Transmitter Optimization Techniques for Physical Layer Security

Miao Zhang

Doctor of Philosophy

University of York

Electronic Engineering

September 2019

Abstract

Information security is one of the most critical issues in wireless networks as the signals transmitted through wireless medium are more vulnerable for interception. Although the existing conventional security techniques are proven to be safe, the broadcast nature of wireless communications introduces different challenges in terms of key exchange and distributions. As a result, information theoretic physical layer security has been proposed to complement the conventional security techniques for enhancing security in wireless transmissions. On the other hand, the rapid growth of data rates introduces different challenges on power limited mobile devices in terms of energy requirements. Recently, research work on wireless power transfer claimed that it has been considered as a potential technique to extend the battery lifetime of wireless networks. However, the algorithms developed based on the conventional optimization approaches often require iterative techniques, which poses challenges for real-time processing. To meet the demanding requirements of future ultra-low latency and reliable networks, neural network (NN) based approach can be employed to determine the resource allocations in wireless communications.

This thesis developed different transmission strategies for secure transmission in wireless communications. Firstly, transmitter designs are focused in a multiple-input single-output simultaneous wireless information and power transfer system with unknown eavesdroppers. To improve the performance of physical layer security and the harvested energy, artificial noise is incorporated into the network to mask the secret information between the legitimate terminals. Then, different secrecy energy efficiency designs are considered for a MISO underlay cognitive radio network, in the presence of an energy harvesting receiver. In particular, these designs are developed with different channel state information assumptions at the transmitter. Finally, two different power allocation designs are investigated for a cognitive radio network to maximize the secrecy rate of the secondary receiver: conventional convex optimization framework and NN based algorithm.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1

Introduction

1.1 Background and Motivations

Mobile devices and wireless technologies have become an indispensable part of daily activities of people and they continue to underpin the lives of people and the infrastructures of modern networked society that people live in today [1, 2]. In 1870s, the existence of electromagnetic waves was proved by James Clerk Maxwell [3]. Based on the continued development of micro-electronic circuits, the first generation (1G) commercial cellular network was introduced in the late 1970s and fully established in 1980s, which used analogue radio signals to provide voice services [1]. Then, in 1990s 1G networks were replaced by the second generation (2G) cellular networks. The 2G networks were developed based on digital technologies, and it provided both voice and data services and allowed users to experience roaming facilities for the first time [4]. The third generation (3G) wireless cellular network was introduced in 1998 and commercially started operating in 2001 [4]. In contrast to the 2G networks, the 3G networks provided high speed wireless Internet access, which made the multimedia applications feasible on mobile devices, i.e., video chats, live TV, navigation services, email, online games, and music. Then the fourth generation (4G) cellular networks commercially deployed in 2010, where 4G users could receive higher speed, high quality and capacity mobile services with improved security and lower cost than the services offered by the 3G networks [5]. The fifth generation (5G) is the next generation commercial mobile network, which is expected to roll-out by 2020 [6] with unleashing enormous potential opportunities for information communication technologies - driven vertical industries. The key requirements for 5G and beyond wireless networks include unprecedented higher data
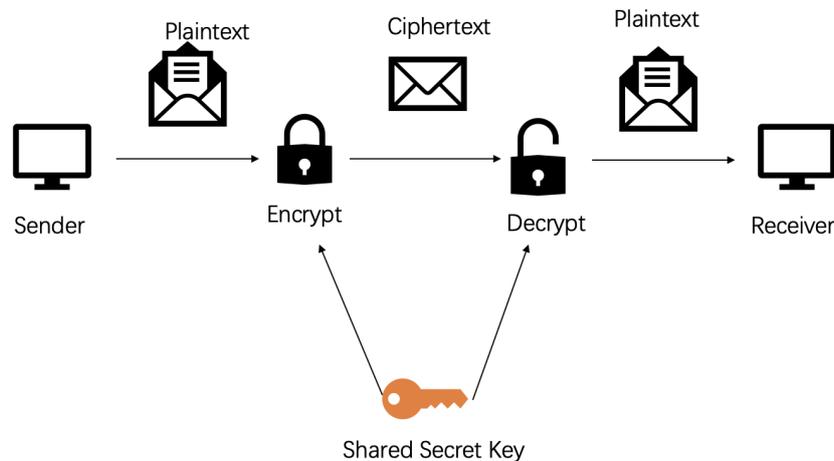
Fig. 1.1 The functions involved in a symmetric encryption system. The secret key cryptology operates in both senders and recipients.

rates with higher energy efficiency (EE), lower cost, ultra reliability and low latency as well as massive connectivity to support proliferation of Internet of Things (IoT).

Over the past decade, many state-of-the-art techniques have been developed to provide a better quality of experience (QoE) to the end users, for example, Wireless Fidelity (WiFi), the Global Positioning System (GPS), wireless sensor networks (WSN), mobile payments, etc. However, the explosive growth of mobile data traffic and newly emerging high data rate applications and services introduce an ever-growing demand for much higher capacity and lower end-to-end latency [7], which poses different challenges on the available wireless resources and wireless network infrastructures. As the current wireless communication technologies do not have the capabilities to meet the unprecedented requirements of future wireless networks, novel disruptive technologies and intelligent radio resource management techniques need to be developed for 5G and beyond wireless networks. Based on the diverse set of requirements, researchers have defined the most relevant performance targets of 5G networks which include security and energy saving [8].

On the other hand, information security has become one of the major issues in wireless networks when people completely rely on wireless transmissions for exchanging and authenticating their confidential informations, such as Internet mobile banking, card transactions and mobile payments. Due to the broadcast nature of the wireless transmissions, the signals transmitted through wireless medium are more vulnerable for interception. Therefore, the ability to share confidential information reliably in the presence of adversaries is extremely important in future wireless networks.
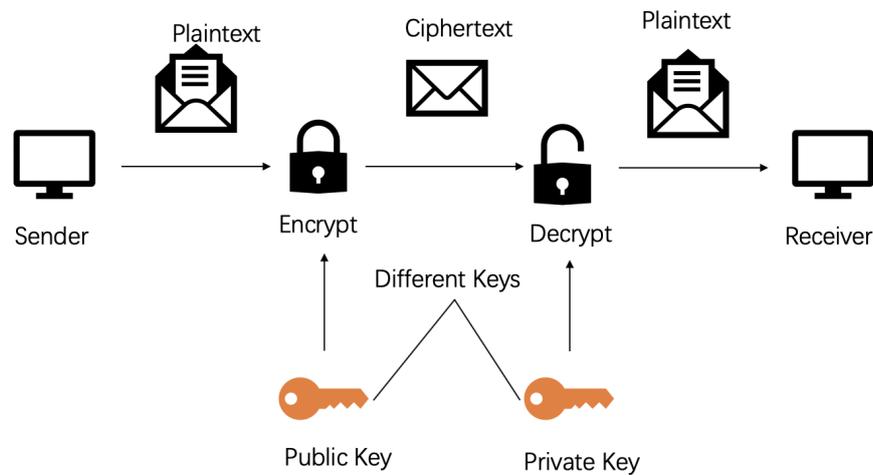
Fig. 1.2 The RSA algorithm.

Conventionally, crypto-based encryption schemes are most widely used to establish secure communications through avoiding intercepting or eavesdropping. The fundamental idea of these methods is to employ different types of secret keys at the application layer to encrypt or decrypt confidential messages [9]. The transmitter encrypts the plain text into cipher text by exploiting the secret key and then sends the cipher text to the authorized receiver. When the authorized receiver receives the message from the transmitter, it cannot read the information directly due to the cipher text format of the message. The receiver has to use the secret key to decrypt the message and transform the message into plain text. If the receiver is not the intended destination of the transmitter, then its secret key cannot decrypt the message. There are two types of conventional security methods: symmetric encryption system and asymmetric encryption system. As shown in Fig 1.1, the symmetric encryption system employs the same secret key at both senders and recipients. For example, both data encryption standard (DES) and advanced encryption standard (AES) are symmetric encryption standards. However, symmetric encryption system is considered vulnerable in wireless communications as the broadcast nature of electromagnetic waves as well as the characteristics of conventional security which completely rely on computation complexity of some intractable mathematical problems. It is more and more vulnerable to be intercepted due to the recent development of advanced processors with high computational capabilities.

The Rivest–Shamir–Adleman (RSA) is one of the most widely used asymmetric crypto-systems. RSA is made of the initial letters of the surnames of Ron Rivest, Adi Shamir, and Leonard Adleman, who first publicly described the algorithm in 1977. RSA is one of the first practical public-key crypto-systems for confidential data transmission. In this asymmetric crypto-system, the public key is utilized only for data encryption whereas the private key

is used only for cipher text decryption. In RSA, this asymmetry is developed based on the practical difficulty of factoring the product of two large prime numbers. As shown in Fig 1.2, the sender exploits the public key to encrypt the information while the receiver employing the private key to decrypt it. The receiver creates and then publishes a public key based on two large prime numbers, along with an auxiliary value, which must be kept secret [10]. Anyone can use the public key to encrypt a message, but the cipher text encrypted by the public key can only be decrypted by the receiver [10].

Despite the existing conventional security techniques are being difficult to break or intercept, the broadcast nature of wireless transmissions introduces different challenges in terms of secret key exchange and distributions in these security techniques [11–18], which are developed based on high complexity of some intractable mathematical problems. As a result, information theoretic based physical layer security has been considered recently to complement the conventional cryptographic methods while providing additional security in wireless transmissions. Different from conventional methods, physical layer security is not an independent part of communication systems. It utilizes the physical characteristics of wireless channels to implement secure transmission in a communication system [19–24]. The theoretical foundation was proposed firstly by Shannon, termed as the perfect secrecy [25], and then it was developed by Wyner [26]. In [26], Wyner introduced a wiretap channel to accomplish secure transmission. The condition to realize confidential communication is that the signal to noise ratio (SNR) or the signal to interference plus noise ratio (SINR) of the legitimate channel is higher than that of the wiretap channel. The secrecy rate is the only indicator to measure the performance of the physical layer security of a communication system. It is the maximum data rate that the transmitter can communicate with the receiver at which the eavesdropper (EVE) cannot intercept anything. The secrecy rate is equal to the difference between the capacity of legitimate channel and the capacity of wiretap channel [26].

Energy is also a critical resource in wireless communication systems. The unprecedented requirements of future wireless networks bring different challenges on the aspects surrounding the energy consumption. Furthermore, the energy consumption, regardless of the underpinning technologies, has a knock-on effect on the environment, carbon footprint, global warming and thus unanticipated financial consequences. On the other hand, the operation time of energy-constrained wireless networks is limited by the life-time of the batteries, i.e., WSN and IoT devices. Although the operation time of wireless devices could be extended through charging or replacing the batteries, it may result in higher expenses and inconvenient, for example, the sensors embedded in buildings and human bodies [27]. In general, conventional energy harvesting (EH) methods collect energy from different external natural sources,
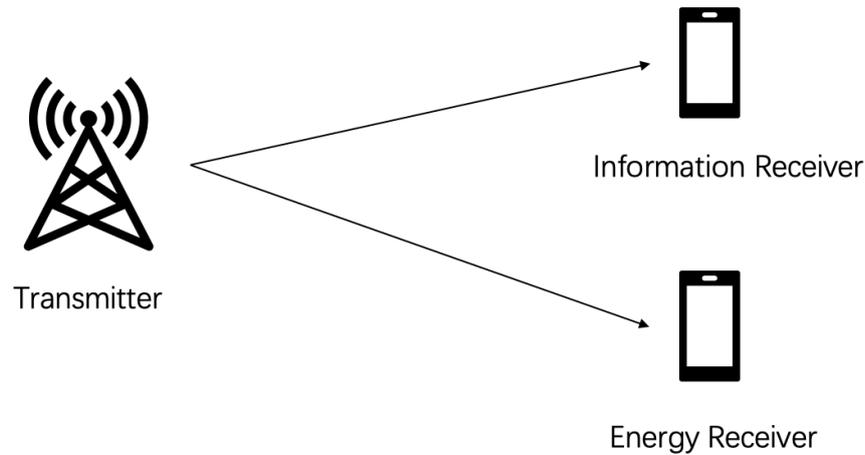
Fig. 1.3 A communication network with separated ER and IR.

(a) Time Switching   (b) Power Splitting
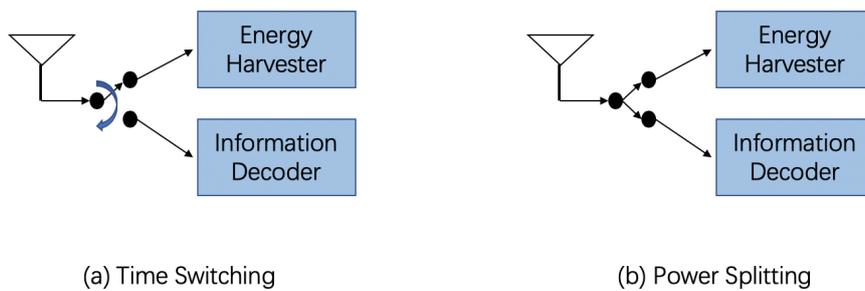
Fig. 1.4 Two practical designs for the co-located ER and IR.

including wind, solar as well as sea waves [28, 29]. However, these external energy resources are always not reliable and unpredictable, and completely depend on the present nature of the environments. Furthermore, it is difficult to incorporate these natural energy sources on mobile devices, due to the size of harvesting devices and the geographical limitations. In contrast to the other renewable energy sources, wireless power transfer (WPT) facilitates practical design and implementation of EH especially in mobile devices [12]. Due to the fact that radio frequency (RF) signals can carry both information and energy, WPT has become a promising technology to address the power issue of wireless systems, which enables wireless devices collecting energy from the information carrying RF signals [30, 31]. WPT has been implemented in practical scenarios in terms of near-field electromagnetic induction and far-field electromagnetic radiation. For example, passive RF identification is widely used in short distance applications while transmitting energy from satellites to the ground or to other

satellites through electromagnetic waves or lasers can be seen as long range implementations [27].

Simultaneous wireless information and power transfer (SWIPT) is one of the promising techniques to implement WPT. Two types of SWIPT framework are investigated in the literature [27, 32, 33], namely separated and co-located energy receiver (ER) and information receiver (IR). A simple communication system with separated receivers is shown in Fig. 1.3, where the transmitter simultaneously sends information and power to IR and ER, respectively. For the case of co-located ER and IR, there exists two different practical designs: time splitting (TS) and power splitting (PS) approaches. As shown in Fig. 1.4 (a), the receiving antenna periodically switches between the energy and the IR in TS approach whereas for PS approach shown in Fig. 1.4 (b), the received signal is split into two separate streams with different power levels and sent to the ER and the IR, respectively.

Recently, artificial intelligence (AI) techniques have been recognized as the potential solution approaches to solve different challenging problems for which the solutions are difficult to determine due to complicated structure of the problems [34]. Furthermore, AI has become as one of the fastest growing fields in many research topics [35] and the fundamental idea of AI was first proposed by Alan Turing in his classic paper "Computing machinery and intelligence" in 1950 [36]. In this seminal work, Turing raised a question "Can machines think?" and proposed a test, namely Turing's test, to answer it. In the following decades, the research of AI went through ups and downs, since some problems were proved infeasible to be implemented with the technical conditions available at that time [34]. Currently, both the number and the performance of AI applications have been significantly improved, due to two factors: (1) the availability of big data, which is significantly beneficial for improving machine learning algorithms; (2) the increased processing capability of computers due to recent advancements of processors, which made it possible for computers to process a huge amount of data within a short time frame [34].

Machine learning is one of the key techniques to implement AI, which enables computers to gather knowledge from its computations and make decisions according to the environment [35, 37, 38]. The performance of machine learning tasks often heavily depends on the available training data. In particular, training process is one of the key steps in machine learning, as the computers learn how to provide a reasonable output results from its inputs [39]. In other words, training is the process that significantly helps the computers to learn the relationship between the inputs and the desired outputs. There are many machine learning frameworks available in the literature [40, 35, 37, 38], including linear regression, logistic regression, decision tree, neural network (NN), reinforcement learning and so on. In this
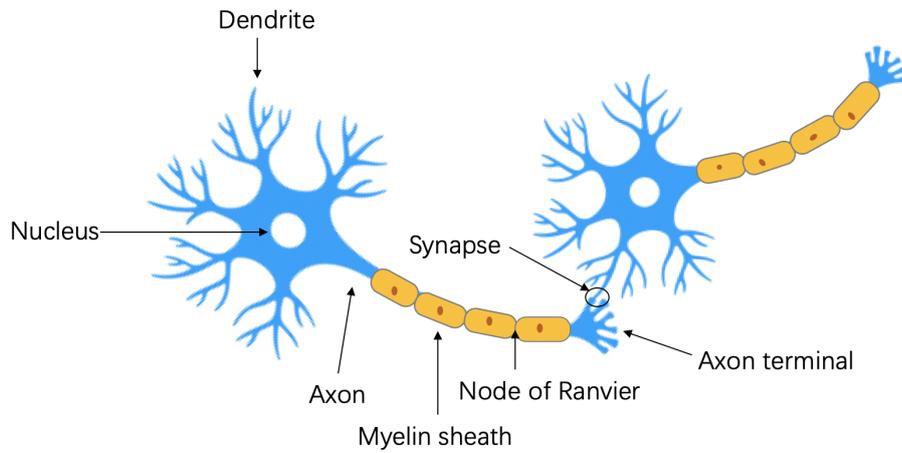
Fig. 1.5 A biological neuron model.
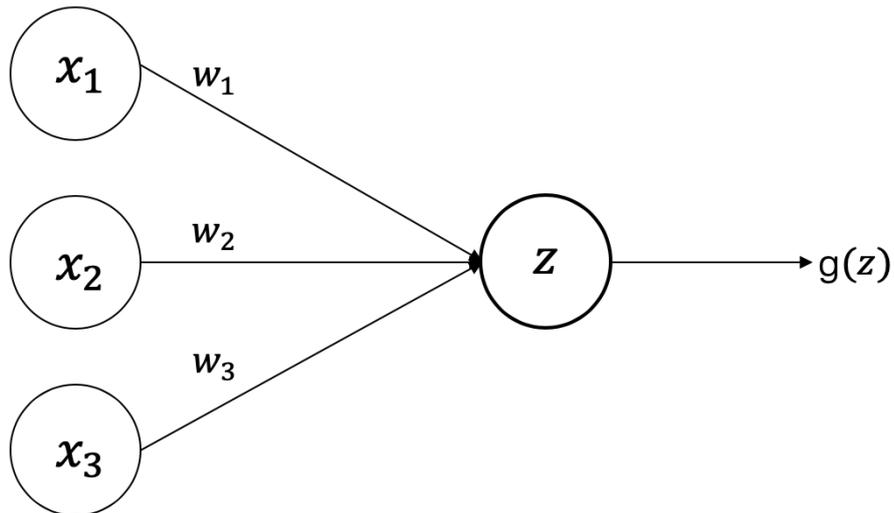
Fig. 1.6 A simple mathematical neuron model.

thesis, only NN approach is exploited to solve resource allocation problems in wireless communications, however, other machine learning techniques for wireless communications will be explored in the future work.

NN is one of the most well-known and commonly used frameworks in machine learning [41], as it is developed to imitate the structure and the function of biological brains. As shown in Fig. 1.5, a biological *neuron* has a number of input wires and these are called *dendrites*. which receive inputs from other *neurons*. Furthermore, a *neuron* also has an output wire called the *axon*, which is used to send signals to other *neurons*. In other words, a *neuron* is a computational unit that receives a number of input signals from *dendrites*, then performs some
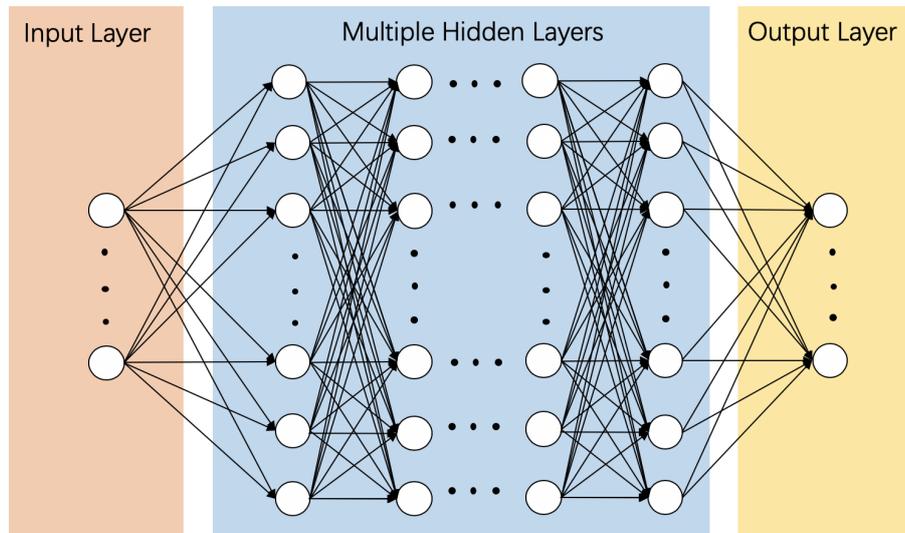
Fig. 1.7 A NN model.

computation and sends its computed results to all of its connected *neurons* through its *axons* [39]. A simple mathematical neuron model is provided in Fig. 1.6, which mimics biological *neurons* to provide an output to its connected neurons based on its inputs. In particular, a NN consists of a number of simple, highly interconnected computation units called neurons, which have the capability to represent this process of biological neuron systems [39].

A NN model is shown in Fig. 1.7, which consists of multiple layers and multiple neurons in each layer. In particular, a NN should consist of the following basic elements [39]:

1. The input layer which is used to process the input data and forward the data to the corresponding neurons. The size of input layer is related to the size of input data.

2. The input weight matrix that presents the weights between the connected neurons in the input layer and that in the hidden layer.

3. The hidden layer(s) contain(s) single or multiple layers with a number of neurons to mimic the procedure of biological brains.

4. The hidden weight matrix that presents the weights between the connected neurons in the hidden layer(s).

5. The output layer contains some neurons to present the output data.

6. The output weight matrix that presents the weights between the connected neurons in the hidden layer and that in the output layer.

Several advantages of employing NN rather than other machine learning techniques are summarized as follows:

1. NN has the potential capabilities to provide a solution within a short time frame with reduced computational complexity [42], compared to that of the other machine learning techniques such as support vector machine and Gaussian process which sometimes require to store all the available data for their processing [42, 43]. However, employing mini-batch gradient descent algorithm to train a NN requires only a sub-set of the available large data set at each training step.

2. A single NN model can be trained to meet the objectives of multiple tasks [42], whereas it is difficult for other machine learning techniques to achieve those multi-objectives with the same model.

3. Furthermore, NN has the capability to automatically extract features from the data with highly complicated structure and to establish the hidden internal relationships, which could help to significantly reduce processing requirements of massive amount of data set [44].

1.2   Thesis outline and Contributions

Motivated by the growing challenges of security, EE and low latency in 5G and beyond wireless networks, the aim of this thesis is to develop different resource allocation techniques to improve security and EE in future wireless networks. This thesis consists of seven chapters and the main contributions of the remaining chapters are summarized as follows:

Chapter 2 presents the fundamental concepts and related literature review. Firstly, the mathematical expression of secrecy rate is provided through illustrating a wiretap channel model. Furthermore, the achievable secrecy rate in a single-input single-output (SISO) system and a multiple-input single-output (MISO) system are presented, respectively. Then, the fundamental concept and mathematical representations of SWIPT are presented. At last, the recent work on physical layer security, SWIPT and machine learning techniques are discussed.

In Chapter 3, the required mathematical background is provided. The fundamental concepts of convex optimization are first presented. Then, various convex optimization problems, such as linear programming (LP), quadratic programming (QP), quadratic constrained quadratic programming (QCQP), second-order cone programming (SOCP) and semidefinite

programming (SDP), as well as duality and Karush Kuhn Tucker (KKT) conditions are discussed. Furthermore, the mathematical expressions for several machine learning approaches are also presented, including linear regression, logistic regression, NN and reinforcement learning.

In Chapter 4, a MISO secure SWIPT system with PS is considered, where the transmitter and legitimate users are equipped with multiple and single antennas, respectively. It is assumed that all the legitimate users can simultaneously process information decoding and EH. A non-linear EH model is adopted in the legitimate users to capture the characteristics of practical EH circuits. Furthermore, it is assumed that all the EVEs are purely passive, hence their channel state information (CSI) is not available at the transmitter, and artificial noise (AN) is exploited to mask the signal intended to the legitimate users. In this work, the aim is to maximize the power of AN to confuse the unknown EVEs while satisfying quality of service (QoS) and EH requirements at the legitimate users. Two robust joint beamforming and PS designs with the AN approach are developed: (1) For the bounded channel uncertainties, the problem can be transformed as a SDP through semidefinite relaxation (SDR) and a linear matrix inequality (LMI) representation; (2) For the statistical channel uncertainties, both SDR and Bernstein-type inequality are exploited to reformulate the original problem into a convex one.

Chapter 5 investigates different secrecy energy efficiency (SEE) optimization problems in a MISO underlay cognitive radio (CR) network. Different SEE maximization problems for an underlay MISO CR network with EH requirement are solved. In particular, a multi-antenna secondary transmitter (SU-Tx) simultaneously sends confidential information and energy to a secondary receiver (SU-Rx) and an ER, respectively. This secondary SWIPT communication is established by sharing the spectrum that allocated for communication for primary user terminals. Transmit beamforming design is considered to maximize the achievable SEE under the constraints of secrecy rate on the SU-Rx, interference leakage on the primary receiver (PU-Rx) as well as EH requirement on ER. Furthermore, the ER is considered to be a potential passive EVE due to the broadcast nature of wireless transmission.

