [119] D. Cozzolino, G. Poggi, M. Nieener, and L. Verdoliva, "Zero-shot detection of ai-generated images," in *Springer Nature Switzerland*, Sep 2024, pp. 54–72. [Online]. Available: http://dx.doi.org/10.1007/978-3-031-72649-1_4
- [120] R. Zhang, H. Wang, Z. Zhao, Z. Guo, X. Yang, Y. Diao, and M. Wang, "Adversarially robust ai-generated image detection for free: An information theoretic perspective," <https://arxiv.org/abs/2505.22604>, 2025.
- [121] N. Zhong, Y. Xu, S. Li, Z. Qian, and X. Zhang, "Patchcraft: Exploring texture patch for efficient ai-generated image detection," <https://arxiv.org/abs/2311.12397>, 2023.
- [122] C. Tan, H. Liu, Y. Zhao, S. Wei, G. Gu, P. Liu, and Y. Wei, "Rethinking the up-sampling operations in cnn-based generative network for generalizable deepfake detection," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jun 2024, pp. 28 130–28 139. [Online]. Available: <http://dx.doi.org/10.1109/CVPR52733.2024.02657>
- [123] D. S. Chinta, S. Kamineni, R. P. Chatragadda, and S. Kamepalli, "Analyzing image classification on ai-generated art vs human created

- art using deep learning models,” in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*. IEEE, Jul 2024, pp. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/ICEEICT61591.2024.10718485>
- [124] F. F. Khan, D. Kim, D. Jha, Y. Mohamed, H. H. Chang, A. Elgammal, L. Elliott, and M. Elhoseiny, “Ai art neural constellation: Revealing the collective and contrastive state of ai-generated and human art,” in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, Jun 2024, pp. 7470–7478. [Online]. Available: <http://dx.doi.org/10.1109/CVPRW63382.2024.00742>
- [125] M. Li and M. Stamp, “Detecting ai-generated artwork,” <https://arxiv.org/abs/2504.07078>, 2025.
- [126] J. J. Bird, C. M. Barnes, and A. Lotfi, “Ai generated art: Latent diffusion-based style and detection,” in *Springer Nature Switzerland*, 2024, pp. 157–169. [Online]. Available: http://dx.doi.org/10.1007/978-3-031-47508-5_13
- [127] K. Harsanto, M. T. Zaidan, H. Lucky, and D. Suhartono, “Detecting ai-generated artworks with deep learning,” in *2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*. IEEE, Nov 2024, pp. 472–476. [Online]. Available: <http://dx.doi.org/10.1109/ICIMCIS63449.2024.10956228>
- [128] S. T. Oo, P. P. Htun, L. Y. W. Lwin, A. N. Htwe, and N. Aye, “Classifying real versus ai generated images using cifar-10 network and painting image data,” in *2025 IEEE Conference on Computer Applications (ICCA)*. IEEE, Mar 2025, pp. 1–5. [Online]. Available: <http://dx.doi.org/10.1109/ICCA65395.2025.11011094>
- [129] H. S. Ohal, T. S. Balote, M. H. Joshi, H. C. Suryawanshi, and N. V. Suryawanshi, “Leveraging convolutional neural networks for ai artwork classification,” in *2025 International Conference on Emerging Smart Computing and Informatics (ESCI)*. IEEE, Mar 2025, pp. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/ESCI63694.2025.10988229>
- [130] K. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*. Springer Science & Business Media, 2000.
- [131] International Telecommunication Union - Radiocommunication Sector (ITU-R), “Parameter values for the hdtv standards for production and international programme exchange,” ITU-R, Tech. Rep. BT.709-6, 2015, last updated 2015. [Online]. Available: <https://www.itu.int/rec/R-REC-BT.709>
- [132] M. Li, X. Chen, X. Li, B. Ma, and P. Vitanyi, “The similarity metric,” *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3250–3264, Dec 2004. [Online]. Available: <http://dx.doi.org/10.1109/TIT.2004.838101>

- [133] G. Wallace, "The jpeg still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992. [Online]. Available: <http://dx.doi.org/10.1109/30.125072>
- [134] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, 1991. [Online]. Available: <http://dx.doi.org/10.1109/18.61115>
- [135] S. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, Mar 1982. [Online]. Available: <http://dx.doi.org/10.1109/TIT.1982.1056489>
- [136] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," Stanford InfoLab, Tech. Rep. Technical Report 2006-13, Jun 2006. [Online]. Available: <http://ilpubs.stanford.edu:8090/778/>
- [137] N. Cristianini and E. Ricci, "Support vector machines: 1992; boser, guyon, vapnik," in *Springer US*, 2008, pp. 928–932. [Online]. Available: http://dx.doi.org/10.1007/978-0-387-30162-4_415
- [138] D. Boswell, "Introduction to support vector machines," *Department of Computer Science and Engineering, University of California San Diego*, vol. 11, pp. 16–17, 2002.
- [139] S. Suthaharan, *Support Vector Machine*. Springer US, 2016, pp. 207–235. [Online]. Available: http://dx.doi.org/10.1007/978-1-4899-7641-3_9
- [140] K. Pal and B. V. Patel, "Data classification with k-fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques," in *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, Mar 2020, pp. 83–87. [Online]. Available: <http://dx.doi.org/10.1109/ICCMC48092.2020.ICCMC-00016>
- [141] L. Pinheiro Cinelli, M. Araújo Marins, E. A. Barros da Silva, and S. Lima Netto, *Variational Autoencoder*. Springer International Publishing, 2021, pp. 111–149. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-70679-1_5
- [142] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," <https://arxiv.org/abs/2103.00020>, 2021.
- [143] (2025) Automatic111/stable-diffusion-webui. [Online]. Available: <https://github.com/AUTOMATIC111/stable-diffusion-webui>
- [144] W. R. Tan, C. S. Chan, H. E. Aguirre, and K. Tanaka, "Improved artgan for conditional synthesis of natural image and artwork," *IEEE Transactions on*

- Image Processing*, vol. 28, no. 1, pp. 394–409, Jan 2019. [Online]. Available: <http://dx.doi.org/10.1109/TIP.2018.2866698>
- [145] (2025) Wikiart - visual art encyclopedia. [Online]. Available: <https://www.wikiart.org>
- [146] Y. Bakreski, “The impact of technological advancements on art-parallels between the renaissance, the discovery of photography, and the digital revolution,” *Journal of Arts and Humanities*, vol. 11, no. 04, pp. 17–25, 2022.
- [147] “Artificio/wikiart,” <https://huggingface.co/datasets/Artificio/WikiArt>, 2025.
- [148] (2025) Imagemagick. [Online]. Available: <https://imagemagick.org/>
- [149] (2025) Ieee dataport. [Online]. Available: <https://dx.doi.org/10.21227/jb0b-w582>