The main contributions of Chapter 5 are summarized as follows: (1) With the assumption of perfect CSI at the SU-Tx, the transmit beamforming design is formulated as a SEE maximization problem with the required set of constraints. The original SEE maximization problem is not convex in nature due to its non-linear fractional objective function which becomes challenging to realize an optimal solution. To circumvent this non-convex issue, the original problem is reformulated into a tractable form by exploiting different mathematical techniques including non-linear fractional programming [45] and difference of concave (DC)

functions programming [46]. Furthermore, the optimality of the solution has been proven. (2) Next, the SEE maximization with the statistical CSI available at the SU-Tx is studied. In particular, the transmit beamforming vectors is designed to maximize the SEE, while satisfying the constraints on interference leakage, outage probability on secrecy rate and EH requirements. This problem appears to be challenging to solve due to the probabilistic constraints, which are difficult to mathematically measure in the design. These outage constraints are expressed into a set of closed-form expressions. Then, the original problem is efficiently solved through an iterative approach by exploiting different mathematical techniques including non-linear fractional and DC programming. Furthermore, the solution obtained by the proposed method always yields rank-one and shows a similar performance of the SDR approach. (3) Next, a robust design with an ellipsoidal based channel uncertainties in all set of channels is considered. This robust SEE maximization problem is non-convex in its original form and it is first reformulated into a series of SDP by employing the SDR and non-linear fractional programming [45]. However, the reformulated problem still remains non-convex due to the channels uncertainties. In order to overcome this non-convexity issue, S-Procedure [47] is exploited to convert this problem into a convex one. Moreover, the method to construct a rank-one optimal covariance matrix is proposed. This confirms the optimality of the proposed robust SEE maximization based beamforming design.

In Chapter 6, a secure transmission in a CR network is considered. This secure network consists of one primary transmitter (PU-Tx), one PU-Rx, one SU-Tx, one SU-Rx and one EVE. All of these terminals are equipped with a single antenna. The aim is to enhance the quality of secure communications between SU-Tx and SU-Rx in the presence of an EVE. In particular, the optimal power allocation is determined to maximize the achievable secrecy rate under the constraints of total transmit power of the SU-Tx and the interference leakage to the PU-Rx. The contributions to this work are: (1) firstly, none of existing research work considered of developing a NN framework to solve the secrecy rate maximization problems in an underlay CR network; (2) secondly, a NN based algorithm is proposed to simultaneously solve the secrecy rate maximization problem with perfect and imperfect CSI at the SU-Tx. The key advantage of the developed approach is that the same NN based algorithm can be employed to solve both the robust and the non-robust secrecy rate maximization problems with imperfect and perfect CSI, respectively, whereas, in conventional optimization approaches, these problems need to be formulated into two completely different optimization framework. Furthermore, to avoid over-fitting, two regularization techniques are also embedded into the proposed NN. To generate the required training set, the conventional optimization framework is utilized and then train the NN with this training set to determine appropriate weights of the connections. These weighted connections establish a mathematical relationship between the

input and the corresponding output parameters. After completing the training process, the performance of the proposed NN based approach is evaluated against the the conventional optimization approaches; (3) finally, the performance of both schemes are compared in terms of achieved secrecy rate and the required computation time to demonstrate the effectiveness and superiority of the proposed NN scheme.

Finally, Chapter 7 concludes this thesis and identifies interesting future research directions on this topic.

1.3   Publication List

Journal papers:

L. Ni, X. Da, H. Hu, Y. Huang, R. Xu and M. Zhang, "Outage Constrained Robust Transmit Design for Secure Cognitive Radio with Practical Energy Harvesting," *IEEE Access*, vol. 6, pp. 71444-71454, Dec. 2018.

M. Zhang, K. Cumanan, J. Thiyagalingam, W. Wang, A. G. Burr, Z. Ding and O. A. Dobre, "Energy Efficiency Optimization for Secure Transmission in MISO Cognitive Radio Network with Energy Harvesting," *IEEE Access*, vol. 7, pp. 126234-126252, Sep. 2019.

L. Ni, X. Da, H. Hu, M. Zhang and K. Cumanan "Outage Constrained Robust Secrecy Energy Efficiency Maximization for EH Cognitive Radio Networks", *IEEE Wireless Commun. Lett.*, minor revision.

M. Zhang, K. Cumanan, J. Thiyagalingam, L. Ni, W. Wang, A. G. Burr, Z. Ding and O. A. Dobre, "Exploiting Deep Learning for Secure Transmission in an Underlay Cognitive Radio Network," to be submitted to *IEEE Trans. Commun.*.

W. Wang, X. Li, M.Zhang, K. Cumanan and D. K. Ng, "Joint Transceiver Designs for Energy-Constrainted UAV secure Communications with Imperfect Location of Eavesdroppers", submitted to *IEEE Trans. Commun.*.

Conference papers:

M. Zhang, K. Cumanan, and A. G. Burr, "Secrecy rate maximization for MISO multi-

casting SWIPT system with power splitting scheme," in *Proc. IEEE SPAWC*, Edinburgh, Jul. 2016, pp. 1-5.

M. Zhang, K. Cumanan, and A. G. Burr, "Secure Energy Efficiency Optimization for MISO Cognitive Radio Network with Energy Harvesting," in *Proc. IEEE WCSP*, Nanjing, Oct. 2017, pp. 1-6.

M. Zhang, K. Cumanan, L. Ni, H. Hu, A. G. Burr and Z. Ding, "Robust Beamforming for AN Aided MISO SWIPT System with Unknown Eavesdroppers and Non-linear EH Model," in *Proc. IEEE GLOBECOM WORKSHOP*, Abu Dahbi, Dec. 2018, pp. 1-7.

M. Xu, W. Wang, M. Zhang, K. Cumanan, G. Zhang and Z. Ding, "Joint Optimization of Energy Consumption and Time Delay in Energy-Constrained Fog Computing Networks," in *Proc. IEEE GLOBECOM*, Hawaii, Dec. 2019, pp. 1-6.

M. Zhang, K. Cumanan, W. Wang, A. G. Burr, Z. Ding, S. Lambotharan and O. A. Dobre, "Energy Efficiency Optimization for Secure Transmission in a MIMO NOMA System," submitted to *IEEE WCNC 2020*.

Chapter 2

Fundamental Concepts and Literature Review

In this chapter, the fundamental concepts of physical layer security and SWIPT are presented. First, the basic concepts of information theoretic physical layer security is studied and the secrecy capacity of different communication systems are defined. Then, the concept of SWIPT is presented with different receiver architectures. Finally, the literature on physical layer security, SWIPT, SEE and machine learning techniques for wireless communications are reviewed which provide recent developments to address different challenges in future wireless networks.

2.1 Fundamental Concepts

This section introduces the fundamental concepts and performance metric of information theoretic physical layer security and SWIPT.

2.1.1 Physical Layer Security

Information security represents one of the major issues in wireless networks, as the signals transmitted through the wireless medium are more vulnerable for interception. Although the existing traditional security techniques may remain non-contestable, the broadcast nature of wireless communications introduces a number of challenges, particularly in terms of secret key exchange and distributions [32, 22, 48, 11]. To further enhance the security of
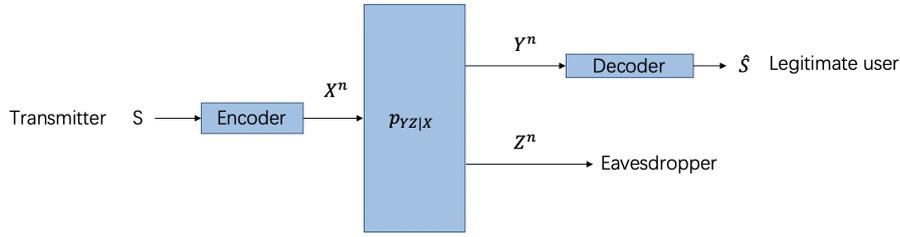
Fig. 2.1 The wiretap channel with confidential information.

wireless networks, information theoretic based physical layer security has been recently proposed to complement the conventional security techniques in wireless transmissions. In particular, this approach exploits the difference between the legitimate and wiretap channels in establishing secure wireless transmission [49]. In other words, physical layer security can only be implemented when the SNR or SINR of the legitimate channel is better than that of the wiretap channel [25, 26, 50]. In contrast to the conventional security methods, physical layer security provides a different paradigm, which is named as secrecy rate and it can be implemented by exploiting the physical layer characteristics of wireless channels [51]. This approach was first theoretically proved by Shannon [25], and then secrecy capacity of wiretap and related channels were developed by Wyner [26] and Csiszar [50]. In recent years, physical layer security has been recognized as a promising technique to establish confidential communications between transmitters and legitimate users [11, 22, 52–55, 18, 56–58].

To investigate information theoretic physical layer security, a basic model, namely, the *wiretap channel* is illustrated in Fig. 2.1 [26]. This secrecy system was initially proposed in the literature [26], and then developed in [50]. Three terminals are considered in this model: one transmitter, one legitimate user and one EVE. The transmitter intends to send a private message set S to the legitimate user at rate R, and wants the message to be confidential. The message set is defined as $S = \{1, 2, ..., 2^{nR}\}$, and the entropy is $H(S) = nR$, where R is the rate of communication and n is the block length of communication. The encoding function $f_n : S \to X^n$ maps each message $s \in S$ to a codeword $x^n \in X^n$. It is assumed that the transmitter communicates with the legitimate user through a discrete memoryless channel, which outputs two discrete sequences Y^n and Z^n at the legitimate user and the EVE, respectively. The transition probability is defined as $p_{YZ|X}$. At the legitimate user, the decoding function $g_n : Y^n \to S$ maps the observation Y^n to an estimated message $\hat{S}$. The quality of security is measured by the entropy rate with respect to (w.r.t.) the observation of EVE, which is also known as *equivocation rate*:

$$R_e = \frac{1}{n}H(S|Z^n). \tag{2.1}$$

The rate R can be called the *secrecy rate*, if and only if for any $\epsilon > 0$, there exists the code sequence $(2^{nR}, n)$ such that

$$Pr[\hat{S} \neq S] < \epsilon, \tag{2.2}$$

$$\frac{1}{n}H(S|Z^n) \geq R - \epsilon. \tag{2.3}$$

The inequalities defined in (2.2) and (2.3) are the reliability and secrecy conditions for this system, respectively. Let the symbol $I(\cdot; \cdot)$ denote the mutual information, one can have

$$\frac{1}{n}I(S; Z^n) = \frac{1}{n}H(S) - \frac{1}{n}H(S|Z^n) = R - R_e \leq \epsilon, \tag{2.4}$$

which states that the EVE cannot decode anything. Hence, the expression for the secrecy rate for a discrete memoryless channel can be written as [50]:

$$C = \max_{U \to X \to YZ}[I(U; Y) - I(U; Z)], \tag{2.5}$$

where U is an auxiliary variable that should meet the Markov chain condition, i.e., $p_{YZ|UX} = p_{YZ|X}$. Note that the auxiliary variable U is introduced to represent the portion of source message corresponding to certain decodability level at the legitimate user [58].

Consider a SISO network as shown in Fig. 2.2, with one transmitter, one legitimate user and one EVE. All of these terminals are equipped with a single antenna. The transmitter intends to send a confidential message to the legitimate user while the EVE attempting to intercept the confidential message. The channels between the transmitter and the legitimate user as well as the EVE are denoted by h_s and h_e, respectively. The noise at the legitimate user and the EVE are denoted by n_s ($\mathbb{E}\{|n_s|^2\} = \sigma_s^2$) and n_e ($\mathbb{E}\{|n_e|^2\} = \sigma_e^2$), respectively. Furthermore, assume that P is the transmit power of the transmitter. The received signal at the legitimate user and the EVE can be mathematically expressed respectively, as

$$y_s = \sqrt{P}h_s x + n_s, \; y_e = \sqrt{P}h_e x + n_e, \tag{2.6}$$

where x ($\mathbb{E}\{|x|^2\} = 1$) denotes signal sent from the transmitter. Based on these assumptions, the achievable secrecy rate of the legitimate user is written as [59]

$$C = \left[\log_2\left(1 + \frac{|h_s|^2 P}{\sigma_s^2}\right) - \log_2\left(1 + \frac{|h_e|^2 P}{\sigma_e^2}\right)\right]^+, \tag{2.7}$$

where $[x]^+$ denotes $\max[0, x]$.
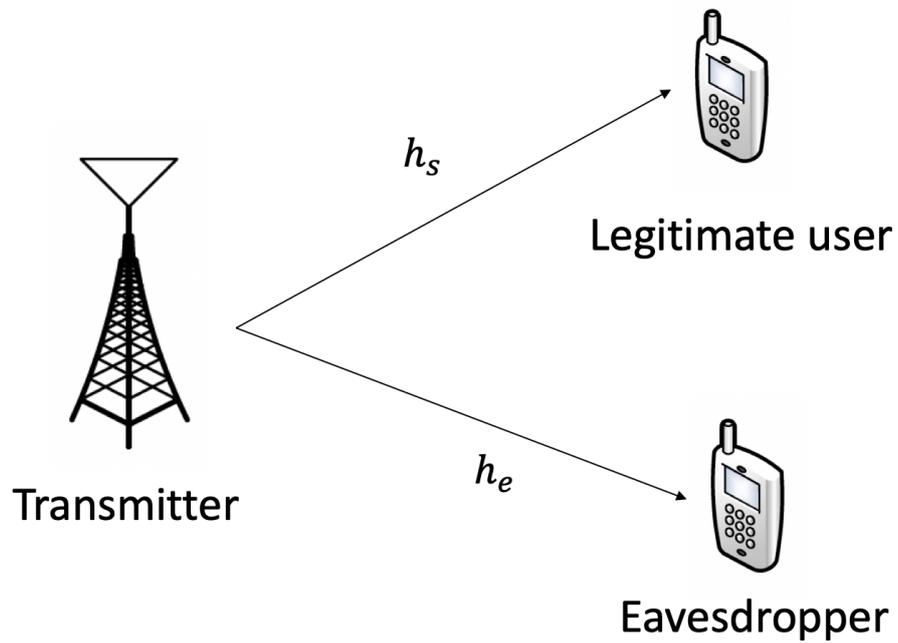
Fig. 2.2 A SISO system with one transmitter, one legitimate user and one EVE.
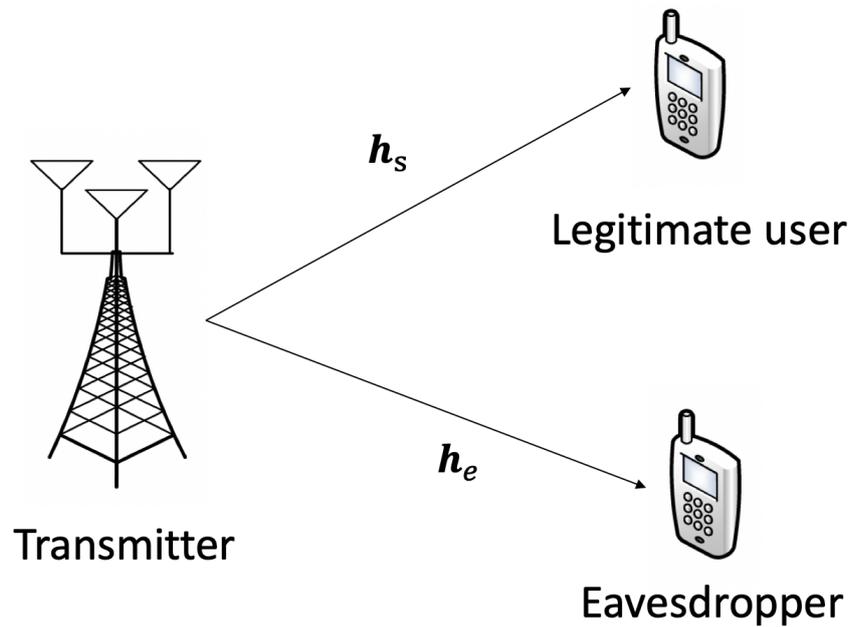
Fig. 2.3 A MISO system with one multi-antenna transmitter, one legitimate user and one EVE.

Next, the achievable secrecy rate of MISO system is investigated. A simplified MISO system model is shown in Fig. 2.3, where a transmitter with N_T antennas intends to send a confidential message to the single antenna legitimate user. At the same time, a single antenna EVE attempts to eavesdrop the information. The channels between the transmitter and the legitimate user and EVE are denoted as $\mathbf{h}_s \in \mathcal{C}^{N_T \times 1}$ and $\mathbf{h}_e \in \mathcal{C}^{N_T \times 1}$, respectively. The parameters $n_s \sim \mathcal{CN}(0, \sigma_s^2)$ and $n_e \sim \mathcal{CN}(0, \sigma_e^2)$ represent the additive white Gaussian noise (AWGN) at the legitimate user and the EVE, respectively. The received signals at the legitimate user and the EVE can be written respectively, as

$$y_s = \mathbf{h}_s^H \mathbf{q} x + n_s, \qquad (2.8)$$

$$y_e = \mathbf{h}_e^H \mathbf{q} \mathbf{x} + n_e, \qquad (2.9)$$

where $\mathbf{q} \in \mathcal{C}^{N_T \times 1}$ denotes the signal vector transmitted from the transmitter and x ($\mathbb{E}\{|x|^2\} = 1$) denotes signal sent from the transmitter, respectively. Thus, the achievable secrecy rate of the legitimate user is defined as [59]

$$C = \left[\log_2(1 + \frac{1}{\sigma_s^2} \mathbf{h}_s^H \mathbf{q}\mathbf{q}^H \mathbf{h}_s) - \log_2(1 + \frac{1}{\sigma_e^2} \mathbf{h}_e^H \mathbf{q}\mathbf{q}^H \mathbf{h}_e) \right]^+ . \qquad (2.10)$$

2.1.2 Simultaneous Wireless Information and Power Transfer

The power is extremely valuable resource and should be efficiently used in wireless communication systems, especially for the batteries powered energy-constrained devices and networks [27]. The slow improvement of battery technologies introduces a critical challenge for designing mobile devices [32], for example, the basic requirement is to design mobile devices much lighter and thinner, however, these designs require ultra high energy density batteries, which are not possible under current technical conditions. Furthermore, these increasing requirements in wireless devices violate one of the key demanding requirements of the 5G networks which is to achieve green and low-carbon communications. WPT is an emerging technology, which facilitates the mobile devices to collect energy from external energy sources without any wired connections [30, 31, 60]. The energy source to enable this technique can be the information carrying RF signals radiated by the transmitters [30, 31, 60]. As such, SWIPT facilitates practical design and implementation of this technique, especially in mobile devices [12, 61].

Two types of receiver designs can be employed in practical SWIPT systems, namely, separated and co-located IR and ER. A simplified MISO SWIPT network is shown in Fig. 2.4, where the IR receives the information whereas ER harvests energy from the transmitter.
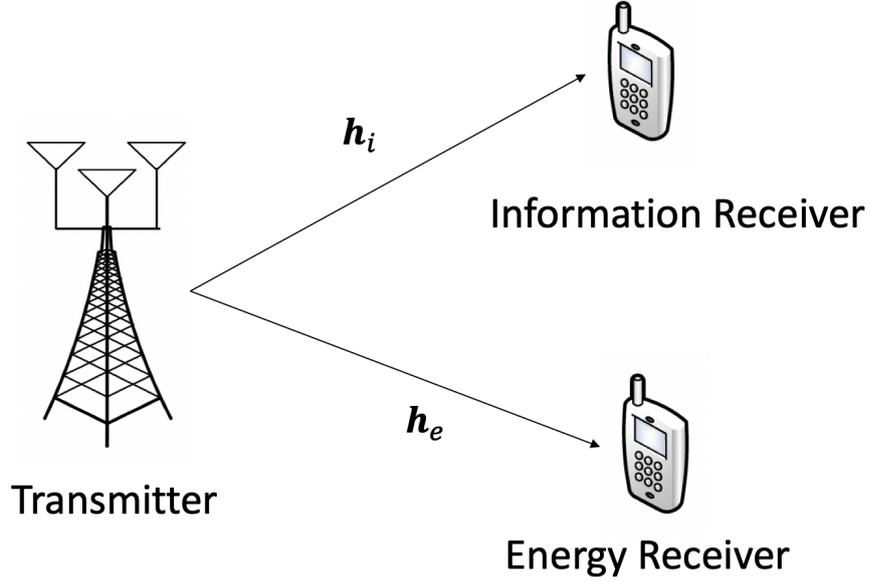
Fig. 2.4 The system model of a MISO SWIPT system with separated IR and ER.

It is assumed that the transmitter is equipped with N_T antennas, while both the IR and ER are equipped with a single antenna. The channels between the transmitter and the IR as well as the ER are denoted as $\mathbf{h}_i \in \mathcal{C}^{N_T \times 1}$ and $\mathbf{h}_e \in \mathcal{C}^{N_T \times 1}$, respectively. The received signal at the ER can be written as

$$y_e = \mathbf{h}_e^H \mathbf{q} x + n_e, \tag{2.11}$$

where $\mathbf{q} \in \mathcal{C}^{N_T \times 1}$ denotes the transmit beamforming vector, x $(\mathbb{E}\{|x|^2\} = 1)$ denotes signal sent from the transmitter and $n_e \sim \mathcal{CN}(0, \sigma_e^2)$ represents the noise at the ER. The received RF power can be expressed as

$$P_e = \mathbf{h}_e^H \mathbf{q} \mathbf{q}^H \mathbf{h}_e + \sigma_e^2, \tag{2.12}$$

Furthermore, the harvested energy at the ER can be written as

$$E_e = \zeta_e(\mathbf{h}_e^H \mathbf{q} \mathbf{q}^H \mathbf{h}_e + \sigma_e^2), \tag{2.13}$$

where $\zeta_e \in (0, 1]$ represents the EH efficiency of the ER.

Another simplified MISO SWIPT communication scenario is presented in Fig. 2.5, where the same receiver can perform both information decoding and EH. The transmitter

is equipped with N_T antennas whereas the receiver consists of only a single antenna. The channel between the transmitter and the receiver is denoted as $\mathbf{h}_i \in \mathcal{C}^{N_T \times 1}$. In co-located receiver structure, there are two types of designs available in the literature [27], namely, PS and TS. In this thesis, the PS approach is considered, where a power splitter is embedded into the receiver. The structure of a PS based receiver design is shown in Fig. 2.6. The functions of a PS based receiver are demonstrated as follows: First, the received RF signal is corrupted by an AWGN $n_i \sim \mathcal{CN}(0, \sigma_i^2)$ at the RF band. This RF band signal is then forwarded to the power splitter. Next, the power splitter separates the signal into two streams with the PS ratio $\rho \in (0, 1]$. After the splitter, ρ part of the signal power is fed to the ER and the left $1 - \rho$ part of the signal power is forwarded to the information decoder. Furthermore, the signal split to the information decoder is first converted to a complex baseband signal and then been sampled and digitalized by an analog-to-digital converter for further decoding [62, 63], where an additional baseband noise is introduced by this conversion. This additional noise can be modelled as $n_p \sim \mathcal{CN}(0, \sigma_p^2)$. Hence, the received signal at the IR and the ER can be expressed, respectively, as

$$y_{IR} = \sqrt{1 - \rho}(\mathbf{h}_i^H \mathbf{q} x + n_i) + n_p, \qquad (2.14)$$

$$y_{ER} = \sqrt{\rho}(\mathbf{h}_i^H \mathbf{q} x + n_i), \qquad (2.15)$$

where $\mathbf{q} \in \mathcal{C}^{N_T \times 1}$ denotes the transmit beamforming vector and where x $(\mathbb{E}\{|x|^2\} = 1)$ denotes signal sent from the transmitter. The received RF power at the ER can be written as

$$P_i = \rho(\mathbf{h}_i^H \mathbf{q} \mathbf{q}^H \mathbf{h}_e + \sigma_i^2). \qquad (2.16)$$

On the other hand, the harvested energy can be written as

$$E_i = \zeta_{eh} \rho(\mathbf{h}_i^H \mathbf{q} \mathbf{q}^H \mathbf{h}_e + \sigma_i^2), \qquad (2.17)$$

where $\zeta_{eh} \in (0, 1]$ represents the EH efficiency.

Recently, more realistic and practical non-linear EH models have been considered in the literature [64–66], which can take into account the non-linear characteristics in practical RF-EH circuits. In linear EH models, the output power can be defined as the product of the received RF power and the EH efficiency ratio. However, non-linear EH models take into account more realistic parameters. The output power of a non-linear EH model can be expressed as [64]

$$E_{out} = \frac{\Psi - M\Omega}{1 - \Omega}, \qquad (2.18)$$
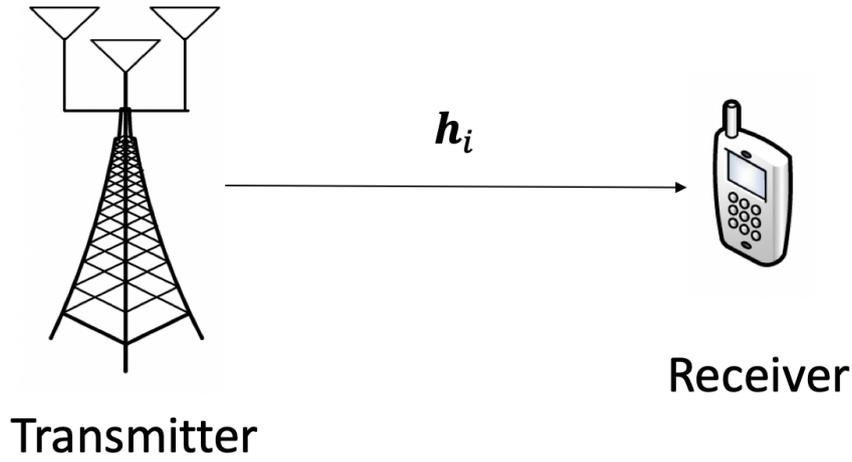
Fig. 2.5 The system model of MISO SWIPT system with co-located IR and ER.
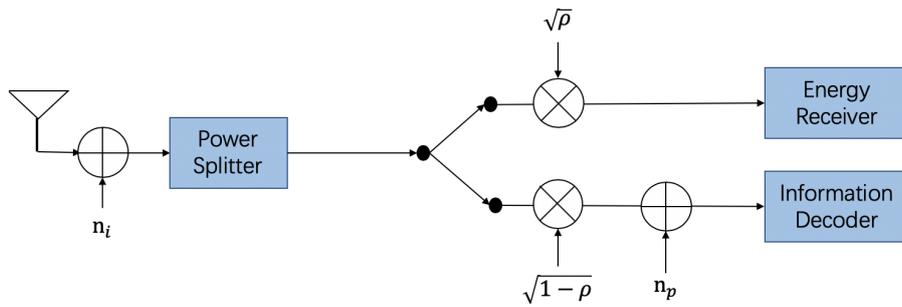
Fig. 2.6 The structure of a PS receiver.

where $\Omega = \frac{1}{1+\exp(ab)}$. The M is a constant that denotes the maximum power that the ER can harvest, this is due to the fact that the RF-EH circuit saturates if the received RF power is exceedingly large. The parameter Ψ is defined as $\Psi = \frac{M}{1+\exp(-a(P_{in}-b))}$, where P_{in} is the input RF power of the EH circuit which is defined in (2.12) for the case of separated IR and ER, and in (2.16) for the scenario of co-located IR and ER with PS scheme. Furthermore, a is the non-linear charging rate w.r.t. the input power and b is a minimum turn-on voltage of the EH circuit [64].

2.2 Literature Review

This section presents literature review on physical layer security, SWIPT, energy efficient strategies and machine learning for wireless communications under each corresponding subsection.

2.2.1 Physical Layer Security

Over the last decade, researchers developed several approaches to improve achievable secrecy rate, for example, relay aided system, cooperative jamming, AN approach and device to device (D2D) communication systems, etc [32, 52, 67–81]. Convex optimization and multi-antenna techniques have been exploited to design the efficient transmit beamforming by solving secrecy rate maximization or transmit power minimization problems. The beamforming design for physical layer security over $\eta - \mu$ fading channels with and without co-channel interference, has been investigated in [67]. Note that the $\eta - \mu$ fading channel model was first proposed in[82], which is modelled as following the $\eta - \mu$ distribution. This channel model can be used to better represent the the small-scale variations of the fading signal in a non-line-of-sight condition [67]. The physical layer security performance of a relay aided full duplex network is studied in [81]. In [68], the sum secrecy rate maximization problem is studied in a relay assisted CR network. An optimal power allocation scheme in a decode-and-forward scenario has been proposed to maximize the sum secrecy rate in [70]. The dual antenna selection has been studied to improve secrecy performance and transmission capacity for a full duplex CR network in [72]. A SDP based transmission technique for MISO secure channels has been proposed in [69]. In [71], a joint source and relay beamforming design is proposed for an amplify-and-forward assisted relay multiple-input multiple-output (MIMO) system to maximize the secrecy rate. Generalized singular value decomposition (GSVD) has been utilized in [73] to determine the optimal power allocation.

Cooperative jamming has been considered as an alternative approach to further improve quality of secure transmissions [32, 52, 74–76]. The aim of any types of jamming signals is to generate more interference to the potential EVEs [52, 76], which in turn will improve the available secrecy rate of the communication system. The optimal cooperative jamming design is studied in [74] to enhance the secrecy performance of a relay assisted wireless network. The secrecy rate maximization problem has been investigated in a MIMO wireless network with a cooperative jammer in [75]. In [52], a novel cooperative jamming approach, namely harvest-and-jam helper, is proposed for an amplify-and-forward relay assisted wireless network to

maximize the secrecy rate, whereas a multi-antenna cooperative jammer is employed in a MIMO network to improve the secrecy performance. Furthermore, a wireless powered D2D secure transmission is studied in [76], where a cooperative jammer is employed to introduce interference to the EVEs.

AN is another type of jamming signal that introduces more interference to EVEs by embedding it in the transmit signal [78, 79]. An isotropic AN approach has been proposed in [78] by exploiting orthogonal projection of the intended signal to the legitimate users, whereas the secrecy rate maximization problem is investigated in [79] by jointly designing transmit beamforming and AN covariance matrix. The joint signal and AN beaforming design is considered to maximize the secrecy rate in a multicast MISO SWIPT system in [11]. In [83], the AN power is maximized to interfere with purely unknown EVEs in a MISO SWIPT network. In [84], the joint beamforming and PS design is studied to minimize the total power in an AN aided secure MIMO SWIPT system. Furthermore, the harvested energy is maximized in a secure MISO SWIPT system with the help of both AN and cooperative jamming in [32].

Generally, the CSI of all the legitimate channels and wiretap channels are assumed to be perfectly known at the transmitter [11, 85, 86]. However, there are practical difficulties in having the perfect CSI at the transmitter due to the channel estimation and quantization errors. Ignoring the effect of CSI uncertainties in the design for wireless networks can lead to solutions that may violate critical constraints and results in a poor performance in realistic channel conditions. Therefore, it is important to incorporate the channel uncertainties in the beamforming designs. Furthermore, it is more challenging for the transmitter to design the beamforming in different types of secrecy rate optimizations when the transmitter cannot obtain the perfect knowledge of CSI.

To circumvent the inevitable channel uncertainties, robust design is a well-known approach, which can be classified into two groups, the worst-case robust design [69, 87–94], and the outage probability based design [95–98]. In the worst-case design, it is assumed that the CSI errors belong to some known bounded uncertainty sets whereas in the outage probability based design, the channel errors are random with a certain statistical distribution and constraints can be satisfied with certain outage probabilities. In fact, the bounded robust optimization is generally conservative owing to its worst-case criterion while probabilistic SINR constrained beamforming provides a soft SINR control. The robust transmit covariance matrix design for a MISO communication system with multiple antenna EVEs has been investigated in [69], where SDR and S-Procedure are employed to solved this robust problem. In [99], the robust transmit covariance matrix design has been proposed to minimize the

EVE's SINR in a MIMO system. The work in [100] considers the robust beaforming design for maximizing the secrecy rate in a CR MISO network. The work in [22] solves a robust optimization problem in a MIMO system at low SNRs in which the original problem is formulated as a SDP by employing S-Procedure. AN aided robust optimization designs have been investigated in [79, 99]. Furthermore, the outage probabilistic based robust designs have been considered in the literature when the transmitter has only the statistical knowledge of channel uncertainties. In this approach, the required quality of services are satisfied at the users with a set of predefined outage probabilities [18, 101]. In [18], a robust outage optimization problem for MIMO system has been studied, in which a Bernstein-type inequality based conservative approximation approach has been utilized to formulate the non-convex outage secrecy rate constraint. The work in [101] developed a robust design in a MISO system, in which the outage probability of the secrecy rate is minimized by employing a single-stream beamforming and transmitting AN in the null space of the legitimate channel.

2.2.2   Simultaneous Wireless Information and Power Transfer

Recently, SWIPT over wireless channels has been studied extensively in the literature [30, 31, 27, 54, 53, 63, 102–106, 33]. Varshney first proposed the idea of transmitting information and power at the same time in [30]. A capacity energy function is presented to characterize the fundamental trade-off between the rates at which energy and reliable information can be transmitted. In [31], Grover and Sahai extended the work in [30] to frequency selective channels with AWGN. It was shown in [31] that a non-trivial trade-off exists for information transfer versus energy transfer via power allocation. Wireless information and power transfer subject to co-channel interference was studied in [102], in which optimal designs to achieve different different outage energy trade-offs as well as rate-energy trade-offs are derived. Different from the traditional view of taking interference as an undesired factor that jeopardizes the wireless channel capacity, in [102] interference was utilized as a source for EH. Unlike [30, 31, 102], which considered point to point single antenna transmission, [27, 107] considered MIMO systems for SWIPT. In particular, the work in [27] studied the performance limits of a three node MIMO broadcasting system, where one receiver harvests the energy and another receiver decodes information from the signals sent by a common transmitter. The work in [107] extended the scheme in [27] by considering imperfect CSI at the transmitter.

In [53], a secure SWIPT transmission system with two type of EVEs (passive and potential EVEs) coexist has been firstly proposed. The transmit beamforming, AN beamforming and energy beamforming have been jointly designed to minimize the required total transmit power.

The work in [103] considered resource allocation design for maximizing the harvested energy in a MISO SWIPT system with multiple-antenna ERs. In addition, the authors presented that there always exists a rank-one solution of the optimal transmit covariance matrix and proposed an algorithm to construct an equivalent rank-one solution.

The CSI in [27, 103] was assumed to be perfectly known at the transmitter, however, it is difficult to obtain the CSI in practical schemes due to the channel estimation and quantization errors. The work in [54, 53, 104, 105] investigated the robust secure communication with MISO SWIPT systems by incorporating all the channel uncertainties. SDR has been employed to solve the secrecy rate maximization problems in [104, 105]. In [104], the authors proposed a suboptimal method to guarantee the rank-one solution for the relaxed problem, whereas in [105], the optimal solution of the proposed problem termed to be rank-two. Furthermore, a MISO SWIPT system was studied in [33], a two step algorithm was proposed to circumvent the rank-one issue, while SDR is exploited to obtain a rank-one solution in the AN added scheme. In [106], the resource allocation in AN added orthogonal frequency division multiple access (OFDMA) SWIPT system has been studied, in which a weighted sum secrecy rate maximization problem is investigated for the IRs and to satisfy the harvested energy requirements at the ERs. SWIPT for multi-user systems was studied in [108], where it was shown in [108] that for multiple access channels with a received energy constraint, time-sharing is necessary to achieve the maximum sum rate when the received energy constraint is sufficiently large; while for the multi-hop channel with a harvesting relay, the transmission strategy depends on the quality of the second link. In [109], a robust EE design for a relay embedded MIMO system with SWIPT is studied. The secure beamforming design for a cooperative MISO CR non-orthogonal multiple access network with SWIPT has been investigated in [110].

Most existing works that adopt the linear EH model assume that the output direct current power is independent of the input power. However, in practice, the EH circuit results in a non-linear power conversion due to the rectifier in the RF-EH conversion circuit which is the key element in WPT implementation [111]. Therefore, the assumption of a linear EH model in the literature may not be able to incorporate the non-linear characteristics in practical EH scenarios [112]. Recently, the practical parametric non-linear EH model has been considered in [112]. In contrast to the linear EH model, the non-linear one includes the characteristics of practical RF-EH conversion circuits. A power allocation strategy for a SWIPT system with non-linear EH model is proposed in [112], whereas beamforming designs are investigated with a non-linear EH model for multiple antenna systems in [64, 65, 113, 114].

2.2.3 Energy Efficient Strategies

The unprecedented requirements of higher data rate, EE and information security pose different challenges in the development of future wireless communication systems and networks. Furthermore, the exponential growth of the number of wireless devices with different newly emerging high data rate applications introduce major concerns on the energy consumption with different impacts in terms of both environmental and economic aspects [115–118, 117, 118]. Different statistics confirm that information communication technologies based industries contribute to two percent of the world energy consumption [119], and this will grow rapidly with the ever increasing number of wireless devices and the emerging new communication technologies. This exponential growth in energy consumption will result in different undesirable impacts on the natural green environments. Hence, energy efficient communications have recently become as a demanding approach to address not only these issues but also due to the limited battery life of mobile devices and slow development of energy storage technologies. Furthermore, the energy efficient designs based on the EE performance metric have been considered as one of the key requirements in the development of future wireless systems. These designs take the EE performance metric into account rather than achievable rate or transmission power metrics.

Most of existing work on physical layer security in the literature considers either secrecy rate maximization with a total transmit power constraint [11, 120, 121] or power minimization to meet the secrecy rate requirements [33, 18, 122]. However, these designs do not take into account the SEE as the power consumptions and their impacts become as major concerns in future secure communications. Therefore, from the energy efficient perspective in physical layer security, the SEE is considered as a suitable performance metric to measure the efficient utilization of the power consumption in a secure communication system. The SEE is defined as the ratio between the achieved secrecy rate and the total power consumption at the transmitter. Several research work on SEE has been investigated in the literature [123–127, 12]. In [123], a transmitter design for maximizing SEE has been investigated in a MISO system with imperfect CSI at the transmitter, whereas a SEE maximization design has been considered for a MISO system under both statistical and imperfect CSI assumptions in [124]. The transmit beamforming design has been proposed to maximize the SEE in a MISO system with the assumption of having statistical CSI at the transmitter. The robust SEE maximization for MIMO SWIPT system has been studied in [126]. Furthermore, in [127] the SEE maximization problem with integrated services (one multicast and one confidential message) in a MISO system has been considered. In [12], the transmit beamforming design is studied to maximize the SEE for an underlay MISO CR SWIPT network.

2.2.4   Machine Learning for Wireless Communications

The explosive growth of complexity and diversity of wireless communication systems and related applications in recent years poses different challenges to mobile service providers [128]. As an emerging solution to address these issues, machine learning embedded wireless communications has drawn a great deal of research interest from both academia and relevant industries due to their benefits of less computational complexity and lower computation time [129–131]. Machine learning based approaches can be applied to solve a wide range problems in wireless communications, form radio access technology selection [132] to different resource optimization problems [61], as well as channel estimation and signal detection problems [133]. In the following, literature review of machine learning for wireless communications is provided.

In the literature, many research work have demonstrated that machine learning techniques can be exploited to solve different types of complicated problems in different domains of wireless communications. The possibility of employing machine learning in 5G networks has been discussed in [129], including embedding machine learning to massive MIMO and smart grid networks. In [134], the opportunities and challenges of combining AI and future mobile networks have been reviewed. The NN based spectrum efficiency and EE maximization techniques have been proposed for CR network in [61]. In [135], the advantages and disadvantages of different learning algorithms have been discussed for self-organizing networks (SONs). Deep learning based channel estimation and signal detection techniques in orthogonal frequency division multiplexing (OFDM) systems have been investigated in [133]. The reinforcement learning based framework and several conventional techniques have been investigated for traffic offloading problems in a stochastic heterogeneous cellular network in [136]. A learning based approach for wireless resource management has been presented in [137] whereas a reinforcement learning based resource allocation technique has been developed for vehicle to vehicle communications in [138]. A big data driven mobile network optimization and several techniques in analytics have been presented in [130]. The authors have provided an overview on Apache Spark (an open-source platform for cluster computing) based learning framework to analyze the mobile big data in [139], which has exploited the real world dataset to validate their proposed approaches. In [140], a survey on deep learning applications for network traffic control systems is provided. Furthermore, this work also discusses several unsolved issues which may be potential avenues for future research work.

The work in [141] introduces the trend in wireless networks to incorporate NN techniques to cope with the challenges associated with different design problems. In this work, an

overview on NN models, as well as the survey of NN applications in wireless networks have been presented. An overview on exploiting different advanced deep learning techniques for IoT data analytics is presented in [142]. Furthermore, this survey also highlights some research challenges and potential directions for future work. The feasibility and advantages of applying deep learning in mobile sensing is discussed in [143]. A comprehensive survey of the applications of deep learning techniques for wireless networks is provided in [144]. Deep learning applications in different layers in a communication system have been investigated, including modulation, coding, resource allocation, path search, and so on. Furthermore, some unsolved research issues are discussed in detail to highlight different avenues for future research. A survey of deep learning applications in mobile and wireless networks is presented in the literature [128], including mobile data analysis, user localization, network control, signal processing and network security.

Recently, machine learning approaches based signal processing has drawn a significant attention from relevant research communities [145–151]. In [145], a deep NN based framework is proposed for signal detection problems in a MIMO system, where the simulation results demonstrate that the proposed NN can offer high accuracy with low complexity. The convolutional NN has been exploited to achieve higher accuracy of user positioning in a massive MIMO system in [147]. The transmit power optimization and the inter-cell interference cancellation problems have been solved by exploiting a non-iterative NN in [146]. Furthermore, deep learning algorithms embedded physical layer designs are investigated in [148], in which unsupervised deep auto-encoder network is exploited to optimize the encoding and decoding processes in a single user MIMO network over a Rayleigh fading channel. In [149], deep learning has been utilized to recover the sparse signal with better accuracy from noisy linear measurements in MIMO systems. A recurrent NN approach is also exploited for a modulation recognition task in [150]. In [151], several deep learning implementations have been presented for modulation classification problems in an end-to-end communication system, where several challenges and avenues for future directions are also identified.

2.3   Summary

In this chapter, the fundamental concepts of information theoretic physical layer security have been provided. Furthermore, the achievable secrecy rate of SISO and MISO secure communication systems have been defined. Then, the underlying concepts and practical designs of a SWIPT system have been presented. Finally, the literature review on physical

layer security, SWIPT, energy efficient transmission designs and machine learning based resource allocation techniques has been provided. Through reviewing the literature on physical layer security and SWIPT, the research gaps in these areas were explored, which motivates the work in Chapter 4. Next, the literature review on energy efficient transmission designs arises research interests of combing physical layer security and energy efficient for CR networks. Furthermore, recent work on machine learning based resource allocation techniques pointed out a promising research direction for future wireless networks.

Chapter 3

Mathematical Background

In this Chapter, the fundamental concepts of convex optimization and machine learning techniques are provided. In particular, the first section introduces the fundamentals of convex optimization, including convex sets, cones, functions and different types of convex optimization problems. In the second section, several machine learning approaches are presented, such as linear regression, logistic regression, NN and reinforcement learning.

3.1 Convex Optimization

Convex optimization is a special set of mathematical optimization techniques [152], which have been widely used to solve numerous problems in wireless communications and signal processing [75, 22, 153, 152, 85, 54]. Most of these problems are formulated into a constrained optimization framework, which can be either naturally convex or can be reformulated into a convex form through some mathematical manipulations [154, 155]. Standard convex optimization problems can be solved efficiently to determine the optimal solution through interior point methods [152, 156]. In particular, convex optimization techniques have also influenced many practical engineering implementations, such as computer designs, analysis tools, real-time and automatic control systems, as it validates the optimality of the solution through examining KKT conditions and duality gaps [152]. It is worthy to point out that determining whether the optimization problem is convex or not is a crucial step that has to be carried out prior to use any convex optimization techniques to solve the problem.

3.1.1 Fundamental of Convex Optimization

This section provides the fundamentals of convex optimization.

3.1.1.1 Convex Sets

A set $S \in \mathbb{R}^n$ can be classified as a convex set if $\mathbf{x}, \mathbf{y} \in S$ and $\forall \theta \in [0, 1]$, the following relation holds [152]:

$$\theta \mathbf{x} + (1 - \theta)\mathbf{y} \in S. \tag{3.1}$$

In a convex set, the line segment between any two points from the set should be in the same set.

3.1.1.2 Convex Cones

A set C is a convex cone, if C is convex and for every $\mathbf{x} \in C$, $\theta_1 \geq 0$, and $\theta_2 \geq 0$ and $\alpha \geq 0$, $a\mathbf{x} \in C$ holds [152]. This can be mathematically expressed as

$$\theta_1 \mathbf{x} + \theta_2 \mathbf{y} \in C, \ \mathbf{x}, \mathbf{y} \in C. \tag{3.2}$$

Convex cones result in various forms of engineering applications, in which the most common convex cones are given as

1. Non-negative orthant: $\mathbb{R}^n_+$.

2. Second order cone (SOC): $C = \{(t, \mathbf{x}) \ t \geq ||\mathbf{x}||\}$.

3. Positive semidefinite matrix cone: $C = \mathbb{S}^n_+ = \{\mathbf{Y} | \mathbf{Y} \text{ is symmetric and } \mathbf{Y} \geq 0\}$.

3.1.1.3 Convex Functions

A function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is convex if dom $f(\mathbf{x})$ is a convex set and if for all $\mathbf{x}_1, \mathbf{x}_2 \in$ dom $f(\mathbf{x})$ the following inequality holds [152]:

$$f(\theta \mathbf{x}_1 + (1 - \theta)\mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1 - \theta)f(\mathbf{x}_2), \quad \forall \theta \in [0, 1]. \tag{3.3}$$

In other words, the line segment between $(\mathbf{x}_1, f(\mathbf{x}_1))$ and $(\mathbf{x}_2, f(\mathbf{x}_2))$ lies above the graph of $f(\mathbf{x})$. The function $f(\mathbf{x})$ is concave if $-f(\mathbf{x})$ is convex. Suppose $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is

continuously differentiable, $\mathbf{x}, \mathbf{y} \in$ dom $f(\mathbf{x})$ and dom $f(\mathbf{x})$ is convex. Then $f(\mathbf{x})$ is convex if the flowing inequality holds

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \tag{3.4}$$

Furthermore, if $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, then $f(\mathbf{x})$ is convex if the flowing inequality holds

$$\boldsymbol{\nabla}^2 f(\mathbf{x}) \succeq \mathbf{0}, \quad \forall \mathbf{x} \in \mathbb{R}^n. \tag{3.5}$$

The inequality in (3.5) indicates that the convexity of $f(\mathbf{x})$ can be given by demonstrating its Hessian is positive semidefinite on its domain [152]. Based on the above, a linear function is always convex, while a quadratic function $\mathbf{x}^T \mathbf{P} \mathbf{x} + \mathbf{a}^T \mathbf{x} + b$ is convex if and only if $\mathbf{P} \succeq \mathbf{0}$.

3.1.2   Convex Optimization Problems

The standard mathematical formulation of a convex problem that finds an $\mathbf{x}$ through minimizing $f_0(\mathbf{x})$ with all the conditions satisfied can be defined as the following form [152]:

$$\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & f_0(\mathbf{x}) \\
\text{s.t.} \quad & f_i(\mathbf{x}) \leq 0, \quad i = 1, ..., m, \\
& h_i(\mathbf{x}) = 0, \quad i = 1, ..., p,
\end{aligned} \tag{3.6}$$

where the vector $\mathbf{x} \in \mathbb{R}^n$ contains the *optimization variables* of the problem, and the function f_0 is the *objective function* or *cost function*. The inequalities $f_i(\mathbf{x}) \leq 0, \ i = 1, ..., m$ are called the *inequality constraints*, and $h_i(\mathbf{x}) = 0, \ i = 1, ..., p$ are called the *equality constraints*. Furthermore, the problem is known as an *unconstrained problem* if there is no constraints. The domain of the problem (3.6) is a set of points, for which the objective and the constraints are defined and is denoted as

$$D = \bigcap_{i=0}^{m} \text{dom} f_i \cap \bigcap_{i=0}^{p} \text{dom} h_i. \tag{3.7}$$

In other words, if a point $\mathbf{x} \in D$ satisfies all the constraints $f_i(\mathbf{x}) \leq 0 \quad i = 1, ..., m$ and $h_i(\mathbf{x}) = 0 \quad i = 1, ..., p$, then it is a feasible point. Thus, the problem (3.6) is said to be *feasible* if there exists a feasible point, and otherwise it is *infeasible*. The $\mathbf{x}^*$ is optimal if and

only if

$$f_0(\mathbf{x}^*) \leq f_0(\mathbf{x}), \quad \forall \mathbf{x} \in D. \tag{3.8}$$

A problem can be classified as a convex optimization problem if the following conditions are satisfied: (1) the objective function must be convex; (2) the inequality and the equality constraint functions must be convex and affine, respectively [152]. Note that an affine function is a function that composed of a linear function and a constant, i.e., $\mathbf{a}^T\mathbf{x} = b$. In the following, general forms of the canonical optimization problem formulations are provided.

3.1.2.1 Linear Programming

When the objective and constraint functions are all affine, then the convex optimization problem is known as a LP [152]. A standard form of LP can be defined as follows:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x} + d \\ s.t. \quad & \mathbf{Gx} \preceq \mathbf{h}, \\ & \mathbf{Ax} = \mathbf{b}, \end{aligned} \tag{3.9}$$

where $\mathbf{G} \in \mathbb{R}^{m \times n}$ and $\mathbf{A} \in \mathbb{R}^{p \times n}$.

3.1.2.2 Quadratic Programming

The convex optimization problem is called a QP, when the objective function is quadratic and the constraint functions are affine. A QP takes the following form [152]:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^T\mathbf{Px} + \mathbf{q}^T\mathbf{x} + r \\ s.t. \quad & \mathbf{Gx} \preceq \mathbf{h}, \\ & \mathbf{Ax} = \mathbf{b}, \end{aligned} \tag{3.10}$$

where $\mathbf{P} \in \mathbb{S}_+^n$, $\mathbf{G} \in \mathbb{R}^{m \times n}$ and $\mathbf{A} \in \mathbb{R}^{p \times n}$. In QP, a convex quadratic function is minimized over a polyhedron. LP is a special case of QP, which can be obtained by setting $\mathbf{P} = \mathbf{0}$ in the objective function of (3.10).

3.1.2.3   Quadratically Constrained Quadratic Programming

When both objective and constraint functions are quadratic, then the convex optimization problem is called QCQP. A QCQP can be expressed as follows [152]:

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^T\mathbf{P}_0\mathbf{x} + \mathbf{q}_0^T\mathbf{x} + r_0 \\
s.t. \quad & \mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i^T\mathbf{x} + r_i \leq 0, \quad i = 1, ..., m, \\
& \mathbf{A}\mathbf{x} = \mathbf{b},
\end{aligned}
\tag{3.11}
$$

where $\mathbf{P}_i \in \mathbb{S}_+^n, \ i = 1, ..., m$. In a QCQP, a quadratic convex function is minimized over a feasible region that is the intersection of ellipsoids. It is easily observed that a LP can be obtained through setting $\mathbf{P}_i = \mathbf{0}, \ i = 0, 1, ..., m$ in (3.11).

3.1.2.4   Second Order Cone Programming

A SOCP is defined in the following form [152]:

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{f}^T\mathbf{x} \\
s.t. \quad & ||\mathbf{A}_i\mathbf{x} + \mathbf{b}_i||_2 \leq \mathbf{c}_i^T\mathbf{x} + d_i, \quad i = 1, ..., m, \\
& \mathbf{F}\mathbf{x} = \mathbf{g},
\end{aligned}
\tag{3.12}
$$

where $\mathbf{A}_i \in \mathbb{R}^{n_i \times n}$ and $\mathbf{F} \in \mathbb{R}^{p \times n}$. The constraint $||\mathbf{A}\mathbf{x} + \mathbf{b}||_2 \leq \mathbf{c}^T\mathbf{x} + d$, where $\mathbf{A} \in \mathbb{R}^{k \times n}$, is called a second order cone constraint. A QCQP can be obtained by setting $\mathbf{c}_i = \mathbf{0}, \ i = 0, 1, ..., m$ and squaring both sides of the constraints. Similarly, a LP can be obtained if $\mathbf{A}_i = \mathbf{0}, \ i = 1, ..., m$.

3.1.2.5   Semidefinite Programming

The standard form of SDP can be expressed as [152]

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x} \\
s.t. \quad & \mathbf{x}_1\mathbf{F}_1 + \mathbf{x}_2\mathbf{F}_2 + ... + \mathbf{x}_n\mathbf{F}_n + \mathbf{G} \preceq \mathbf{0}, \\
& \mathbf{A}\mathbf{x} = \mathbf{b},
\end{aligned}
\tag{3.13}
$$

where $\mathbf{G}, \mathbf{F}_1, ..., \mathbf{F}_n \in \mathbb{S}^{k \times k}$ are symmetric matrices and $\mathbf{A} \in \mathbb{R}^{p \times n}$. The first constraint in (3.13) is called LMI. A LP can be obtained if the matrices $\mathbf{G}, \mathbf{F}_1, ..., \mathbf{F}_n$ are all diagonal.

3.1.3 Duality and KKT Conditions

The fundamental of duality is to take the constraints in (3.6) into account by augmenting the objective function with a weighted sum of the constraint functions. The Lagrangian $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ for the original problem in (3.6) can be defined as [152]

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{v}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}), \tag{3.14}$$

where λ_i and v_i are the Lagrange multipliers associated with the i-th inequality $f_i(\mathbf{x}) \leq 0$ and the equality $h_i(\mathbf{x}) = 0$ constraints, respectively. The objective function in (3.6), i.e., $f_0(\mathbf{x})$ is called the primal objective and the optimization variable $\mathbf{x}$ is the primal variable, whereas Lagrange multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{v}$ associated with (3.14) are known as the dual variables. The Lagrange dual objective or the Lagrange dual function $g : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ is defined as the minimum value of the Lagrangian over $\mathbf{x}$: for $\boldsymbol{\lambda} \in \mathbb{R}^m, \boldsymbol{v} \in \mathbb{R}$, and it is given as

$$g(\boldsymbol{\lambda}, \boldsymbol{v}) = \inf_{\mathbf{x} \in D} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}) \right). \tag{3.15}$$

The dual function is always concave regardless of the convexity of the original problem, which is due to the fact that the dual function is the pointwise infimum of a family of affine functions of $(\boldsymbol{\lambda}, \boldsymbol{v})$ [152]. The dual function $g(\boldsymbol{\lambda}, \boldsymbol{v})$ yields a lower bound on the optimal value of $f_0(\mathbf{x}^*)$ of the problem (3.6) [152], which can be equivalently written as

$$g(\boldsymbol{\lambda}, \boldsymbol{v}) \leq f_0(\mathbf{x}^*), \tag{3.16}$$

which indicates that for any feasible set $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{v})$, the following inequality holds:

$$\begin{aligned} f_0(\mathbf{x}) &\geq f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}) \\ &\geq \inf_{\mathbf{x} \in D} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}) \right) \\ &= g(\boldsymbol{\lambda}, \boldsymbol{v}). \end{aligned} \tag{3.17}$$

Duality gap is defined as the the difference between the primal objective $f_0(\mathbf{x})$ and the dual objective $g(\boldsymbol{\lambda}, \boldsymbol{v})$. A weak duality holds if $g(\boldsymbol{\lambda}, \boldsymbol{v}) < f_0(\mathbf{x}^*)$, while a strong duality holds if $g(\boldsymbol{\lambda}, \boldsymbol{v}) = f_0(\mathbf{x}^*)$. To obtain the best lower bound of the original problem, the dual problem defined as follows should be solved:

$$\underset{\boldsymbol{\lambda}, \boldsymbol{v}}{\text{maximize}} \quad g(\boldsymbol{\lambda}, \boldsymbol{v})$$
$$s.t. \quad \boldsymbol{\lambda} \succeq \mathbf{0}. \tag{3.18}$$

The Lagrange dual problem is always a convex problem, since the objective function in (3.18) is always a concave function and the constraints are convex. The KKT conditions listed below, which confirm the optimality of the solutions, are written as follows [152]:

1. Primal constraints: $f_i(\mathbf{x}) \leq 0, \ i = 1, ..., m, \ h_i(\mathbf{x}) = 0, \ i = 1, ..., p$.

2. Dual constraints: $\boldsymbol{\lambda} \succeq \mathbf{0}$.

3. Complementary slackness: $\lambda_i f_i(\mathbf{x}) = 0, \quad i = 1, ..., m$.

4. Gradient of Lagrangian w.r.t. $\mathbf{x}$ vanishes:

$$\boldsymbol{\nabla} f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \boldsymbol{\nabla} f_i(\mathbf{x}) + \sum_{i=1}^{p} v_i \boldsymbol{\nabla} h_i(\mathbf{x}) = 0. \tag{3.19}$$

For general optimization, these KKT conditions are necessary but not sufficient conditions for optimality. The KKT conditions hold if strong duality holds for any optimization problems. However, for convex optimization problems, if the KKT conditions hold, then the strong duality holds between the primal and the dual problems in which both primal and dual variables will be optimal [152].

3.2 Machine Learning

Machine learning is the scientific study of algorithms and statistical models, which are used by computers to learn and improve automatically from the environment without being explicitly programmed [38, 35, 37]. Recently, with the massive availability of training data and the recent development of advanced high speed processors, the applications of machine learning have grown dramatically across different research domains [35, 157, 158]. Machine

learning has been applied in numerous fields to solve different challenging problems in day to day lives of people, including finance, healthcare, traffic and information technologies [157]. One of the attractive features of machine learning techniques in wireless communications is that they have the potential capabilities to meet the requirements of ultra low latency and high reliability of 5G and future wireless networks, which address the issues of the high computational complexity and computation time of conventional convex optimization techniques. Recently, machine learning based optimization approaches have been applied to solve different complicated problems in wireless communication [138, 133, 61, 137]. The following subsections introduce different machine learning techniques which have been widely employed in different domains of wireless communications.

3.2.1   Linear Regression

Linear regression is an approach that constructs an appropriate linear function which can capture the relationship between inputs and the corresponding outputs [38, 158]. In this subsection, the underlying fundamental concepts of this linear regression technique are described. For a given set of N input variables, the linear regression can be modelled as a following single function that combines the input variables linearly [38]:

$$h(x) = \theta_0 + \theta_1 x_1 + ... + \theta_N x_N, \tag{3.20}$$

where $\theta_0, ..., \theta_N$ are the weights that need to be determined to approximate the relationship between the inputs and the outputs. By introducing a new parameter $x_0 = 1$, the equation in (3.20) can be written in a compact form as follows:

$$h(\mathbf{x}) = \sum_{i=0}^{N} \theta_i x_i = \boldsymbol{\theta}^T \mathbf{x}, \tag{3.21}$$

where $\boldsymbol{\theta} = [\theta_0, ..., \theta_N]$ and $\mathbf{x} = [x_0, ..., x_N]$. Assume that there are M training examples available, the dataset $\{(\mathbf{x}^{(m)}, y^{(m)}); m = 1, ...M\}$ is called a training set, where $\mathbf{x}^{(m)}$ is the m-th input vector and $y^{(m)}$ is the m-th target value that this approach need to approximate. It is assumed that the relationship between the target parameter $y^{(m)}$ and the inputs is defined through a deterministic function $h(\mathbf{x})$ as [159]

$$y^{(m)} = h(\mathbf{x}^{(m)}) + \varepsilon^{(m)} = \boldsymbol{\theta}^T \mathbf{x}^{(m)} + \varepsilon^{(m)}, \tag{3.22}$$

where $\varepsilon^{(m)}$ is an error term that captures either un-modelled effects or random noise. Furthermore, $\varepsilon^{(m)}$ is independently and identically distributed. Based on the Gaussian distribution with zero mean and variance σ^2, the probability density function of $\varepsilon^{(m)}$ can be expressed as [159]

$$p(\varepsilon^{(m)}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(\varepsilon^{(m)})^2}{2\sigma^2} \right). \tag{3.23}$$

Hence, the likelihood function can be written as follows [159]:

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{m=1}^{M} p(y^{(m)}|x^{(m)}; \boldsymbol{\theta}) \\ &= \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)})^2}{2\sigma^2} \right). \end{aligned} \tag{3.24}$$

The basic concept of maximum likelihood is to choose the parameters $\boldsymbol{\theta}$ to maximize $L(\boldsymbol{\theta})$. Based on the strictly increasing characteristic of logarithmic functions, the logarithm of the likelihood function $L(\boldsymbol{\theta})$ can be considered as

$$\begin{aligned} l(\boldsymbol{\theta}) &= \log L(\boldsymbol{\theta}) \\ &= \log \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)})^2}{2\sigma^2} \right), \\ &= \sum_{m=1}^{M} \log \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)})^2}{2\sigma^2} \right), \\ &= M \log \left(\frac{1}{\sqrt{2\pi}\sigma} \right) - \frac{1}{2\sigma^2} \sum_{m=1}^{M} (y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)})^2. \end{aligned} \tag{3.25}$$

Since $M \log(\frac{1}{\sqrt{2\pi}\sigma})$ and $\frac{1}{\sigma^2}$ are constants, maximizing $l(\boldsymbol{\theta})$ is equivalent to minimizing the following part of the log-likelihood function in (3.25):

$$\frac{1}{2} \sum_{m=1}^{M} (y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)}))^2. \tag{3.26}$$

As a result, the cost function (also known as loss function in the literature) can be defined as [159]

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{m=1}^{M} (y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)})^2. \tag{3.27}$$

In order to minimize the cost function, a well-known gradient descent iterative algorithm is considered. In particular, this iterative algorithm begins with a set of initial values for the weights $\boldsymbol{\theta}$, and iteratively updates $\boldsymbol{\theta}$ as

$$\boldsymbol{\theta} = \boldsymbol{\theta} - \alpha \frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \tag{3.28}$$

where α is the learning rate. The partial derivative $\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ can be expressed as

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\theta}} J(\boldsymbol{\theta}) &= \frac{\partial}{\partial \boldsymbol{\theta}} \frac{1}{2} \sum_{m=1}^{M} (y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)})^2 \\
&= \sum_{m=1}^{M} (y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)}) \cdot \frac{\partial}{\partial \boldsymbol{\theta}} (y^{(m)} - \boldsymbol{\theta}^T \mathbf{x}^{(m)}) \\
&= \sum_{m=1}^{M} (\boldsymbol{\theta}^T \mathbf{x}^{(m)} - y^{(m)}) \mathbf{x}^{(m)},
\end{aligned} \tag{3.29}$$

which gives the definition of the updating steps of the weights as follows:

$$\boldsymbol{\theta} - \alpha \sum_{m=1}^{M} (\boldsymbol{\theta}^T \mathbf{x}^{(m)} - y^{(m)}) \mathbf{x}^{(m)}. \tag{3.30}$$

3.2.2   Logistic Regression

Logistic regression is another type of machine learning technique which is used to determine the parameters of a logistic model. This approach has been widely used in different classification problems [38, 158]. In this subsection, a binary logistic model is considered, in which the output y can take only two possible values, 0 and 1. The hypotheses $h(\mathbf{x})$ in logistic regression can be defined as [158]

$$h(\mathbf{x}) = g(\boldsymbol{\theta}^T \mathbf{x}), \tag{3.31}$$

where $g(z) = \frac{1}{1+e^{-z}}$, is defined as the sigmoid function. Note that $g(z)$ tends to approach 1 as $z \to +\infty$, while $g(z)$ tends to approach to 0 as $z \to -\infty$. Hence, the hypotheses $h(x)$ is always bounded between 0 and 1. Considering M training examples and

$$P(y = 1|\mathbf{x}; \boldsymbol{\theta}) = h(\mathbf{x}), \tag{3.32}$$

$$P(y = 0|\mathbf{x}; \boldsymbol{\theta}) = 1 - h(\mathbf{x}), \tag{3.33}$$

the likelihood function can be expressed

$$
\begin{aligned}
L(\boldsymbol{\theta}) &= \prod_{m=1}^{M} p(y^{(m)}|\mathbf{x}^{(m)}\boldsymbol{\theta}) \\
&= \prod_{m=1}^{M} (h(\mathbf{x}^{(m)}))^{y^{(m)}} (1 - h(\mathbf{x}^{(m)}))^{1-y^{(m)}}.
\end{aligned}
\tag{3.34}
$$

The log likelihood can be written as

$$
\begin{aligned}
l(\boldsymbol{\theta}) &= \log L(\boldsymbol{\theta}) \\
&= \sum_{m=1}^{M} y^{(m)} \log h(\mathbf{x}^{(m)}) + (1 - y^{(m)}) \log(1 - h(\mathbf{x}^{(m)})).
\end{aligned}
\tag{3.35}
$$

Different from the derivation provided for linear regression, a gradient ascent method can be exploited in logistic regression to determine the parameters $\boldsymbol{\theta}$ to maximize the likelihood function. First, the partial derivatives of the log likelihood function are given as follows [158]:

$$
\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta}) &= \sum_{m=1}^{M} \left(y^{(m)} \frac{1}{g(\boldsymbol{\theta}^T \mathbf{x}^{(m)})} - (1 - y^{(m)}) \frac{1}{1 - g(\boldsymbol{\theta}^T \mathbf{x}^{(m)})} \right) \frac{\partial}{\partial \boldsymbol{\theta}} g(\boldsymbol{\theta}^T \mathbf{x}^{(m)}) \\
&= \sum_{m=1}^{M} (y^{(m)} - h(\mathbf{x}^{(m)})) \mathbf{x}^{(m)}.
\end{aligned}
\tag{3.36}
$$

Based on these derivatives, the stochastic gradient ascent rule can be defined as follows [158]:

$$\boldsymbol{\theta} = \boldsymbol{\theta} + \alpha \sum_{m=1}^{M} (y^{(m)} - h(\mathbf{x}^{(m)})) \mathbf{x}^{(m)}, \tag{3.37}$$

where α is the learning rate.

3.2.3   Neural Network
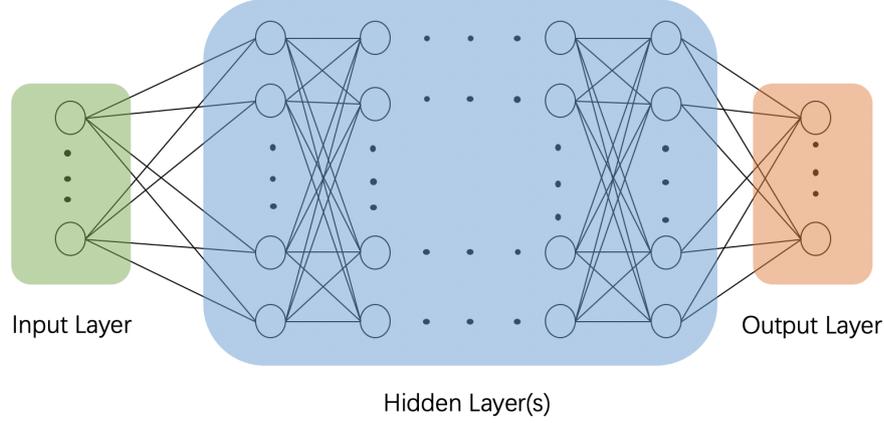
Fig. 3.1 A NN framework.

The fundamental idea behind the development of a NN is to mimic the information processing in biological systems [39], which has been used broadly to deal with a wide range of different realistic problems [38]. A typical NN consists of three layers: input layer, hidden layer(s) and output layer. Fig. 3.1 shows a feed-forward NN, in which the data are passed through the NN starting from the input layer, then the information moves in only one direction through the hidden layers, and finally to the output layer. This process can be expressed mathematically as [35]

$$\mathbf{z}^{(l+1)} = \mathbf{W}^{(l)}\mathbf{a}^{(l)} + \mathbf{b}^{(l)}, \tag{3.38}$$

$$\mathbf{a}^{(l+1)} = g(\mathbf{z}^{(l+1)}), \tag{3.39}$$

where $\mathbf{z}^{(l+1)}$ is the linear transformation of given inputs at the $l+1$-th layer, whereas $\mathbf{a}^{(l+1)}$ is output activation value of the $l+1$-th layer. Note that $g(\mathbf{z})$ denotes any activation function. Different activation functions can be chosen based on the types of problems. The activation functions include [39]:

$$g(\mathbf{z}) = \frac{1}{1 + e^{-\mathbf{z}}} \quad \text{(sigmoid)}, \tag{3.40}$$

$$g(\mathbf{z}) = \max(\mathbf{0}, \mathbf{z}) \quad \text{(Rectified Linear Unit (ReLU))}, \tag{3.41}$$

$$g(\mathbf{z}) = \frac{e^{\mathbf{z}} - e^{-\mathbf{z}}}{e^{\mathbf{z}} + e^{-\mathbf{z}}} \quad \text{(tanh)}. \tag{3.42}$$

The algorithmic procedure to determine the weights and the bias is called back-propagation. It is assumed that $J(\mathbf{W}, \mathbf{b})$ is the loss function of the NN, note that the loss function can be

different based on different activation functions. The following hold based on the chain rule:

$$\frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{W}^{(l)}} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}^{(l+1)}} \frac{\partial \mathbf{z}^{(l+1)}}{\partial \mathbf{W}^{(l)}}$$
$$= \boldsymbol{\delta}^{(l+1)}(\mathbf{a}^{(l)})^T, \tag{3.43}$$

$$\frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{b}^{(l)}} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}^{(l+1)}} \frac{\partial \mathbf{z}^{(l+1)}}{\partial \mathbf{b}^{(l)}}$$
$$= \boldsymbol{\delta}^{(l+1)}. \tag{3.44}$$

Since $\mathbf{a}^{(l)}$ can be calculated from feed-forward process, then, $\boldsymbol{\delta}^{(l)}$ can be derived based on the chain rule as follows:

$$\boldsymbol{\delta}^{(l)} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}^{(l)}} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}^{(l+1)}} \frac{\partial \mathbf{z}^{(l+1)}}{\partial \mathbf{a}^{(l)}} \frac{\partial \mathbf{a}^{(l)}}{\partial \mathbf{z}^{(l)}}$$
$$= [(\mathbf{W}^{(l)})^T \boldsymbol{\delta}^{(l+1)}] \cdot g'(\mathbf{z}^{(l)}). \tag{3.45}$$

Starting from the output layer, the $\boldsymbol{\delta}^{(l)}$ can be calculated backward layer by layer until the input layer. Finally, by assuming α as the learning rate, the weights matrix $\mathbf{W}^{(l)}$ and the bias vector $\mathbf{b}^{(l)}$ for the l-th layer can be updated respectively as follows [35]:

$$\mathbf{W}^{(l)} = \mathbf{W}^{(l)} - \alpha[\boldsymbol{\delta}^{(l+1)}(\mathbf{a}^{(l)})^T], \tag{3.46}$$

$$\mathbf{b}^{(l)} = \mathbf{b}^{(l)} - \alpha\boldsymbol{\delta}^{(l+1)}, \tag{3.47}$$

3.2.4   Reinforcement Learning

Reinforcement learning is another area of machine learning technique, in which machine performs a set of suitable actions to maximize some notion of cumulative reward in a particular situation. The underlying mathematical principle of reinforcement learning is Markov decision processes. In particular, a Markov decision process is a tuple $(\mathbf{S}, \mathbf{A}, \{P_{sa}\}, \gamma, R)$, where $\mathbf{S}$ is a set of states, $\mathbf{A}$ is a set of actions and P_{sa} are the state transition probabilities. The parameter $\gamma \in [0, 1)$ denotes the discount factor, whereas $R : \mathbf{S} \times \mathbf{A} \to \mathbb{R}$ is the reward function. In some state s_0, action $a_0 \in \mathbf{A}$ is carried out. As a result of a_0, the state randomly transits to another state s_1, according to $s_1 \sim P_{s_0 a_0}$. Then, another action a_1 is carried out, as a result, the state randomly changes again, to $s_2 \sim P_{s_1 a_1}$ and then action a_2 takes place and so on. Hence, the total reward is defined as [160]

$$R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots, \tag{3.48}$$

In reinforcement learning, the aim is to choose actions to maximize the expected value of the total pay-off [160]:

$$E[R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots]. \tag{3.49}$$

A policy is any function mapping from the states to the actions, i.e., for a state s, the action $a = \pi(s)$. The value function for a policy π can be defined as

$$V^\pi(s) = E[R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots | s_0 = s, \pi]. \tag{3.50}$$

$V^\pi(s)$ denotes the expected sum value of the rewards starting from state s and taking actions according to policy π. With a fixed policy π, the value function $V^\pi(s)$ satisfies the following Bellman equations [160]:

$$V^\pi(s) = R(s) + \gamma \sum_{s' \in \mathbf{S}} P_{s\pi(s)}(s') V^\pi(s'). \tag{3.51}$$

Then, the optimal value of the function can be expressed as [160]

$$V^*(s) = R(s) + \max_{a \in \mathbf{A}} \gamma \sum_{s' \in \mathbf{S}} P_{s\pi(s)}(s') V^\pi(s'). \tag{3.52}$$

3.3   Summary

This chapter has presented the required mathematical background for this thesis. In the first section, basic elements of convex optimization techniques, as well as duality and KKT conditions have been discussed. In the second section, some fundamental machine learning approaches have been provided, including linear regression, logistic regression, NN and reinforcement learning. In this thesis, convex optimization techniques are utilized to justify the convexity of optimization problems and transform non-convex problems into convex forms. In particular, different optimization problems, including AN power maximization, SEE maximization and secrecy rate maximization problems, are solved through convex optimization techniques. Specifically, Lagrangian dual problems and KKT conditions are exploited to prove the optimality of relaxed problems. Furthermore, NN is employed to map the relationship between the input and output parameters for a secrecy transmission network, which has the capability to provide near optimal performance with reduced computational complexity.

Chapter 4

Robust Beamforming for AN Aided MISO SWIPT System with Unknown Eavesdroppers and Non-linear EH Model

4.1 Introduction

It is difficult for the transmitter to obtain perfect CSI in wireless networks due to the channel estimation and quantization errors. In beamforming designs for secure transmissions, the secrecy performance may significantly degrade if the channel uncertainties are not taken into account in the designs. Therefore, it is important to incorporate the channel uncertainties in the beamforming designs. Recently, robust beamforming designs have been considered to address this channel uncertainty issue, where two types of robust designs are studied, namely, the worst-case robust design [69, 87–89], and the outage probability based design [95–97]. In the worst-case robust design, the channel uncertainties are assumed to be norm-bounded, whereas in the outage probability design, the CSI errors are modelled as random variables with known statistical properties. Furthermore, in most existing works on secure transmission schemes, it is assumed that the transmitter has perfect or imperfect knowledge on the CSI of EVEs [113, 53, 55, 89]. However, it is not always possible to obtain the CSI of EVEs in practical scenarios, for example, the EVEs might be purely passive during the transmissions. Most recently, robust beamforming designs with unknown EVEs have been studied in [161, 162], where the power of AN or jamming signal is maximized to confuse the unknown EVEs as much as possible under the QoS requirement at the legitimate user.

In this chapter, a beamforming design is considered for downlink transmission of a multi-user MISO system where each legitimate user employs a PS SWIPT technique. The transmitter intends to send confidential information to its legitimate users in the presence of purely unknown EVEs. The transmitter is equipped with multiple antennas, while the legitimate users and the EVEs have a single antenna. Furthermore, it is assumed that all the legitimate users can simultaneously perform information decoding and EH. A non-linear EH model is adopted in the legitimate users to capture the characteristics of practical EH circuits. In addition, it is assumed that all the EVEs are purely passive, hence their CSI is not available at the transmitter, and AN is exploited to mask the signal intended for the legitimate users. In this work, the aim is to maximize the power of AN to confuse the unknown EVEs while satisfying QoS and EH requirements at the legitimate users. Two robust joint beamforming and PS designs with the AN approach are considered: (1) For the bounded channel uncertainties, the problem is transformed as a SDP through SDR and a linear matrix inequality representation; (2) For the statistical channel uncertainties, both SDR and Bernstein-type inequality are exploited to reformulate the original problem into a convex form.

4.2   System Model

In this work, the downlink transmission of an AN aided MISO SWIPT system with K legitimate users and J EVEs is considered, as shown in Fig. 4.1. It is assumed that the legitimate transmitter is equipped with N_T antennas whereas both legitimate users and EVEs are equipped with single antennas. The channels between the transmitter and the k-th legitimate user, and the j-th EVE are denoted as $\mathbf{h}_{s,k} \in \mathcal{C}^{N_T \times 1}$ and $\mathbf{h}_{e,j} \in \mathcal{C}^{N_T \times 1}$, respectively. The $\mathbf{q}_k \in \mathcal{C}^{N_T \times 1}$ and $\mathbf{v} \in \mathcal{C}^{N_T \times 1}$ represent the information signal vector and the AN vector. In addition, the elements in AN vector $\mathbf{v}$ follow a zero mean complex Gaussian distribution with a covariance matrix $\mathbf{V} \succeq 0$. The received signal at the k-th legitimate user and the j-th EVE can be expressed as

$$y_{s,k} = \sum_{l=1}^{K} \mathbf{h}_{s,k}^H \mathbf{q}_l x_k + \mathbf{h}_{s,k}^H \mathbf{v} + n_{s,k}, \qquad (4.1)$$

$$y_{e,j} = \sum_{l=1}^{K} \mathbf{h}_{e,j}^H \mathbf{q}_l x_k + \mathbf{h}_{e,j}^H \mathbf{v} + n_{e,j}, \qquad (4.2)$$

respectively, where $n_{s,k}$ and $n_{e,j}$ are the joint effects of thermal noise and signal processing noise, respectively, at the k-th legitimate receiver and the j-th EVE, which are modelled as
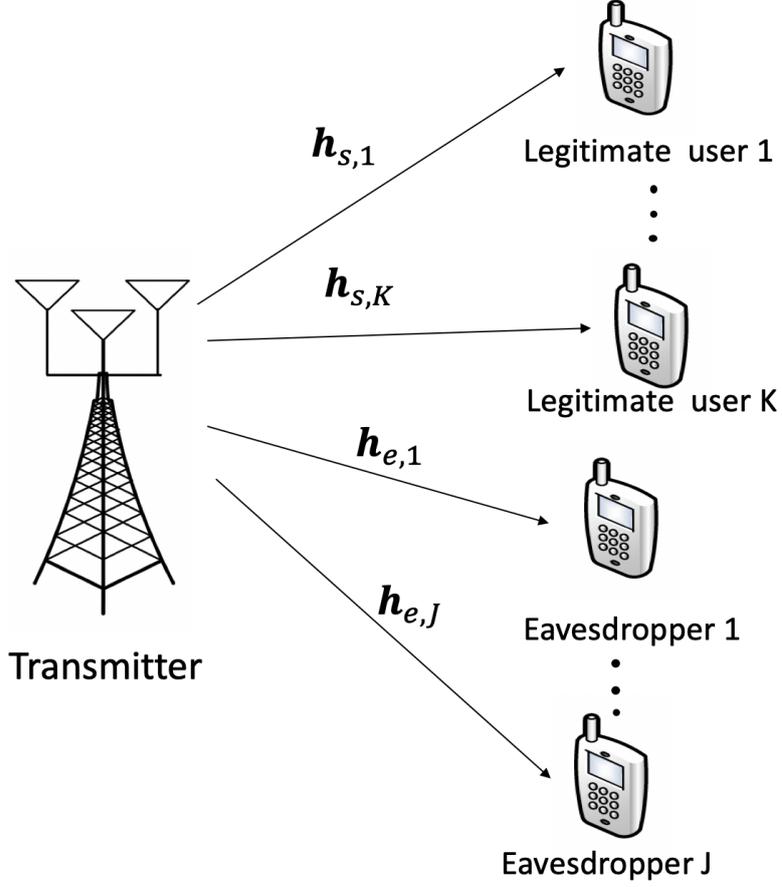
Fig. 4.1 A MISO SWIPT system with K single-antenna legitimate users in the presence of J single-antenna EVEs.

AWGN with zero mean and variance $\sigma_{s,k}^2$ and $\sigma_{e,j}^2$. x_k ($\mathbb{E}\{|x_k|^2\} = 1$) denotes signal from the transmitter intended to the k-th legitimate user. By employing PS approach, the received signal at the information decoding circuit of k-th legitimate user can be written as follows:

$$y_{s,k}^{ID} = \sqrt{\rho_{s,k}}(\sum_{l=1}^{K} \mathbf{h}_{s,k}^H \mathbf{q}_l x_l + \mathbf{h}_{s,k}^H \mathbf{v} + n_{s.k}) + n_{sp,k}, \tag{4.3}$$

where $\rho_{s,k} \in (0, 1]$ denotes the PS factor of the k-th legitimate user. Furthermore, $n_{sp,k} \sim \mathcal{CN}(0, \sigma_{sp,k}^2)$ is the noise introduced by the information decoding circuit of the k-th legitimate user. The received SINR at the k-th legitimate receiver and the j-th EVE can be written, respectively, as

$$\Gamma_{s,k} = \frac{\mathbf{h}_{s,k}^H \mathbf{q}_k \mathbf{q}_k^H \mathbf{h}_{s,k}}{\sum_{i \neq k} \mathbf{h}_{s,k}^H \mathbf{q}_i \mathbf{q}_i^H \mathbf{h}_{s,k} + \mathbf{h}_{s,k}^H \mathbf{V} \mathbf{h}_{s,k} + \sigma_{s,k}^2 + \frac{\sigma_{sp,k}^2}{\rho_{s,k}}}, \tag{4.4}$$

and

$$\Gamma_{e,j} = \frac{\mathbf{h}_{e,j}^H \mathbf{q}_k \mathbf{q}_k^H \mathbf{h}_{e,j}}{\sum_{i \neq k} \mathbf{h}_{e,j}^H \mathbf{q}_i \mathbf{q}_i^H \mathbf{h}_{e,j} + \mathbf{h}_{e,j}^H \mathbf{V} \mathbf{h}_{e,j} + \sigma_{e,j}^2}. \tag{4.5}$$

Furthermore, the received signal for the EH circuit of the k-th legitimate user is given by

$$y_{s,k}^{EH} = \sqrt{1 - \rho_{s,k}} \Big(\sum_{l=1}^{K} \mathbf{h}_{s,k}^H \mathbf{q}_l + \mathbf{h}_{s,k}^H \mathbf{v} + n_{s,k} \Big). \tag{4.6}$$

The received RF power at the k-th legitimate user can be expressed as

$$P_k = (1 - \rho_{s,k}) \Big(\sum_{l=1}^{K} \mathbf{h}_{s,k}^H \mathbf{q}_l \mathbf{q}_l^H \mathbf{h}_{s,k} + \mathbf{h}_{s,k}^H \mathbf{V} \mathbf{h}_{s,k} + \sigma_{s,k}^2 \Big). \tag{4.7}$$

The linear EH model is widely adopted in the literature [33, 32, 163, 48], where the output power of the k-th legitimate user can be given by the following linear model:

$$E_k^l = \zeta_k P_k, \tag{4.8}$$

where ζ_k represents the EH efficiency of the k-th legitimate user. In this work, a non-linear parametric EH model is adopted, which means the RF to direct current conversion efficiency depends on the input power level. Let the parameter a_k be the non-linear charging rate w.r.t. the input power and b_k be a minimum turn-on voltage of the EH circuit based parameter. Then, the output direct current power at the k-th legitimate user can be given by

$$E_k^{nl} = \frac{\Psi_k - M_k \Omega_k}{1 - \Omega_k}, \tag{4.9}$$

where $\Omega_k = \frac{1}{1 + \exp(a_k b_k)}$, $\Psi_k = \frac{M_k}{1 + \exp(-a_k(P_k - b_k))}$ and M_k is a constant denoting the maximum output direct current power at the k-th legitimate user.

4.3 Problem Formulation

As it is assumed that the CSI of EVEs is unavailable at the transmitter, the best approach is to exploit AN to degrade intercepting capability of EVEs. The aim is to maximize the AN power while ensuring the worst-case SINRs and energy harvesting requirements are satisfied. This design problem to jointly optimize beamforming vectors and AN covariance matrix can

be mathematically formulated as

$$\max_{\mathbf{V},(\mathbf{q}_k,\rho_{s,k},\forall k)} \text{tr}(\mathbf{V})$$

$$s.t.\ \Gamma_{s,k} \geq \gamma, \forall k,$$

$$E_k \geq \bar{E}_s, \forall k,$$

$$\sum_{l=1}^{K} \|\mathbf{q}_l\|_2^2 + \text{tr}(\mathbf{V}) \leq P_{\text{total}},$$

$$0 < \rho_{s,k} \leq 1, \mathbf{V} \succeq \mathbf{0}. \tag{4.10}$$

Firstly, the information beamforming vector $\mathbf{q}_k$ is tackled, by defining a new rank-one matrix $\mathbf{Q}_k = \mathbf{q}_k \mathbf{q}_k^H$. Then the original problem becomes

$$\max_{\mathbf{V},(\mathbf{Q}_k,\rho_{s,k},\forall k)} \text{tr}(\mathbf{V}) \tag{4.11a}$$

$$s.t.\ \frac{\mathbf{h}_{s,k}^H \mathbf{Q}_k \mathbf{h}_{s,k}}{\sum_{i \neq k} \mathbf{h}_{s,k}^H \mathbf{Q}_i \mathbf{h}_{s,k} + \mathbf{h}_{s,k}^H \mathbf{V} \mathbf{h}_{s,k} + \sigma_{s,k}^2 + \frac{\sigma_{sp,k}^2}{\rho_{s,k}}} \geq \gamma, \tag{4.11b}$$

$$(1 - \rho_{s,k})(\sum_{l=1}^{K} \mathbf{h}_{s,k}^H \mathbf{Q}_l \mathbf{h}_{s,k} + \mathbf{h}_k^H \mathbf{V} \mathbf{h}_k + \sigma_{s,k}^2) \geq \omega_k, \tag{4.11c}$$

$$\text{tr}(\sum_{l=1}^{K} \mathbf{Q}_l + \mathbf{V}) \leq P_{\text{total}}, \mathbf{V} \succeq \mathbf{0}, \mathbf{Q}_k \succeq \mathbf{0}, \forall k, \tag{4.11d}$$

$$0 < \rho_{s,k} \leq 1, \tag{4.11e}$$

$$\text{rank}(\mathbf{Q}_k) = 1. \tag{4.11f}$$

where ω_k represents the required received power under the non-linear EH model, which is given by

$$\omega_k = b_k - \frac{\ln(\frac{M_k}{\bar{E}_s + (M_k - \bar{E}_s)\Omega_k} - 1)}{a_k}. \tag{4.12}$$

By employing SDR, the problem defined in (4.11) is relaxed by dropping the non-convex rank constraint $\text{rank}(\mathbf{Q}_k) = 1$, the relaxed problem becomes:

$$\max_{\mathbf{V},(\mathbf{Q}_k,\rho_{s,k},\forall k)} \text{tr}(\mathbf{V})$$

$$s.t.\ (4.11b)-(4.11e). \tag{4.13}$$

Proposition 4.1: Provided that the problem (4.13) is feasible, the optimal solution will be always rank-one.

Proof: Please refer to Section 4.8.1. ∎

4.4 Robust Design with Ellipsoidal Channel Uncertainties

By taking into account the inevitable channel estimation and quantization errors, it is assumed that the transmitter has imperfect CSI of the legitimate users. The imperfect CSI are modelled based on the deterministic models [164, 162] and the actual channel of the k-th legitimate user can be written as follows:

$$\mathbf{h}_{s,k} = \hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k}, \tag{4.14}$$

where $\hat{\mathbf{e}}_{s,k}$ represents the channel uncertainties, and $\hat{\mathbf{h}}_{s,k}$ denotes the estimated channel vector. In this deterministic model, the Euclidean norms of channel errors are bounded by a set of thresholds as

$$||\hat{\mathbf{e}}_{s,k}||_2 = ||\mathbf{h}_{s,k} - \hat{\mathbf{h}}_{s,k}||_2 \leq \epsilon_{s,k}, \tag{4.15}$$

where $\epsilon_{s,k} \geq 0$ denotes the upper bound of channel uncertainties. By incorporating these uncertainties, the original problem can be formulated as the following robust optimization problem:

$$\max_{\mathbf{V},(\mathbf{Q}_k,\rho_{s,k},\forall k)} \text{tr}(\mathbf{V}) \tag{4.16a}$$

$$s.t. \ (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})^H \mathbf{Q}_k (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k}) \geq \gamma \Bigg[\sum_{i \neq k} (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})^H \mathbf{Q}_i (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})$$

$$+ (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})^H \mathbf{V}(\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k}) + \sigma_{s,k}^2 + \frac{\sigma_{sp,k}^2}{\rho_{s,k}} \Bigg], \tag{4.16b}$$

$$(1 - \rho_{s,k}) \Bigg[\sum_{l=1}^{K} (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})^H \mathbf{Q}_l (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k}) + (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})^H \mathbf{V}(\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})$$

$$+ \sigma_k^2 \Bigg] \geq \omega_k, \tag{4.16c}$$

$$\text{tr}(\sum_{l=1}^{K} \mathbf{Q}_l + \mathbf{V}) \leq P_{\text{total}}, \tag{4.16d}$$

$$0 < \rho_{s,k} \le 1, \mathbf{Q}_k \succeq \mathbf{0}, \forall k, \tag{4.16e}$$

$$\mathbf{V} \succeq \mathbf{0}, ||\hat{\mathbf{e}}_{s,k}||_2 \le \epsilon_{s,k}. \tag{4.16f}$$

Then, the channel errors are eliminated and the constraints in (4.16b) and (4.16c) are transformed into LMI forms. In order to convert (4.16b) into LMI form, a new variable $\mathbf{W}_k = \frac{1}{\gamma}\mathbf{Q}_k - \sum_{i \ne k} \mathbf{Q}_i - \mathbf{V}$ is first defined, then the following lemma is employed:

Lemma 4.1: (*Schur complement* [152]): Let $\mathbf{X}$ be a complex hermitian matrix,

$$\mathbf{X} = \mathbf{X}^H = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_2^H & \mathbf{B}_3 \end{bmatrix}. \tag{4.17}$$

Thus, $\mathbf{B}_4 = \mathbf{B}_3 - \mathbf{B}_2^H \mathbf{B}_1^{-1} \mathbf{B}_2$ is the Schur complement of $\mathbf{B}_1$ in $\mathbf{X}$ and the following statements hold: (1) $\mathbf{X} \succeq 0$, if and only if $\mathbf{B}_1 \succeq 0$ and $\mathbf{B}_4 \succeq 0$, (2) if $\mathbf{B}_1 \succeq 0$ then $\mathbf{X} \succeq 0$ if and only if $\mathbf{B}_4 \succeq 0$.

By utilizing Lemma 4.1, the constraint in (4.16b) can be rewritten as the following semidefinite constraint:

$$\begin{bmatrix} \rho_{s,k} & \sigma_{sp,k} \\ \sigma_{sp,k} & (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})^H \mathbf{W}_k (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k}) - \sigma_{s,k}^2 \end{bmatrix} \succeq \mathbf{0}. \tag{4.18}$$

However, this semidefinite constraint in (4.18) is not convex due to the channel uncertainties $\hat{e}_{s,k}$. In order to make this constraint tractable, the following lemma is utilized to convert it into LMIs:

Lemma 4.2: [165] If $\mathbf{D} \succeq 0$, then the following LMI

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 + \mathbf{A}_3 \mathbf{X} \\ (\mathbf{A}_2 + \mathbf{A}_3 \mathbf{X})^H & \mathbf{A}_4 + \mathbf{X}^H \mathbf{A}_5 + \mathbf{A}_5^H \mathbf{X} + \mathbf{X}^H \mathbf{A}_6 \mathbf{X} \end{bmatrix} \succeq \mathbf{0}, \tag{4.19}$$

$$\forall \mathbf{X} : \mathbf{I} - \mathbf{X}^H \mathbf{D} \mathbf{X} \succeq \mathbf{0}, \tag{4.20}$$

is equivalent to the following inequality. There exists $\lambda \ge 0$ such that

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{A}_3 \\ \mathbf{A}_2^H & \mathbf{A}_4 & \mathbf{A}_5^H \\ \mathbf{A}_3^H & \mathbf{A}_5 & \mathbf{A}_6 \end{bmatrix} - \lambda \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{D} \end{bmatrix} \succeq \mathbf{0}. \tag{4.21}$$

To proceed, it is set that $\mathbf{X} = \hat{\mathbf{e}}_{s,k}, \mathbf{D} = 1/\epsilon_{s,k}^2 \mathbf{I}, \mathbf{A}_1 = \rho_{s,k}, \mathbf{A}_2 = \sigma_{sp,k}, \mathbf{A}_3 = \mathbf{0}_{1 \times N_T}, \mathbf{A}_4 = \hat{\mathbf{h}}_{s,k}^H \mathbf{W}_k \hat{\mathbf{h}}_{s,k} - \sigma_{s,k}, \mathbf{A}_5 = \mathbf{W}_k \hat{\mathbf{h}}_{s,k}, \mathbf{A}_6 = \mathbf{W}_k$. Now the norm-bounded error vector $\hat{\mathbf{e}}_{s,k}$, i.e., $||\hat{\mathbf{e}}_{s,k}||_2 \leq \epsilon_{s,k}$ can be rewritten as $\mathbf{I} - \hat{\mathbf{e}}_{s,k}^H \frac{\mathbf{I}}{\epsilon_{s,k}^2} \hat{\mathbf{e}}_{s,k} \succeq \mathbf{0}$, then adopting Lemma 4.2, the constraint (4.18) can be equivalently expressed as

$$\begin{bmatrix} \rho_{s,k} & \sigma_{sp,k} & \mathbf{0}_{1 \times N_T} \\ \sigma_{sp,k} & \hat{\mathbf{h}}_{s,k}^H \mathbf{W}_k \hat{\mathbf{h}}_{s,k} - \sigma_{s,k}^2 - \lambda_k & \hat{\mathbf{h}}_{s,k}^H \mathbf{W}_k \\ \mathbf{0}_{N_T \times 1} & \mathbf{W}_k \hat{\mathbf{h}}_{s,k} & \mathbf{W}_k + \frac{\lambda_k}{\epsilon_{s,k}^2} \mathbf{I} \end{bmatrix} \succeq \mathbf{0}, \tag{4.22}$$

where $\lambda_k \geq 0$ is an auxiliary variable. Similarly, defining $\mathbf{M}_k = \sum_{l=1}^K \mathbf{Q}_l + \mathbf{V}$ and applying Lemma 4.1, the constraint in (4.16c) can be equivalently reformulated as

$$\begin{bmatrix} (1 - \rho_{s,k}) & \sqrt{\omega_k} \\ \sqrt{\omega_k} & (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k})^H \mathbf{W}_k (\hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k}) + \sigma_{s,k}^2 \end{bmatrix} \succeq \mathbf{0}. \tag{4.23}$$

By setting $\mathbf{X} = \hat{\mathbf{e}}_{s,k}, \mathbf{D} = 1/\epsilon_{s,k}^2 \mathbf{I}, \mathbf{A}_1 = (1 - \rho_{s,k}), \mathbf{A}_2 = \sqrt{\omega_k}, \mathbf{A}_3 = \mathbf{0}_{1 \times N_T}, \mathbf{A}_4 = \hat{\mathbf{h}}_{s,k}^H \mathbf{M}_k \hat{\mathbf{h}}_{s,k} + \sigma_{s,k}, \mathbf{A}_5 = \mathbf{M}_k \hat{\mathbf{h}}_{s,k}, \mathbf{A}_6 = \mathbf{M}_k$ and employing Lemma 4.2, the constraint (4.23) can be recast as

$$\begin{bmatrix} (1 - \rho_{s,k}) & \sqrt{\omega_k} & \mathbf{0}_{1 \times N_T} \\ \sqrt{\omega_k} & \hat{\mathbf{h}}_{s,k}^H \mathbf{M}_k \hat{\mathbf{h}}_{s,k} + \sigma_{s,k}^2 - t_k & \hat{\mathbf{h}}_{s,k}^H \mathbf{M}_k \\ \mathbf{0}_{N_T \times 1} & \mathbf{M}_k \hat{\mathbf{h}}_{s,k} & \mathbf{M}_k + \frac{t_{s,k}}{\epsilon_{s,k}^2} \mathbf{I} \end{bmatrix} \succeq \mathbf{0}, \tag{4.24}$$

where $t_k \geq 0$ is an auxiliary variable. Hence, the original problem (4.16) can be reformulated as

$$\max_{\mathbf{V}, (\mathbf{Q}_k, \rho_{s,k}, \lambda_k, t_k, \forall k)} \operatorname{tr}(\mathbf{V})$$

$$s.t. \ (4.11d), (4.11e), (4.22), (4.24),$$

$$\lambda_k \geq 0, t_k \geq 0. \tag{4.25}$$

The problem (4.25) is convex and can be solved efficiently via CVX [152].

4.5 Robust Design with Statistical Channel Uncertainties

In this subsection, another robust design is developed to handle statistical channel uncertainties which are modelled as Gaussian random variables with known statistical

distributions. In particular, it is assumed that the channel uncertainties have the following distribution:

$$\hat{\mathbf{e}}_{s,k} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Theta}_{s,k}), \tag{4.26}$$

where $\boldsymbol{\Theta}_{s,k} \succeq \mathbf{0}$ is the given covariance matrix for $\hat{\mathbf{e}}_{s,k}$. Then the robust problem can be written with statistical channel uncertainties as

$$\max_{\mathbf{V},(\mathbf{Q}_k, \rho_{s,k}, \forall k)} \mathrm{tr}(\mathbf{V}) \tag{4.27a}$$

$$s.t. \ \mathrm{Prob}(\Gamma_{s,k} \geq \gamma) \geq 1 - p_{s,k} \tag{4.27b}$$

$$\mathrm{Prob}(E_k \geq \bar{E}_s) \geq 1 - q_{s,k}, \tag{4.27c}$$

$$\mathbf{h}_{s,k} = \hat{\mathbf{h}}_{s,k} + \hat{\mathbf{e}}_{s,k}, \hat{\mathbf{e}}_{s,k} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Theta}_{s,k}), \tag{4.27d}$$

$$(4.11d), (4.11e), \tag{4.27e}$$

where $p_{s,k} \in (0,1]$ and $q_{s,k} \in (0,1]$ are the predefined outage probabilities for the SINR and EH requirements of the k-th legitimate user, respectively. These thresholds ensure the probabilities that the k-th legitimate user is served with a satisfiable SINR and EH no less than $(1 - p_{s,k}) \times 100\%$ and $(1 - q_{s,k}) \times 100\%$, respectively. The problem (4.27) is non-convex due to the outage probability constraints in (4.27b) and (4.27c) which do not have closed-form expressions. To circumvent this issue, the following lemma is exploited:

Lemma 4.3: (*Bernstein-type Inequality* [166]): Let $\mathbf{e} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$, $\mathbf{Q} \in \mathbb{H}^n$, $\mathbf{r} \in \mathbb{C}^n$ and $s \in \mathbb{R}$, for any $0 < p \leq 1$,

$$\mathrm{Prob}\{\mathbf{e}^H \mathbf{Q} \mathbf{e} + 2Re(\mathbf{e}^H \mathbf{r}) + s \geq 0\} \geq 1 - p, \tag{4.28}$$

is equivalent to the following set of convex constraints:

$$\mathrm{tr}(\mathbf{Q}) - \sqrt{-2\ln(p)}t_1 + \ln(p)t_2 + s \geq 0,$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{Q}) \\ \sqrt{2}\mathbf{r} \end{bmatrix} \right\|_2 \leq t_1,$$

$$t_2 \mathbf{I} + \mathbf{Q} \succeq \mathbf{0}, t_2 \geq 0, \tag{4.29}$$

where t_1 and t_2 are new set of slack variables. The following equations are first defined:

$$\mathbf{B}_k = \boldsymbol{\Theta}_{s,k}^{\frac{1}{2}} \mathbf{W}_k \boldsymbol{\Theta}_{s,k}^{\frac{1}{2}}, \mathbf{r}_k = \boldsymbol{\Theta}_{s,k}^{\frac{1}{2}} \mathbf{W}_k \hat{\mathbf{h}}_{s,k}, \ s_k = \hat{\mathbf{h}}_{s,k}^H \mathbf{W}_k \hat{\mathbf{h}}_{s,k} - \sigma_{s,k}^2 - \sigma_{sp,k}^2 / \rho, \tag{4.30}$$

$$\mathbf{E}_k = \boldsymbol{\Theta}_{s,k}^{\frac{1}{2}} \mathbf{M}_k \boldsymbol{\Theta}_{s,k}^{\frac{1}{2}}, \mathbf{g}_k = \boldsymbol{\Theta}_{s,k}^{\frac{1}{2}} \mathbf{M}_k \hat{\mathbf{h}}_{s,k}, \; l_k = \hat{\mathbf{h}}_{s,k}^H \mathbf{M}_k \hat{\mathbf{h}}_{s,k} + \sigma_{s,k}^2 - \frac{\omega_k}{1 - \rho_k}. \quad (4.31)$$

Then, by applying Lemma 4.1 and Lemma 4.3, the constraint in (4.27b) is equivalent to

$$\begin{bmatrix} f_{s,k} + \hat{\mathbf{h}}_{s,k}^H \mathbf{W}_k \hat{\mathbf{h}}_{s,k} - \sigma_{s,k}^2 & \sigma_{sp,k} \\ \sigma_{sp,k} & \rho_{s,k} \end{bmatrix} \succeq \mathbf{0},$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{B}_k) \\ \sqrt{2}\mathbf{r}_k \end{bmatrix} \right\|_2 \leq x_{s,k},$$

$$y_{s,k}\mathbf{I} + \mathbf{B}_k \succeq \mathbf{0}, \; y_{s,k} \geq 0, \quad (4.32)$$

where $f_{s,k} = \mathrm{tr}(\mathbf{B}_k) - \sqrt{-2\ln(p_{s,k})}x_{s,k} + \ln(p_{s,k})y_{s,k}$, and $x_{s,k}$ and $y_{s,k}$ are slack variables. Similarly, the constraint in (4.27c) can be equivalently written as

$$\begin{bmatrix} g_{s,k} + \hat{\mathbf{h}}_{s,k}^H \mathbf{M}_k \hat{\mathbf{h}}_{s,k} + \sigma_{s,k}^2 & \sqrt{\omega_k} \\ \sqrt{\omega_k} & 1 - \rho_{s,k} \end{bmatrix} \succeq \mathbf{0},$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{E}_k) \\ \sqrt{2}\mathbf{g}_k \end{bmatrix} \right\|_2 \leq m_{s,k},$$

$$n_{s,k}\mathbf{I} + \mathbf{E}_k \succeq \mathbf{0}, \; n_{s,k} \geq 0, \quad (4.33)$$

where $g_{s,k} = \mathrm{tr}(\mathbf{E}_k) - \sqrt{-2\ln(q_{s,k})}m_{s,k} + \ln(q_{s,k})n_{s,k}$, and $m_{s,k}$ and $n_{s,k}$ are slack variables. The statistical channel uncertainty based robust optimization problem in (4.27) is reformulated as

$$\max_{\mathbf{V},(\mathbf{Q}_k,\rho_{s,k},x_{s,k},y_{s,k},m_{s,k},n_{s,k},\forall k)} \mathrm{tr}(\mathbf{V})$$

$$s.t. \; (4.11\mathrm{d}), (4.11\mathrm{e}), (4.32), (4.33). \quad (4.34)$$

The formulation in (4.34) is a convex optimization problem and can be solved efficiently by available convex optimization software. [152].

4.6 Simulation Results

In this section, numerical results are presented to validate the performance of the proposed robust schemes. Here, a wireless communication system with one transmitter and three legitimate users is considered, where the transmitter is equipped with $N_T = 4$ antennas and all the legitimate users are equipped with single antenna. It is assumed that all the
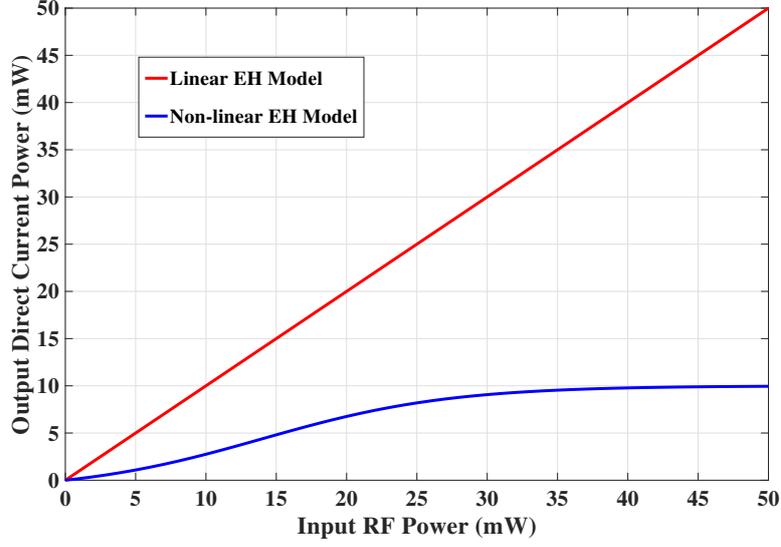
Fig. 4.2 Comparison of the output direct current power for the non-linear EH model and the linear EH model, where a similar performance comparison can be found in [112].

channel models include both large scale and small scale fading. The simplified large scale fading model is given by $D = (\frac{d}{d_0})^{-\Phi_1}$, where d is the distance between the transmitter and the receiver, $d_0 = 10$ m is assumed as the reference distance, and $\Phi_1 = 2$ as the path loss exponent. Furthermore, it is assumed that $d_1 = 10$ m, $d_2 = 14$ m and $d_3 = 18$ m as the distance between the transmitter and the 1st, 2nd and 3rd legitimate user, respectively. All the channel coefficients are modelled as Rician fading, which can be expressed as

$$\mathbf{h}_{s,k} = \sqrt{\frac{K_R}{1+K_R}}\mathbf{h}_{s,k}^{LOS} + \sqrt{\frac{1}{1+K_R}}\mathbf{h}_{s,k}^{NLOS}, \tag{4.35}$$

where $\mathbf{h}_{s,k}^{LOS}$ indicates the line of sight deterministic component with $||\mathbf{h}_{s,k}^{LOS}||_2^2 = D$, $\mathbf{h}_{s,k}^{NLOS} \sim \mathcal{CN}(\mathbf{0}, D\mathbf{I})$ represents the Rayleigh fading component. K_R is the Rician factor which is set to be 3. In addition, for the line of sight component, the far-field uniform linear antenna array is employed to model the channels [167]. The line of sight channel vectors can be modelled as [167]

$$\mathbf{h}_{s,k}^{LOS} = \sqrt{D/N_T}[1 \; e^{j\theta_k} \; ... \; e^{j(N_T-1)\theta_k}]^T$$

with $\theta_k = -\frac{w\pi\Psi\sin(\phi_k)}{wl}$, where wl is the carrier wavelength, $\Psi = wl/2$ is the spacing between successive antenna elements at the transmitter, and ϕ_k is the direction of the k-th legitimate user to the transmitter. It is assumed that $\phi_1 = -30°$, $\phi_2 = 0°$ and $\phi_3 = 30°$. For the
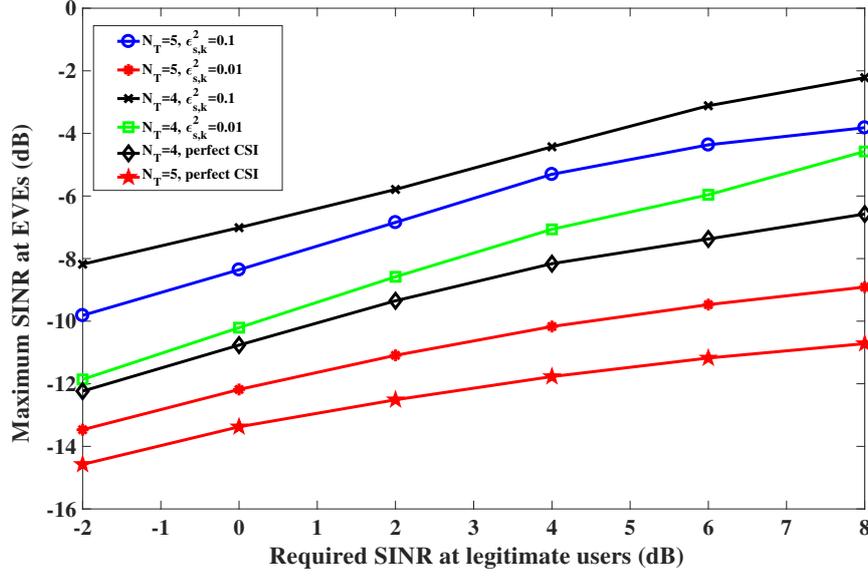
Fig. 4.3 Maximum SINR among EVEs versus the SINR requirement at legitimate users with different antennas and error bounds for the robust design with ellipsoidal channel uncertainties.

non-linear EH model, M is set to be 10 dBm which corresponds to the maximum output direct current power. Besides, $a_k = 150 \ \forall k$ and $b_k = 0.014 \ \forall k$ are adopted as provided in [112, 168]. The available total transmit power P_{total} is assumed to be 30 dBm. The EH efficiency ζ_k is set to be 1 for linear EH model. Note that the CSI of the EVEs is unavailable in both robust designs. To determine the worst-case SINR for EVEs, the channels of the EVEs are randomly generated in the same way as the channels of the legitimate users, by setting all the distance between the transmitter and EVEs to 8 m and the number of EVEs $J = 2$. For the statistical channel uncertainty based robust design, the covariance matrices are set to be $\boldsymbol{\Theta}_{s,k} = \epsilon_{s,k}^2 \mathbf{I}, \ \forall k$. In addition, the SINR and EH outage probability requirements in this design are set to be $p_{s,k} = q_{s,k} = 0.1$. All the numerical results are obtained by averaging over 1000 different random channel realizations. All the parameters described above are used in simulations unless or otherwise specified.

First, Fig. 4.2 provides the comparison of the output direct current power for the non-linear EH model and the linear EH model. As shown in Fig. 4.2, the output power of linear EH model increases as the input RF power increases. However, the output direct power of the non-linear EH model first increases and then becomes a constant. It is due to the fact that the output power of non-linear EH model cannot exceed the maximum output power threshold M, no matter how high the input RF power is.
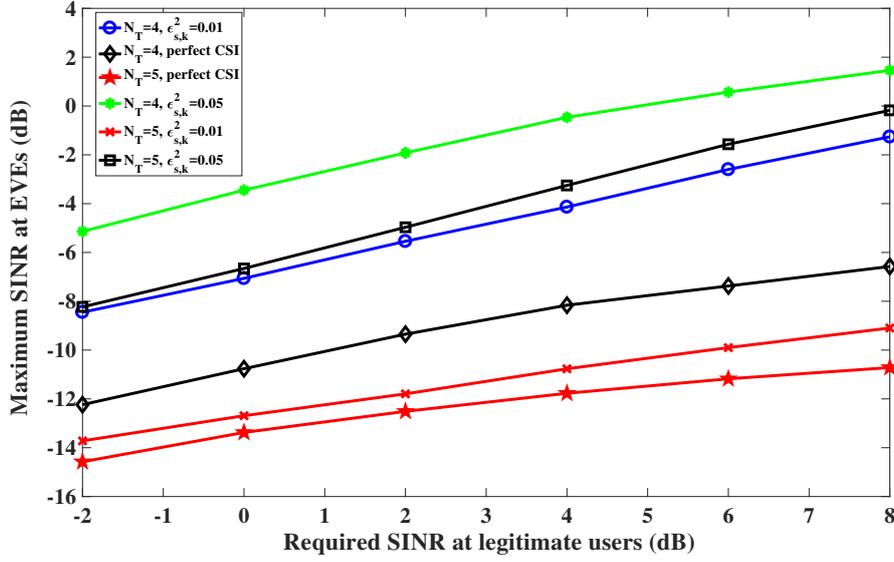
Fig. 4.4 Maximum SINR among EVEs versus the SINR requirement at legitimate users with different antennas and error bounds for the robust design with statistical channel uncertainties.

Fig. 4.3 compares the achieved maximum SINR at the EVEs between the perfect CSI design and robust design with the ellipsoidal channel uncertainties for different target SINRs and EH ($\bar{E}_s = 8$ dBm) requirements. As seen in Fig. 4.3, the largest SINR at the EVEs increases with the SINR targets and the error bounds. The design with perfect CSI achieves the least SINR at the EVEs with the same number of transmit antennas. Besides, it is obvious that the achievable SINR at the EVEs can be reduced with larger number of transmit antennas due to the additional degrees of freedom offered in the beamforming design.

Fig. 4.4 compares the achieved maximum SINR at the EVEs between the perfect CSI design and the robust design with statistical channel uncertainties for different target SINRs and EH ($\bar{E}_s = 8$ dBm) requirements. As seen in Fig. 4.4, the largest SINR at the EVEs increases with the SINR targets and the error bounds. The design with perfect CSI achieves the least SINR at the EVEs with the same number of transmit antennas. Similar to Fig. 4.3, it is obvious that the achievable SINR at the EVEs can be reduced by using a larger number of transmit antennas.

Fig. 4.5 compares non-linear EH model and linear EH model for robust design with ellipsoidal channel uncertainties. The required SINR at legitimate users is set to be 8 dB with the error-bound $\epsilon_{s,k}^2 = 0.01$. The largest SINR at the EVEs increases as EH requirement increases for both EH models. However, there exists a saturation point on the EH requirement $\bar{E}_s$ in the non-linear EH model, as seen in Fig. 4.5, due to the non-linear characteristics of practical RF-EH conversion circuit. Furthermore, for the non-linear EH model, the EH
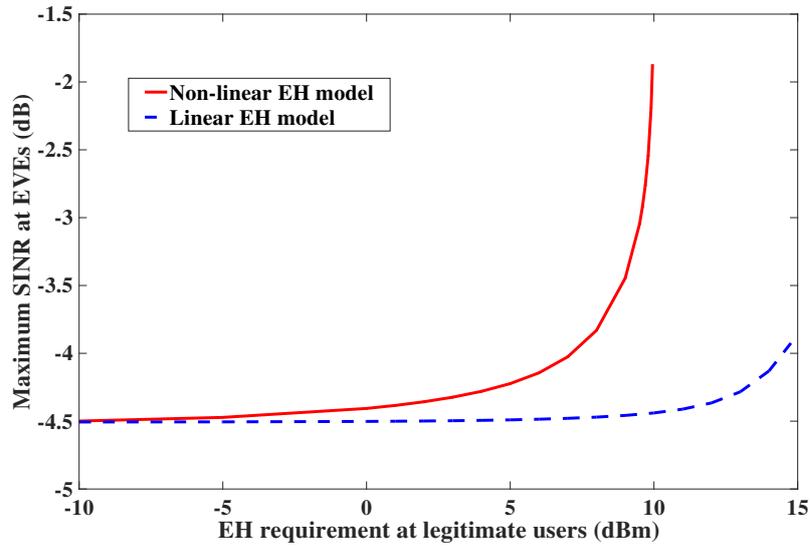
Fig. 4.5 Non-linear EH model vs linear EH model for robust design with ellipsoidal channel uncertainties.
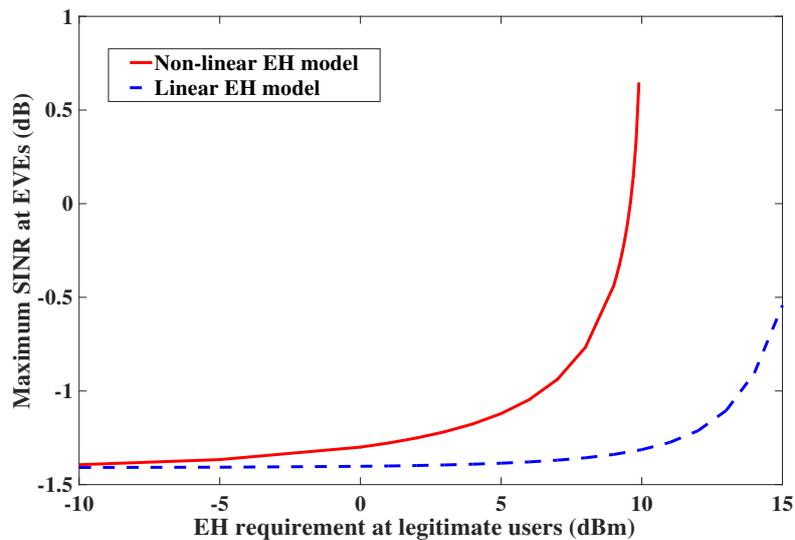
Fig. 4.6 Non-linear EH model vs linear EH model for robust design with statistical channel uncertainties.

requirements can be guaranteed only with $\bar{E}_s < M_k$ due to the maximum output direct current power limitation of the practical RF-EH conversion circuit. On the other hand, adopting the linear EH model may lead to false output direct current power. Hence, employing the non-linear EH model can reflect the characteristics of the practical RF-EH conversion circuit.
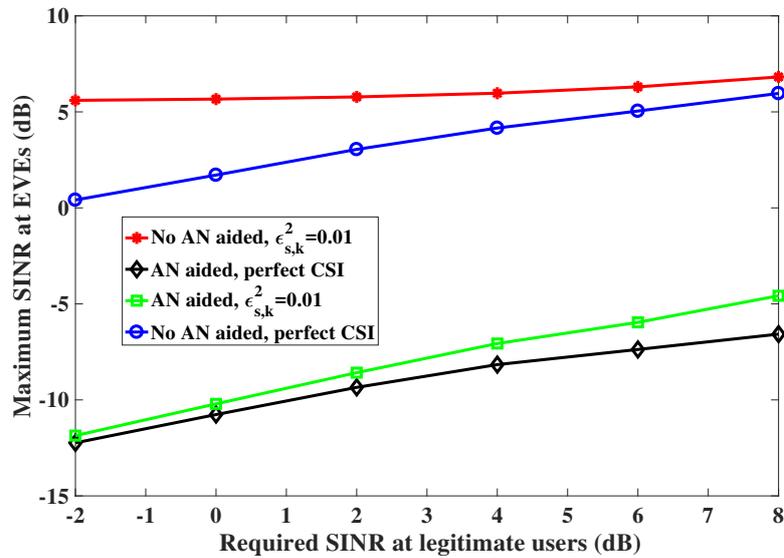
Fig. 4.7 Maximum SINR among EVEs versus the SINR requirement at legitimate users for different transmission schemes under the assumption of ellipsoidal channel uncertainties.
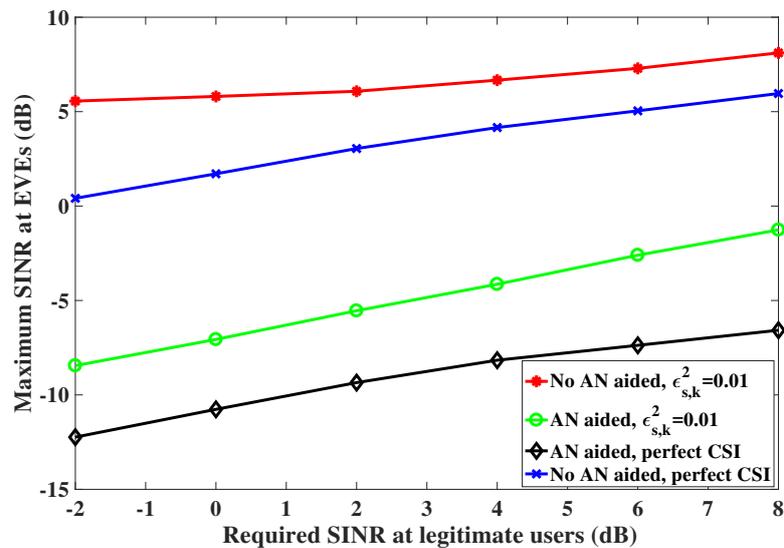
Fig. 4.8 Maximum SINR among EVEs versus the SINR requirement at legitimate users for different transmission schemes with under assumption of statistical channel uncertainties.

Similar performance is shown in Fig. 4.6 for the statistical channel uncertainty based robust design with $\gamma = 8$ dB, and $\epsilon_{s,k}^2 = 0.01$.

Next, Fig. 4.7 compares the achieved maximum SINR at the EVEs versus the SINR requirement at legitimate users for different transmission schemes with ellipsoidal channel uncertainties: AN aided and no AN aided schemes. The problem of no AN aided scheme is formulated as a power minimization problem as follows:

$$\min_{(\mathbf{Q}_k, \rho_{s,k}, \lambda_k, t_k, \forall k)} \text{tr}(\sum_{k=1}^{K} \mathbf{Q}_k)$$

$$s.t. \ (4.11\text{d}), (4.11\text{e}), (4.22), (4.24),$$

$$\lambda_k \geq 0, t_k \geq 0, \tag{4.36}$$

where the AN covariance matrix in the constraints of the above problem is set to $\mathbf{V} = \mathbf{0}$. As seen in Fig. 4.7, the largest SINR at the EVEs increases with the SINR targets. Furthermore, the proposed AN aided design outperforms the no AN scheme in all SINR regions for both perfect and imperfect CSI assumptions. This is due to the fact that: (1) The achieved SINR at EVEs is reduced by introducing AN; (2) The legitimate users have the capability to harvest energy from AN, and hence more power is split to the information decoding circuit. In other words, the same SINR at legitimate users can be achieved with less beamforming power in AN aided design than that of the no AN aided scheme.

Finally, Fig. 4.8 compares the achieved maximum SINR at the EVEs versus the SINR requirement at legitimate users for different transmission schemes with statistical channel uncertainties: AN aided and no AN aided schemes. The problem of no AN aided scheme is formulated as a power minimization problem as follows:

$$\min_{(\mathbf{Q}_k, \rho_{s,k}, x_{s,k}, y_{s,k}, m_{s,k}, n_{s,k}, \forall k)} \text{tr}(\sum_{k=1}^{K} \mathbf{Q}_k)$$

$$s.t. \ (4.11\text{d}), (4.11\text{e}), (4.32), (4.33). \tag{4.37}$$

where the AN covariance matrix in the constraints of the above problem is set to $\mathbf{V} = \mathbf{0}$. Similar to Fig. 4.7, the proposed AN aided design achieves much less SINR at the EVEs in comparison with the no AN aided scheme.

4.7 Summary

This chapter studied different robust beamforming designs in an AN-aided MISO SWIPT system in the presence of multiple purely passive EVEs. By considering practical

non-linear EH models, two robust designs have been developed based on different types of channel uncertainties: bounded and statistical based channel uncertainties. In the robust design with bounded uncertainties, this robust optimization problem was non-convex and was reformulated into a SDP through SDR, Schur complement and a LMI representation techniques. Then, by exploiting both SDR and Bernstein-type inequality approaches, the robust design with statistical channel uncertainties were recast as another convex optimization problem. Simulation results were provided to demonstrate that the proposed AN aided transmit schemes could significantly reduce the achieved SINR at EVEs while comparing with that of the no AN aided designs. The performance comparisons of both robust designs between different number of transmit antennas were also provided, which indicated that the achieved SINR at EVEs can be reduced by introducing more transmit antennas. Furthermore, the performance comparisons between the linear and non-linear EH model for both robust designs were provided, which showed that when the EH requirements are not constants, the resource allocation designs with non-linear EH model are more accurate than that with the linear EH model to reflect practical EH circuits.

4.8 Appendix

4.8.1 Proof of Proposition 4.1

First, the Lagrangian function of problem (4.13) can be derived as:

$$
\mathcal{L}(\mathbf{Q}_k, \mathbf{V}, \mathbf{Z}_k, \mathbf{Y}, \alpha, \beta, \gamma, \mu) = -\mathrm{tr}(\mathbf{V}) - \mathrm{tr}(\mathbf{Z}_k \mathbf{Q}_k) - \mathrm{tr}(\mathbf{Y}\mathbf{V}) - \lambda_k \{ \mathrm{tr}[\mathbf{h}_{s,k} \mathbf{h}_{s,k}^H (\frac{\mathbf{Q}_k}{\gamma} - \sum_{i \neq k} \mathbf{Q}_i
$$

$$
- \mathbf{V}) - \sigma_{s,k}^2 - \sigma_{sp,k}^2/\rho_{s,k}] \} - \mu_k \{ \mathrm{tr}[\mathbf{h}_{s,k} \mathbf{h}_{s,k}^H (\sum_{l=1}^{K} \mathbf{Q}_l + \mathbf{V}) + \sigma_{s,k}^2] - \omega_k/(1 - \rho_{s,k}) \}
$$

$$
+ \alpha[\mathrm{tr}(\sum_{l=1}^{K} \mathbf{Q}_l + \mathbf{V}) - P_{total}], \tag{4.38}
$$

where $\mathbf{Z}_k \in \mathbb{H}_+^{N_T}$, $\mathbf{Y} \in \mathbb{H}_+^{N_T}$, $\lambda_k \in \mathbb{R}_+$, $\mu_k \in \mathbb{R}_+$, $\alpha \in \mathbb{R}_+$ are the Lagrangian multipliers associated with problem (4.13). Then the corresponding KKT conditions [152] can be expressed as:

$$
\frac{\partial \mathcal{L}}{\partial \mathbf{Q}_k} = -\mathbf{Z}_k - (\lambda_k/\gamma + \mu_k)\mathbf{h}_{s,k} \mathbf{h}_{s,k}^H + \alpha \mathbf{I} = \mathbf{0}, \tag{4.39}
$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{V}} = -\mathbf{I} - \mathbf{Y} + (\lambda_k - \mu_k)\mathbf{h}_{s,k}\mathbf{h}_{s,k}^H + \alpha\mathbf{I} = \mathbf{0}, \tag{4.40}$$

$$\mathbf{Z}_k\mathbf{Q}_k = \mathbf{0}, \mathbf{Z}_K \succeq \mathbf{0}, \mathbf{YV} = \mathbf{0}, \mathbf{Y} \succeq \mathbf{0}. \tag{4.41}$$

The following equality holds:

$$(4.39)\text{-}(4.40) = -\mathbf{Z}_k + \mathbf{I} + \mathbf{Y} - \lambda_k(1 + \frac{1}{\gamma})\mathbf{h}_{s,k}\mathbf{h}_{s,k}^H = \mathbf{0}, \tag{4.42}$$

$$\Rightarrow \mathbf{Z}_k = \mathbf{I} + \mathbf{Y} - \lambda_k(1 + \frac{1}{\gamma})\mathbf{h}_{s,k}\mathbf{h}_{s,k}^H, \tag{4.43}$$

$$\Rightarrow \{\mathbf{I} + \mathbf{Y} - \lambda_k(1 + \frac{1}{\gamma})\mathbf{h}_{s,k}\mathbf{h}_{s,k}^H\}\mathbf{Q}_k = \mathbf{0}. \tag{4.44}$$

Hence, the following rank relation holds:

$$\text{rank}(\mathbf{Q}_k) = \text{rank}\left\{\frac{\lambda_k(1 + \frac{1}{\gamma})\mathbf{h}_{s,k}\mathbf{h}_{s,k}^H}{\mathbf{I} + \mathbf{Y}}\mathbf{Q}_k\right\} \le \text{rank}[\lambda_k(1 + \frac{1}{\gamma})\mathbf{h}_{s,k}\mathbf{h}_{s,k}^H] \le 1, \tag{4.45}$$

By eliminating the trivial solution $\mathbf{Q}_k = \mathbf{0}$, one can obtain $\text{rank}(\mathbf{Q}_k) = 1$, which completes the proof of proposition 4.1. ∎

Chapter 5

Energy Efficiency Optimization for Secure Transmission in MISO Cognitive Radio Network with Energy Harvesting

5.1 Introduction

Radio spectrum is a nature resource that enables wireless communications [169]. Conventionally, radio spectrum is allocated to users base on the static spectrum allocation policy [170–175], where every user need to obtain a license to access to a exclusive spectrum. This policy has been working well in the past decades, however, with the exponential growth of wireless devices, one critical issue has been exposed that the spectrum is not enough for allocating to every user [170]. Furthermore, most spectrum that allocated for licensed users are not fully utilized [171]. As a result, dynamic spectrum access has been proposed as a promising solution to the spectrum scarcity and inefficiency issues [176]. In dynamic spectrum access, the secondary users who do not have licenses to access the spectrum are allowed to either use the idle spectrum, or share the spectrum that belongs to primary users who are the license holders [176].

CR has been recognized as a key technique to implement dynamic spectrum access [177]. The concept of CR was firstly proposed in 1999 by Mitola [178]. CR networks have the capability to learn and exchange messages with the surrounding wireless networks to sense and exploit the idle spectrum source. Alternatively, in CR networks, the occurrence of conflicts can be effectively restricted or reduced through sensing the environment and dynamically and autonomously adapt its transmission strategies (i.e., transmit power and carrier frequency)

[169, 179–181, 23, 175, 182, 183]. The operation of CR can be classified into two categories, namely, opportunistic spectrum access and spectrum sharing. In opportunistic spectrum access model, the secondary users have to monitor the radio spectrum before transmission. When the spectrum is idle, the secondary users can access to the spectrum [184]. Different from opportunistic spectrum access, in spectrum sharing model, both primary and secondary users are allowed to transmit messages simultaneously [184]. However, to enabling spectrum sharing, the secondary transmitter has to obtain CSI and then adjust its transmit strategies to control the amount of interference power to the primary receivers [185]. The interference caused by the secondary transmission is acceptable for the primary users, and the maximum value of this interference power is named as the interference leakage tolerance [169].

Most research work on physical layer security transmission techniques consider either fully utilizing the available power to maximize the secrecy rate [11, 75, 83] or minimizing the required power to meet the target secrecy rate [22, 18]. However, these designs might not be able to achieve the maximum SEE due to their objectives which do not take into account the SEE performance metric. Therefore, from the perspective of energy efficient secure communication, the SEE has been considered as an appropriate performance metric to measure the efficient utilization of the limited power resource.

In this chapter, a number of SEE maximization problems for an underlay MISO CR network with EH requirements are solved. In particular, a multi-antenna SU-Tx simultaneously sends confidential information and energy to a SU-Rx and an ER, respectively. This SWIPT communication is established by sharing the spectrum that is allocated for communication for primary user terminals. In this CR network, transmit beamforming design is considered to maximize the achievable SEE under the constraints of secrecy rate on the SU-Rx, interference leakage on the PU-Rx and EH requirement on ER. Furthermore, the ER is considered to be a potential EVE due to the broadcast nature of wireless transmission.

5.2   System Model

A downlink transmission of MISO CR network is considered as shown in Fig. 5.1. This CR network consists of five terminals: an SU-Tx, an SU-Rx, a PU-Tx, a PU-Rx and an ER. The SU-Tx is equipped with N_T antennas, while the rest of the four terminals consist of a single antenna. Furthermore, it is assumed that the multi-antenna SU-Tx intends to send confidential information to the single antenna SU-Rx while satisfying the interference leakage threshold on the PU-Rx. The ER harvests energy from the SU-Tx transmission through WPT technique. However, the ER might as well be a potential EVE and may attempt to intercept
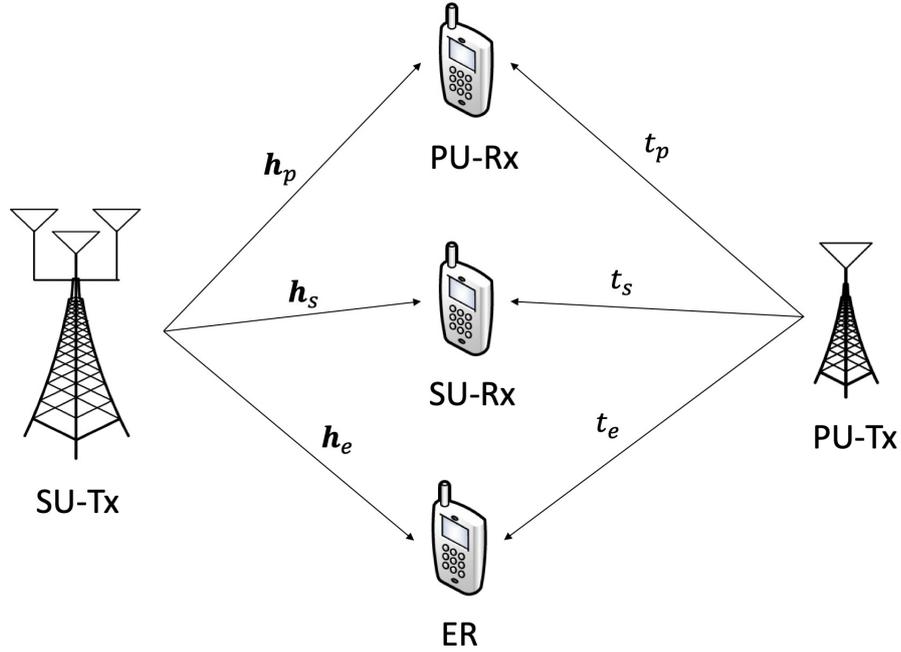
Fig. 5.1 An underlay CR network with a multi-antenna SU-Tx and single antenna PU-Tx, PU-Rx, SU-Rx and ER.

the message intended to the SU-Rx. Therefore, the ER is assumed to be a potential EVE in this CR network. The channel coefficients between the SU-Tx and PU-Rx, SU-Rx and ER are denoted by $\mathbf{h}_p \in \mathcal{C}^{N_T \times 1}$, $\mathbf{h}_s \in \mathcal{C}^{N_T \times 1}$ and $\mathbf{h}_e \in \mathcal{C}^{N_T \times 1}$, respectively. Thus, the received signal at SU-Rx and ER can be expressed as

$$y_s = \mathbf{h}_s^H \mathbf{q} x_s + t_s \sqrt{\Psi} x_p + n_s, \tag{5.1}$$

$$y_e = \mathbf{h}_e^H \mathbf{q} x_s + t_e \sqrt{\Psi} x_p + n_e, \tag{5.2}$$

where $\mathbf{q} \in \mathcal{C}^{N_T \times 1}$ and x_s $(\mathbb{E}\{|x_s|^2\} = 1)$ denote the beamforming vector and the signal from the SU-Tx, respectively. The notations $n_s \sim \mathcal{CN}(0, \sigma_s^2)$ and $n_e \sim \mathcal{CN}(0, \sigma_e^2)$ represent the joint effects of thermal noise and signal processing noise, at the SU-Rx and the ER, respectively. Moreover, it is assumed that Ψ and x_p $(\mathbb{E}\{|x_p|^2\} = 1)$ are the transmit power of PU-Tx and the information signal intended to the PU-Rx, respectively. The $t_s \in \mathcal{C}^{1 \times 1}$ and $t_e \in \mathcal{C}^{1 \times 1}$ denote the channel coefficients between the PU-Tx and the SU-Rx as well as the ER, respectively. Furthermore, the equivalent noise at the SU-Rx and ER, which captures the joint effect of the received interference from the PU-Tx, can be modelled as additive white Gaussian noise with zero mean and σ_{su}^2 and σ_{er}^2 variances, respectively [186, 169, 187]. The

achievable secrecy rate at the SU-Rx is defined as

$$R_s = \log_2\left(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{q}\mathbf{q}^H\mathbf{h}_s\right) - \log_2\left(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{q}\mathbf{q}^H\mathbf{h}_e\right). \tag{5.3}$$

The total transmit power consumption at the SU-Tx is

$$P_t = \frac{||\mathbf{q}||_2^2 + P_c}{\xi}, \tag{5.4}$$

where P_c is the circuit power consumption of the transmitter and the $\xi \in (0,1]$ is the power amplifier efficiency, which is assumed to be one ($\xi = 1$) throughout this chapter without loss of generality. The SEE, which is defined as the ratio between the achievable secrecy rate and the total transmit power consumption, can be expressed as

$$\begin{aligned} \eta &= \frac{R_s}{P_t} \\ &= \frac{\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{q}\mathbf{q}^H\mathbf{h}_s) - \log_2(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{q}\mathbf{q}^H\mathbf{h}_e)}{||\mathbf{q}||_2^2 + P_c}. \end{aligned} \tag{5.5}$$

The interference leakage to the PU-Rx is defined as

$$P_{il} = \mathbf{h}_p^H\mathbf{q}\mathbf{q}^H\mathbf{h}_p. \tag{5.6}$$

Recently, non-linear EH models have been investigated in the literature [64–66], which are more realistic and practical models to reflect the non-linear relationship between the RF and direct current conversion circuits. However, a generic non-linear EH model that could address all the issues in practical EH scenarios is still non-existent [188, 189]. Furthermore, by observing the measurement data presented in [190, 191], the output power of the linear EH model is accurate when compared against the non-linear model, even when considered across wider ranges of the input RF power. The linear EH model is also an effective model and has been widely adopted in the analysis and optimal design for EH systems [33, 32, 163, 48]. By assuming that the input RF power is well within the linear regime of the rectifier, a linear EH model is adopted. The harvested energy at the ER can be written as

$$P_r = \zeta_{eh}(\mathbf{h}_e^H\mathbf{q}\mathbf{q}^H\mathbf{h}_e + \sigma_{er}^2), \tag{5.7}$$

where $\zeta_{eh} \in (0,1]$ represents the EH efficiency of the ER.

5.3   SEE Maximization with Perfect CSI

In this section, an SEE maximization problem is solved with the minimum harvested energy and the maximum interference leakage constraints at the ER and PU-Rx, respectively. It is assumed that perfect CSI of all terminals is available at the SU-Tx. Note that for the passive ER, the perfect CSI can be estimated by the SU-Tx through local oscillator power leakage from the ER's RF front end, which is described in detail in the literature [192]. This SEE maximization problem is formulated as

$$\max_{\mathbf{q}} \frac{\log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{q}\mathbf{q}^H\mathbf{h}_s)-\log_2(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{q}\mathbf{q}^H\mathbf{h}_e)}{\|\mathbf{q}\|_2^2 + P_c} \tag{5.8a}$$

$$s.t. \log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{q}\mathbf{q}^H\mathbf{h}_s) - \log_2(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{q}\mathbf{q}^H\mathbf{h}_e) \geq R_d, \tag{5.8b}$$

$$\zeta_{eh}(\mathbf{h}_e^H\mathbf{q}\mathbf{q}^H\mathbf{h}_e + \sigma_{er}^2) \geq \omega_s, \tag{5.8c}$$

$$\mathbf{h}_p^H\mathbf{q}\mathbf{q}^H\mathbf{h}_p \leq P_f, \|\mathbf{q}\|_2^2 \leq P_{\text{tx}}, \tag{5.8d}$$

where ω_s indicates the minimum EH requirement at the ER. The P_f and P_{tx} are the predefined interference leakage tolerance at the PU-Rx and the maximum available transmit power at the SU-Tx, respectively. Let $\mathbf{Q}_s$ be a rank-one covariance matrix, based on the beamforming vector $\mathbf{q}$ such that $\mathbf{Q}_s = \mathbf{q}\mathbf{q}^H$. Then the original problem can be expressed as

$$\max_{\mathbf{Q}_s} \frac{R_s(\mathbf{Q}_s)}{P_t(\mathbf{Q}_s)} \tag{5.9a}$$

$$s.t. \log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)-\log_2(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e) \geq R_d, \tag{5.9b}$$

$$\zeta_{eh}(\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e + \sigma_{er}^2) \geq \omega_s, \tag{5.9c}$$

$$\mathbf{h}_p^H\mathbf{Q}_s\mathbf{h}_p \leq P_f, \tag{5.9d}$$

$$\text{tr}(\mathbf{Q}_s) \leq P_{\text{tx}}, \mathbf{Q}_s \succeq \mathbf{0}, \text{rank}(\mathbf{Q}_s) = 1, \tag{5.9e}$$

where $R_s(\mathbf{Q}_s) = \log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)-\log_2(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e)$ and $P_t(\mathbf{Q}_s) = \text{tr}(\mathbf{Q}_s) + P_c$.

This is a non-convex problem due to the fractional objective function. To recast this problem as a convex one, non-linear fractional and DC programming are employed in the following subsections.

5.3.1 Non-linear Fractional Programming

The objective function defined in (5.9a) is a fractional function with non-linear terms in the numerator and denominator. Furthermore, the overall problem presented in (5.9) is known as a non-linear fractional problem in the literature [45]. First, it is transformed into a parametric programming.

Theorem 5.1 (From [45]): Let $R_s(\mathbf{Q}_s^*) > 0$ and $P_t(\mathbf{Q}_s^*) > 0$ be the optimal achieved secrecy rate and optimal minimum total power consumption of the problem defined in (5.9), respectively. Further, let $\mathcal{F}$ be the feasible solution set of this problem. Moreover, the transmit covariance matrix $\mathbf{Q}_s^*$ provides the maximum SEE such that

$$\lambda^* = \frac{R_s(\mathbf{Q}_s^*)}{P_t(\mathbf{Q}_s^*)} = \max_{\mathbf{Q}_s \in \mathcal{F}} \left\{ \frac{R_s(\mathbf{Q}_s)}{P_t(\mathbf{Q}_s)} \right\}, \tag{5.10}$$

if and only if $R_s(\mathbf{Q}_s^*)$, $P_t(\mathbf{Q}_s^*)$ and λ^* satisfy the following condition:

$$\max_{\mathbf{Q}_s \in \mathcal{F}} \{R_s(\mathbf{Q}_s) - \lambda^* P_t(\mathbf{Q}_s)\} = R_s(\mathbf{Q}_s^*) - \lambda^* P_t(\mathbf{Q}_s^*) = 0. \tag{5.11}$$

Additionally,

$$\max_{\mathbf{Q}_s, \lambda} \{R_s(\mathbf{Q}_s) - \lambda P_t(\mathbf{Q}_s)\} \tag{5.12}$$

is defined as a parameteric programming with parameter λ [45].

Proof: Please refer to Section 5.8.1 ∎

Lemma 5.1: The objective function defined in (5.12) is a convex, strictly decreasing and continuous function w.r.t. λ.

Lemma 5.2: The equation $\max_{\mathbf{Q}_s, \lambda} \{R_s(\mathbf{Q}_s) - \lambda P_t(\mathbf{Q}_s)\} = 0$ has a unique solution, and let denote it by λ_0. Then, the problems defined in (5.12) and (5.9) have the same optimal solution with the optimal SEE of λ_0.

Proof: Please refer to [45]. ∎

By utilizing Theorem 5.1, the original problem defined in (5.9) is recast into the following parametric programming problem with parameter λ:

$$F(\lambda) = \max_{\mathbf{Q}_s} [R_s(\mathbf{Q}_s) - \lambda P_t(\mathbf{Q}_s)]$$

$$= \max_{\mathbf{Q}_s} \{\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - \log_2(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e) - \lambda[\text{tr}(\mathbf{Q}_s) + P_c]\}$$

$$s.t.\ (5.9\text{b})\text{-}(5.9\text{e}). \tag{5.13}$$

It can be seen that the original problem presented in (5.9) is transformed into a parameterized polynomial subtractive form. As a result, the original problem is reformulated to determine λ^* and $\mathbf{Q}_s^*$, which can satisfy the condition provided in (5.11). Furthermore, by utilizing Dinkelbach's algorithm [45] with an initial value λ_0 of λ, the optimal solutions of (5.9) can be obtained by iteratively solving the following optimization problem:

$$\max_{\mathbf{Q}_s} [R_s(\mathbf{Q}_s) - \lambda_i P_t(\mathbf{Q}_s)]$$

$$s.t.\ (5.9\text{b})\text{-}(5.9\text{e}), \tag{5.14}$$

for a given λ_i at the ith iteration. The value of λ_i can be considered as the SEE obtained at the previous iteration. At each iteration, λ_{i+1} should be updated such that

$$\lambda_{i+1} = \frac{R_s^i(\mathbf{Q}_s)}{P_t^i(\mathbf{Q}_s)}, \tag{5.15}$$

where R_s^i and P_t^i denote the achieved secrecy rate and total power consumption of (5.14) for a given λ_i from the previous iteration, respectively. This iterative process will be terminated when the condition in (5.11) is satisfied. However, in practice the iterative process will be carried out until the following inequality is satisfied:

$$\Delta F(\lambda) = |R_s^i(\mathbf{Q}_s) - \lambda_i P_t^i(\mathbf{Q}_s)| \leq \varepsilon, \tag{5.16}$$

where $\varepsilon > 0$ is the convergence threshold.

5.3.2 DC Programming

In this subsection, the required details on DC programming are provided to solve the SEE maximization problem. DC programming is a well-known optimization approach to solve non-convex problems. In particular, this technique can be exploited in an optimization problem with an objective function defined as a difference of two concave functions. Since the objective function in (5.14) falls under this category, DC programming is utilized to solve this problem.

The fundamental idea of DC programming is to locally linearize the non-concave functions at a feasible point $\mathbf{Q}_s^k$, and to iteratively update the approximation by solving the corresponding approximated problem [46]. The following function is defined to approximate the second term of the objective function in (5.14):

$$f(\mathbf{Q}_s, \mathbf{Q}_s^k) = \log_2(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e) + \frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H(\mathbf{Q}_s - \mathbf{Q}_s^k)\mathbf{h}_e}{(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e)\ln 2}. \tag{5.17}$$

Based on this approximation, the problem defined in(5.14) can be transformed into an equivalent problem as follows:

$$\max_{\mathbf{Q}_s}\ \{\log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)-f(\mathbf{Q}_s,\mathbf{Q}_s^k)-\lambda_i[\mathrm{tr}(\mathbf{Q}_s)+P_c]\}$$
$$s.t.\ (5.9b)\text{-}(5.9e). \tag{5.18}$$

This problem is still non-convex due to the non-convex rank-one constraint. By employing SDR, the problem in (5.9) is relaxed by dropping the rank constraint $\mathrm{rank}(\mathbf{Q}_s) = 1$, which can be defined as:

$$\max_{\mathbf{Q}_s}\ \{\log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)-f(\mathbf{Q}_s,\mathbf{Q}_s^k)-\lambda_i[\mathrm{tr}(\mathbf{Q}_s)+P_c]\}$$
$$s.t.\ (5.9b)\text{-}(5.9d),\ \mathbf{Q}_s \succeq \mathbf{0}, \mathrm{tr}(\mathbf{Q}_s) \leq P_{\mathrm{tx}}. \tag{5.19}$$

Proposition 5.1: Provided that the problem (5.19) is feasible, the optimal solution will be always rank-one.

Proof: Please refer to Section 5.8.2.   ∎

This approximated problem is convex in terms of $\mathbf{Q}_s$. Hence, the suboptimal solution $\mathbf{Q}_s^*$ to the original problem can be obtained by solving the problem in (5.19) and iteratively updating $\mathbf{Q}_s^k$ based on the solution obtained from the previous iteration. Note that DC programming is developed based on the first-order Taylor approximation, where the rest terms are ignored. Hence, the iterative algorithm of DC programming only guarantees to converge to a suboptimal [152]. The algorithm based on DC programming is summarized in Algorithm 1.

Algorithm 1 Iterative algorithm to solve (5.19)

1: Initialize $i = 0$ and choose an initial value λ_0;
2: **repeat**$\leftarrow$ Outer loop (Dinkelbach's algorithm)
3:      Initial $k = 0$, choose an initial value $\mathbf{Q}_s^k = 0$ and $F(\lambda)^{i,k} = 0$;
4:      **repeat**$\leftarrow$ Inner loop (DC programming)
5:          Solve the problem (5.19) with $\lambda = \lambda_i$ and obtain $\mathbf{Q}_s^{k+1}$;
6:          Compute $F(\lambda)^{i,k+1} = \log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s^{k+1}\mathbf{h}_s) - \log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_e^H\mathbf{Q}_s^{k+1}\mathbf{h}_e) -$
            $\lambda_i[\mathrm{tr}(\mathbf{Q}_s^{k+1}) + P_c]$;
7:          $\Delta\mu = F(\lambda)^{i,k+1} - F(\lambda)^{i,k}$;
8:          Update $k = k + 1$;
9:      **until** $|\Delta\mu| \leq \zeta$;
10:     Update λ_{i+1} through (5.15);
11:     $i = i + 1$;
12: **until** (5.16) satisfied;
13: Return $\lambda^* = \lambda_i$, $P_t^*(\mathbf{Q}_s) = P_t^{i-1}(\mathbf{Q}_s)$, $R_s^* = R_s^{i-1}$;
14: Output λ_i.

5.3.3 Computational Complexity

In this subsection, the computational complexity analysis of Algorithm 1 is provided. The problem defined in (5.19) can be solved by a standard interior point method [193]. To analyse the computational complexity of interior point methods, a general conic program is first presented as follows:

$$\min_{\mathbf{z}\in\mathbb{R}^n} \quad \mathbf{c}^T\mathbf{z} \tag{5.20a}$$

$$s.t. \quad \sum_{i=1}^n z_i\mathbf{A}_i^j - \mathbf{B}^j \in \mathbb{S}_+^{k_j} \quad \text{for } j = 1, ..., p, \tag{5.20b}$$

where $\mathbb{S}_+^{k_j}$ is the set of $k \times k$ positive semidefinite matrices. Note that linear constraints are equivalent to LMI constraints, i.e., $\mathbf{a}^T\mathbf{z} - b \geq 0$ is equivalent to $\mathbf{a}^T\mathbf{z} - b \in \mathbb{S}_+^1$. The computational complexity of a generic interior point method for solving (5.20) consists of the following parts [166, 193]:

1. Iteration complexity: The number of iterations required to reach an ϵ-optimal solution for a given $\epsilon > 0$ of (5.20) is on the order of $\sqrt{\Upsilon(\mathbb{k})}\ln(1/\epsilon)$, where $\sqrt{\Upsilon(\mathbb{k})} = \sum_{j=1}^p k_j$ is known as barrier parameter associated with the cone $\mathbb{k} = \prod_{j=1}^p \mathbb{S}_+^{k_j}$.

2. Per-iteration computational complexity: The search direction can be determined by solving n linear equations in n variables in each iteration. This computational

complexity is obtained by calculating the sum of formation and factorization of $n \times n$ coefficient matrix $\mathbf{H}$. The computational complexity of the formation is on the order of

$$C_{form} = n \sum_{j=1}^{p} k_j^3 + n^2 \sum_{j=1}^{p} k_j^2,$$

and that of factorizing $\mathbf{H}$ is on the order of $C_{fact} = n^3$. Therefore, the total computational complexity of this part is on the order of $\Lambda = C_{form} + C_{fact}$.

Combining these two parts, the computational complexity of a generic interior point method for solving (5.20) is on the order of $\sqrt{\Upsilon(\Bbbk)}\Lambda \ln(1/\epsilon)$.

The problem defined in (5.19) has four linear constraints, and one LMI constraint of size N_T. The number of variables n is in the order of N_T^2. To obtain an optimal solution to the problem defined in (5.19), the computational complexity is in the order of $\sqrt{\Upsilon(\Bbbk)}\Lambda \ln(1/\epsilon)$, where $\Upsilon(\Bbbk) = N_T + 4$, $\epsilon > 0$, and $\Lambda = n^3 + n(N_T^3 + 4) + n^2(N_T^2 + 4)$ [166, 193]. The total computational complexity of solving the problem defined in (5.9) can be calculated through multiplying the computational cost of solving (5.19) by the number of both inner and outer loop iterations. Therefore, the computational complexity of solving the problem defined in (5.9) should be in the order of $TK\sqrt{N_T + 4}[n^3 + n(N_T^3 + 4) + n^2(N_T^2 + 4)] \ln(1/e)$, where T and K are the numbers of iterations of the inner and the outer loop, respectively.

5.4   SEE Maximization with Statistical CSI

In this section, a robust design is considered with a realistic assumption that only the statistical CSI of ER is available at the SU-Tx. This scenario could arise when ER is part of neither the primary nor the secondary system. Hence, it is difficult to have ER's CSI at the SU-Tx based on handshaking signals. Furthermore, this assumption can be further supported by the fact that the ER might be silent or passive to hide its existence from the SU-Tx. Hence, it is assumed that only the statistical information on CSI of ER is available at the SU-Tx and the corresponding CSI of ER can be defined as

$$\mathbf{h}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{G}_e), \tag{5.21}$$

where $\mathbf{G}_e \succeq \mathbf{0}$ is the covariance matrix of $\mathbf{h}_e$. With this statistical CSI, the SEE maximization problem can be formulated by introducing a new slack variable δ as follows:

$$\max_{\mathbf{Q}_s, \delta} \delta \tag{5.22a}$$

$$s.t. \ \Pr\{f_s(\mathbf{Q}_s) \geq \delta\} \geq 1 - \alpha, \tag{5.22b}$$

$$\Pr\{g_s(\mathbf{Q}_s) \geq R_d\} \geq 1 - \beta, \tag{5.22c}$$

$$\Pr\{\zeta_{eh}(\mathbf{h}_e^H \mathbf{Q}_s \mathbf{h}_e + \sigma_{er}^2) \leq \omega_s\} \leq \gamma, \tag{5.22d}$$

$$\mathbf{h}_p^H \mathbf{Q}_s \mathbf{h}_p \leq P_f, \tag{5.22e}$$

$$\mathrm{tr}(\mathbf{Q}_s) \leq P_{\mathrm{tx}}, \mathbf{Q}_s \succeq 0, \mathrm{rank}(\mathbf{Q}_s) = 1, \tag{5.22f}$$

where $f_s(\mathbf{Q}_s)$ and $g_s(\mathbf{Q}_s)$ are defined as

$$f_s(\mathbf{Q}_s) = \frac{\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s) - \log_2(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H \mathbf{Q}_s \mathbf{h}_e)}{\mathrm{tr}(\mathbf{Q}_s) + P_c}, \tag{5.23}$$

$$g_s(\mathbf{Q}_s) = \log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s) - \log_2(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H \mathbf{Q}_s \mathbf{h}_e). \tag{5.24}$$

In (5.22b)-(5.22d), the parameters α, β and γ are chosen to define the outage probabilities for SEE, secrecy rate and EH at ER as $0 < \alpha < 0.5$, $0 < \beta < 0.5$ and $0 < \gamma < 0.5$, respectively. Furthermore, the left hand side of the constraint in (5.22b) can be formulated as

$$\Pr\left\{ \log \frac{1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s}{1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H \mathbf{Q}_s \mathbf{h}_e} \geq \delta(\mathrm{tr}(\mathbf{Q}_s) + P_c) \right\}$$

$$= \Pr\left\{ \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H \mathbf{Q}_s \mathbf{h}_e \leq (1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s)2^{-\delta(\mathrm{tr}(\mathbf{Q}_s) + P_c)} - 1 \right\}$$

$$\overset{(a)}{=} 1 - \exp\left(\frac{1 - (1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s)2^{-\delta(\mathrm{tr}(\mathbf{Q}_s) + P_c)}}{\frac{1}{\sigma_{er}^2}\mathrm{tr}(\mathbf{G}_e \mathbf{Q}_s)} \right), \tag{5.25}$$

where equality (a) in (5.25) is derived based on the fact that the random variable $\mathbf{h}_e^H \mathbf{Q}_s \mathbf{h}_e$ follows an exponential distribution with mean $\mathrm{tr}(\mathbf{G}_e \mathbf{Q}_s)$. Similarly, the left hand side of the constraints in (5.22c) and (5.22d) can be recast respectively as

$$(5.22c) \Rightarrow 1 - \exp\left(\frac{1 - (1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s)2^{-R_d}}{\frac{1}{\sigma_{er}^2}\mathrm{tr}(\mathbf{G}_e \mathbf{Q}_s)} \right) \geq 1 - \beta, \tag{5.26}$$

$$(5.22d) \Rightarrow 1 - \exp\left(\frac{\zeta_{eh}\sigma_{er}^2 - \omega_s}{\zeta_{eh}\mathrm{tr}(\mathbf{G}_e \mathbf{Q}_s)} \right) \leq \gamma. \tag{5.27}$$

Hence, the original problem in (5.22) can be rewritten as

$$\max_{\mathbf{Q}_s, \delta} \delta$$

$$s.t. \ \exp\left(\frac{1 - (1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)2^{-\delta(\text{tr}(\mathbf{Q}_s) + P_c)}}{\frac{1}{\sigma_{er}^2}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)}\right) \leq \alpha,$$

$$\exp\left(\frac{1 - (1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)2^{-R_d}}{\frac{1}{\sigma_{er}^2}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)}\right) \leq \beta,$$

$$1 - \exp\left(\frac{\zeta_{eh}\sigma_{er}^2 - \omega_s}{\zeta_{eh}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)}\right) \leq \gamma, \ (5.22e), (5.22f). \tag{5.28}$$

By performing some matrix manipulations (i.e., arithmetic and inequality operations), the following equivalent optimization problem is derived:

$$\max_{\mathbf{Q}_s} \frac{\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - \log_2(1 - \frac{1}{\sigma_{su}^2}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)\ln\alpha)}{\text{tr}(\mathbf{Q}_s) + P_c} \tag{5.29a}$$

$$s.t. \ 1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s \geq 2^{R_d} - \frac{2^{R_d}}{\sigma_{er}^2}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)\ln\beta, \tag{5.29b}$$

$$\zeta_{eh}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)\ln(1 - \gamma) \leq \zeta_{eh}\sigma_{er}^2 - \omega_s, \tag{5.29c}$$

$$(5.22e), (5.22f). \tag{5.29d}$$

In the following subsections, an approach is presented to solve this equivalent problem based on non-linear fractional and DC programming.

5.4.1 Non-linear Fractional Programming

Following a similar approach as in (5.13), first a parametric problem is defined w.r.t. λ, as follows:

$$F_s(\lambda) = \max_{\mathbf{Q}_s} \log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - \log_2(1 - \frac{1}{\sigma_{er}^2}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)\ln\alpha) - \lambda(\text{tr}(\mathbf{Q}_s) + P_c),$$

$$s.t. \ (5.29b), (5.29c), (5.22e), (5.22f). \tag{5.30}$$

The value of λ_i at the ith iteration can be chosen as the SEE obtained at the previous iteration. Furthermore, λ_i should be updated at each iteration as $\lambda_{i+1} = \frac{R_{st}^i(\mathbf{Q}_s)}{P_{st}^i(\mathbf{Q}_s)}$, where $R_{st}^i(\mathbf{Q}_s)$ and $P_{st}^i(\mathbf{Q}_s)$ denote the achieved secrecy rate and total power consumption that are obtained by solving (5.30) for the given λ_i, respectively. In practice, the iterative process will be repeated

until the following inequality is satisfied:

$$\Delta F_s(\lambda) = |R_s^i(\mathbf{Q}_s) - \lambda_i P_t^i(\mathbf{Q}_s)| \leq \varsigma, \tag{5.31}$$

with a small convergence tolerance $\varsigma > 0$.

5.4.2   DC Programming

Next, DC programming is exploited to solve the optimization problem defined in (5.30). Based on this approximation, the problem can be converted into the following equivalent problem:

$$\max_{\mathbf{Q}_s} \{\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - f_s(\mathbf{Q}_s, \mathbf{Q}_s^k) - \lambda_i[\mathrm{tr}(\mathbf{Q}_s) + P_c]\}$$
$$s.t.\ (5.29\text{b}), (5.29\text{c}), (5.22\text{e}), (5.22\text{f}), \tag{5.32}$$

where

$$f_s(\mathbf{Q}_s, \mathbf{Q}_s^k) = \log_2(1 - \frac{1}{\sigma_{er}^2}\mathrm{tr}(\mathbf{G}_e\mathbf{Q}_s^k)\ln\alpha)) - \frac{\frac{\ln\alpha}{\sigma_{er}^2}\mathrm{tr}(\mathbf{G}_e(\mathbf{Q}_s - \mathbf{Q}_s^k))}{(1 - \frac{1}{\sigma_{er}^2}\mathrm{tr}(\mathbf{G}_e\mathbf{Q}_s^k)\ln\alpha)\ln 2}. \tag{5.33}$$

Then, by exploiting SDR, the relaxed problem can be defined by dropping the rank-one constraint $\mathrm{rank}(\mathbf{Q}_s) = 1$ in (5.32) as follows:

$$\max_{\mathbf{Q}_s} \{\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - f_s(\mathbf{Q}_s, \mathbf{Q}_s^k) - \lambda_i[\mathrm{tr}(\mathbf{Q}_s) + P_c]\}$$
$$s.t.\ (5.29\text{b}), (5.29\text{c}), (5.22\text{e}), \mathrm{tr}(\mathbf{Q}_s) \leq P_{\mathrm{tx}}, \mathbf{Q}_s \succeq 0. \tag{5.34}$$

Proposition 5.2: Provided that the problem (5.34) is feasible, the optimal solution will be always rank-one.

   Proof: Please refer to Section 5.8.3.   ∎

This approximated problem is convex in terms of $\mathbf{Q}_s$. Hence, the suboptimal solution $\mathbf{Q}_s^*$ of the original problem in (5.22) can be obtained through iteratively solving problem (5.34) and updating $\mathbf{Q}_s^k$ at each iteration based on the solution obtained from the previous iteration. The proposed algorithm based on DC programming is provided in Algorithm 2.

Algorithm 2 Iterative algorithm for solving (5.34)

1: Initialize $i = 0$ and choose an initial value λ_0 for λ;
2: **repeat**$\leftarrow$ Outer loop (Dinkelbach's algorithm)
3:      Initialize $k = 0$, choose an initial value $\mathbf{Q}_s^k = 0$ and $F_s(\lambda)^{i,k} = 0$;
4:      **repeat**$\leftarrow$ Inner loop (DC programming)
5:          Solve the problem (5.34) with $\lambda = \lambda_i$ and obtain $\mathbf{Q}_s^{k+1}$;
6:          Compute $F_s(\lambda)^{i,k+1} = \log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s^{k+1}\mathbf{h}_s) - \log_2(1 - \frac{1}{\sigma_{er}^2}\text{tr}(\mathbf{G}_e\mathbf{Q}_s^{k+1})\ln\alpha) - \lambda_i(\text{tr}(\mathbf{Q}_s^{k+1}) + P_c)$;
7:          $\Delta\mu = F_s(\lambda)^{i,k+1} - F(\lambda)^{i,k}$;
8:          Update $k = k + 1$;
9:      **until** $|\Delta\mu| \le \zeta$;
10:     Update $\lambda_{i+1} = \frac{R_{st}^i(\mathbf{Q}_s)}{P_{st}^i(\mathbf{Q}_s)}$;;
11:     $i = i + 1$;
12: **until** (5.31) satisfied;
13: Return $\lambda^* = \lambda_i$, $P_{st}^*(\mathbf{Q}_s) = P_{st}^{i-1}(\mathbf{Q}_s)$, $R_{st}^*(\mathbf{Q}_s) = R_{st}^{i-1}(\mathbf{Q}_s)$;
14: Output λ_i.

5.4.3   Computational Complexity

The problem defined in (5.34) has four linear constraints, and one LMI constraint of size N_T. By utilizing the same method in the literature [166, 193], the computational complexity of the solution for (5.22) should be in the order of $TK\sqrt{N_T + 4}[n^3 + n(N_T^3 + 4) + n^2(N_T^2 + 4)]\ln(1/\epsilon)$, where T and K are the numbers of iterations of the inner and the outer loop, respectively.

5.5   SEE Maximization with Imperfect CSI

In this section, a tractable approach is developed to solve the robust SEE maximization problem with imperfect CSI on all set of channels. First, this robust problem is reformulated into a simple parametric problem by exploiting a well-known non-linear fractional programming. Then, it is showed that the optimal solution can be obtained by solving a series of SDP. The actual channel coefficients can be modelled by channel uncertainties as follows:

$$\mathbf{h}_s = \hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s, \; \mathbf{h}_p = \hat{\mathbf{h}}_p + \hat{\mathbf{e}}_p, \; \mathbf{h}_e = \hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e, \tag{5.35}$$

where $\hat{\mathbf{e}}_s$, $\hat{\mathbf{e}}_p$ and $\hat{\mathbf{e}}_e$ represent the channel uncertainties. Furthermore, these channel uncertainties can be defined based on an ellipsoid model as

$$||\hat{\mathbf{e}}_s||_2 = ||\mathbf{h}_s - \hat{\mathbf{h}}_s||_2 \leq \epsilon_s, \tag{5.36}$$

$$||\hat{\mathbf{e}}_p||_2 = ||\mathbf{h}_p - \hat{\mathbf{h}}_p||_2 \leq \epsilon_p, \tag{5.37}$$

$$||\hat{\mathbf{e}}_e||_2 = ||\mathbf{h}_e - \hat{\mathbf{h}}_e||_2 \leq \epsilon_e, \tag{5.38}$$

where $\epsilon_s \geq 0$, $\epsilon_p \geq 0$ and $\epsilon_e \geq 0$ are Euclidean norm-based error bounds. By incorporating these channel uncertainties in the original design, the SEE maximization problem can be reformulated into the following robust optimization framework:

$$\max_{\mathbf{Q}_s} \frac{R_m(\mathbf{Q}_s)}{P_m(\mathbf{Q}_s)} \tag{5.39a}$$

$$s.t. \ \log_2(1 + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{Q}_s(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)) - \log_2(1 + \frac{1}{\sigma_{er}^2}(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e)^H \mathbf{Q}_s(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e)) \geq R_d, \tag{5.39b}$$

$$\zeta_{eh}((\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e)^H \mathbf{Q}_s(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e) + \sigma_{er}^2) \geq \omega_s, \tag{5.39c}$$

$$(\hat{\mathbf{h}}_p + \hat{\mathbf{e}}_p)^H \mathbf{Q}_s(\hat{\mathbf{h}}_p + \hat{\mathbf{e}}_p) \leq P_f, \tag{5.39d}$$

$$\text{tr}(\mathbf{Q}_s) \leq P_{\text{tx}}, \mathbf{Q}_s \succeq \mathbf{0}, \text{rank}(\mathbf{Q}_s) = 1, \tag{5.39e}$$

where $R_m(\mathbf{Q}_s) = \log_2(1 + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{Q}_s(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)) - \log_2(1 + \frac{1}{\sigma_{er}^2}(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e)^H \mathbf{Q}_s(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e))$ and $P_m = \text{tr}(\mathbf{Q}_s) + P_c$. Next a two stage reformulation is presented to solve the above robust problem.

5.5.1 A Two-stage Reformulation of Problem (5.39)

First, Theorem 5.1, Lemma 5.1 and Lemma 5.2 are exploited to recast the problem in (5.39) into the following parametric problem:

$$\max_{\mathbf{Q}_s} \ [R_m(\mathbf{Q}_s) - \lambda P_m(\mathbf{Q}_s)]$$

$$s.t. \ (5.39b)-(5.39e). \tag{5.40}$$

It is obvious that the original problem (5.39) is transformed into a parameterized polynomial subtractive form. As a result, the original problem is reformulated to determine λ^* and $\mathbf{Q}_s^*$ to satisfy the condition defined in (5.11). Furthermore, Dinklebach's algorithm is employed, which iteratively solves the problem defined in (5.40) for a given λ and updates λ in the next

iteration as follows:

$$\lambda_{i+1} = \frac{R_m^i(\mathbf{Q}_s)}{P_m^i(\mathbf{Q}_s)}, \tag{5.41}$$

where $R_m^i(\mathbf{Q}_s)$ and $P_m^i(\mathbf{Q}_s)$ denote the achieved secrecy rate and total power consumption of (5.39) for the given λ_i, respectively, and subscript i denotes the iteration number. This iterative process will be terminated once the following condition is satisfied:

$$\Delta F_m(\lambda) = |R_m^i(\mathbf{Q}_s) - \lambda_i P_m^i(\mathbf{Q}_s)| \leq \rho, \tag{5.42}$$

where $\rho > 0$ is the termination threshold. By introducing a new slack variable τ, the following equivalent problem of (5.40) is derived:

$$F_m(\lambda) = \max_{\mathbf{Q}_s, \tau} \log_2 \left(\frac{1 + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{Q}_s(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)}{\tau} \right) - \lambda[\text{tr}(\mathbf{Q}_s) + P_c]\}$$

$$s.t.\ 1 + \frac{1}{\sigma_{er}^2}(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e)^H \mathbf{Q}_s(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e) \leq \tau, \tag{5.43a}$$

$$\log_2(1 + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{Q}_s(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)) - \log_2(\tau) \geq R_d, \tag{5.43b}$$

$$(5.39c)\text{-}(5.39e). \tag{5.43c}$$

Next, by introducing a new slack variable ν, the above problem is transformed into a two-stage problem, namely, outer problem and inner problem. The outer problem can be defined as

$$F_m(\lambda) = \max_{\nu}\ \{\log_2 \Theta(\lambda, \nu) - \nu\}$$

$$s.t.\ \lambda(\nu_{min} + P_c) \leq \nu \leq \lambda(P_{tx} + P_c), \tag{5.44}$$

where ν_{min} denotes the minimum value of ν which can be obtained by solving the power minimization problem defined in (5.52). Furthermore, the inner problem determines $\Theta(\lambda, \nu)$ for a given ν, which can be defined as

$$\Theta(\lambda, \nu) = \max_{\mathbf{Q}_s, \tau} \frac{1 + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{Q}_s(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)}{\tau}$$

$$s.t.\ \lambda(\text{tr}(\mathbf{Q}_s) + P_c) \leq \nu,\ (5.43a)\text{-}(5.43c). \tag{5.45}$$

However, the outer problem (5.44) is a single variable optimization problem with a set of feasible values $\nu \in [\lambda(\nu_{min} + P_c), \lambda(P_{tx} + P_c)]$. It is well-known that the optimal value can be efficiently obtained through an one-dimensional search. Furthermore, it is worthy note

that the one-dimensional search approach is a kind of exhaustive algorithms, which has the capability to provide global optimal solution [169, 11]. By relaxing the non-convex rank-one constraint, the relaxed inner problem can be defined as

$$\Theta(\lambda, \nu) = \max_{\mathbf{Q}_s, \tau} \frac{1 + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{Q}_s(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)}{\tau}$$

$$s.t.\ (5.43a), (5.43b), (5.39c), (5.39d),$$

$$\lambda(\text{tr}(\mathbf{Q}_s) + P_c) \leq \nu, \text{tr}(\mathbf{Q}_s) \leq P_{\text{tx}}, \mathbf{Q}_s \succeq \mathbf{0}. \tag{5.46}$$

Proposition 5.3: Provided that the problem (5.46) is feasible, the optimal solution will be always rank-one.

Proof: Please refer to Section 5.8.4.   ∎

In the following subsection, an SDP formulation based on the Charnes-Cooper transformation [194] and S-Procedure [47] is presented to solve the above relaxed inner problem.

5.5.2   SDP-based Reformulation of the Inner Problem (5.46)

First, the Charnes-Cooper transformation [194] is introduced as

$$\mathbf{W} = \frac{\mathbf{Q}_s}{\tau}, \Gamma = \frac{1}{\tau}, \tag{5.47}$$

and a slack variable ϱ to recast the inner problem (5.46) as

$$\Theta(\lambda, \nu) = \max_{\mathbf{W}, \Gamma, \varrho} \varrho$$

$$s.t.\ \Gamma + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{W}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s) \geq \varrho, \tag{5.48a}$$

$$\Gamma + \frac{1}{\sigma_{er}^2}(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e)^H \mathbf{W}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s) \leq 1, \tag{5.48b}$$

$$\Gamma + \frac{1}{\sigma_{su}^2}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{W}(\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s) \geq 2^{R_d}, \tag{5.48c}$$

$$\zeta_{eh}((\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e)^H \mathbf{W}(\hat{\mathbf{h}}_e + \hat{\mathbf{e}}_e) + \Gamma\sigma_e^2) \geq \omega_s\Gamma, \tag{5.48d}$$

$$(\hat{\mathbf{h}}_p + \hat{\mathbf{e}}_p)^H \mathbf{W}(\hat{\mathbf{h}}_p + \hat{\mathbf{e}}_p) \leq \Gamma P_f, \tag{5.48e}$$

$$\lambda(\text{tr}(\mathbf{W}) + \Gamma P_c) \leq \Gamma\nu, \text{tr}(\mathbf{W}) \leq \Gamma P_{\text{tx}}, \mathbf{W} \succeq \mathbf{0}. \tag{5.48f}$$

To further proceed with this problem, the following lemma is required:

Lemma 5.3: (*S-Procedure* [47]) Define $f_i(\mathbf{x})$, $i = 1, 2$ such as

$$f_i(\mathbf{x}) = \mathbf{x}^H \mathbf{A}_i \mathbf{x} + 2Re\{\mathbf{b}_i^H \mathbf{x}\} + c_i, \tag{5.49}$$

in which $\mathbf{x} \in \mathcal{C}^{N_T \times 1}$, $\mathbf{A}_i \in \mathcal{C}^{N_T \times N_T}$, $\mathbf{b}_i \in \mathcal{C}^{N_T \times 1}$ and $c_i \in \mathcal{R}$. The implication $f_1(\mathbf{x}) \geq 0 \rightarrow f_2(\mathbf{x}) \geq 0$ holds if and only if there exists a $\vartheta \geq 0$ such that

$$\begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^H & c_2 \end{bmatrix} - \vartheta \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & c_1 \end{bmatrix} \succeq \mathbf{0}. \tag{5.50}$$

By applying Lemma 5.3, the inner problem defined in (5.48) can be recast as

$$\Theta(\lambda, \nu) = \max_{\mathbf{W}, \tau, \varrho} \varrho \tag{5.51a}$$

$$s.t. \begin{bmatrix} \vartheta_1 \mathbf{I} + \frac{1}{\sigma_{su}^2} \mathbf{W} & \frac{1}{\sigma_{su}^2} \mathbf{W} \hat{\mathbf{h}}_s \\ \frac{1}{\sigma_{su}^2} \hat{\mathbf{h}}_s^H \mathbf{W} & t_1 \end{bmatrix} \succeq \mathbf{0}, \tag{5.51b}$$

$$\begin{bmatrix} \vartheta_2 \mathbf{I} - \frac{1}{\sigma_{er}^2} \mathbf{W} & -\frac{1}{\sigma_{er}^2} \mathbf{W} \hat{\mathbf{h}}_e \\ -\frac{1}{\sigma_{er}^2} \hat{\mathbf{h}}_e^H \mathbf{W} & t_2 \end{bmatrix} \succeq \mathbf{0}, \tag{5.51c}$$

$$\begin{bmatrix} \vartheta_3 \mathbf{I} + \frac{1}{\sigma_{su}^2} \mathbf{W} & \frac{1}{\sigma_{su}^2} \mathbf{W} \hat{\mathbf{h}}_s \\ \frac{1}{\sigma_{su}^2} \hat{\mathbf{h}}_s^H \mathbf{W} & t_3 \end{bmatrix} \succeq \mathbf{0}, \tag{5.51d}$$

$$\begin{bmatrix} \vartheta_4 \mathbf{I} + \zeta_{eh} \mathbf{W} & \zeta_{eh} \mathbf{W} \hat{\mathbf{h}}_e \\ \zeta_{eh} \hat{\mathbf{h}}_e^H \mathbf{W} & t_4 \end{bmatrix} \succeq \mathbf{0}, \tag{5.51e}$$

$$\begin{bmatrix} \vartheta_5 \mathbf{I} - \mathbf{W} & -\mathbf{W} \hat{\mathbf{h}}_p \\ -\hat{\mathbf{h}}_p^H \mathbf{W} & t_5 \end{bmatrix} \succeq \mathbf{0}, \tag{5.51f}$$

$$(5.48f), \vartheta_1 \geq 0, \vartheta_2 \geq 0, \vartheta_3 \geq 0, \vartheta_4 \geq 0, \vartheta_5 \geq 0, \tag{5.51g}$$

where $t_1 = \frac{1}{\sigma_{su}^2} \hat{\mathbf{h}}_s^H \mathbf{W} \hat{\mathbf{h}}_s + \Gamma - \varrho - \vartheta_1 \epsilon_s^2$, $t_2 = 1 - \frac{1}{\sigma_{er}^2} \hat{\mathbf{h}}_e^H \mathbf{W} \hat{\mathbf{h}}_e - \Gamma - \vartheta_2 \epsilon_e^2$, $t_3 = \frac{1}{\sigma_{su}^2} \hat{\mathbf{h}}_s^H \mathbf{W} \hat{\mathbf{h}}_s - 2^{R_d} + \Gamma - \vartheta_3 \epsilon_s^2$, $t_4 = \zeta_{eh} \hat{\mathbf{h}}_e^H \mathbf{W} \hat{\mathbf{h}}_e + \zeta_{eh} \Gamma \sigma_e^2 - \omega_s \Gamma - \vartheta_4 \epsilon_e^2$ and $t_5 = \Gamma P_f - \hat{\mathbf{h}}_p^H \mathbf{W} \hat{\mathbf{h}}_p - \vartheta_5 \epsilon_p^2$. This inner problem in (5.51) is convex and can be efficiently solved through existing convex optimization software [152]. Furthermore, υ_{min} can be obtained through solving the following convex problem

$$\upsilon_{min} = \min_{\mathbf{Q}_s, \Gamma} \text{tr}(\mathbf{Q}_s)$$

$$s.t. \ (5.51c) - (5.51f), \ (5.48f),$$

$$\vartheta_2 \geq 0, \vartheta_3 \geq 0, \vartheta_4 \geq 0, \vartheta_5 \geq 0. \tag{5.52}$$

The above developed procedures for solving this robust SEE maximization problem are summarized in Algorithm 3.

Algorithm 3 Dinkelbach's algorithm for solving the SEE maximization problem in (5.39)

 1: Initialization $i = 0$ and $\lambda = \lambda_0$;
 2: **repeat**
 3:     Perform a one-dimensional search over ν to obtain the optimal values $(\nu^*, \Theta(\lambda_i, \nu^*))$ for the outer problem in (5.44), where each $\Theta(\lambda_i, \nu)$ is obtained by solving the inner problem in (5.51);
 4:     Retrieve the corresponding $\mathbf{Q}_s^*$ through (5.47);
 5:     Compute $F(\lambda_i) = R_m^i(\mathbf{Q}_s^*) - \lambda_i P_m^i(\mathbf{Q}_s^*)$;
 6:     Update λ_{i+1} by using (5.41);
 7:     $i = i + 1$;
 8: **until** the condition in (5.42) is satisfied;
 9: Return the optimal value of the inner problem in (5.51) to the outer problem in (5.44);
10: Output the maximum SEE obtained by the outer problem in (5.44).

5.5.3 Computational Complexity

The problem defined in (5.51) has five LMI constraints of size $N_T + 1$, one LMI constraint of size N_T, and seven linear constraints. The number of variables n is on the order of N_T^2. By following the same method in [166, 193], the computational complexity of solving the problem defined in (5.39) should be on the order of $TK\sqrt{6N_T + 12}\{n^3 + n[5(N_T + 1)^3 + N_T^3 + 7] + n^2[5(N_T + 1)^2 + N_T^2 + 7]\}\ln(1/\epsilon)$, where T and K are the numbers of iterations of the one dimensional search and the Dinkelbach's algorithm, respectively.

5.6 Simulation Results

To validate the proposed approaches, a number of simulations are performed based on the proposed model, but with different configurations. An underlay MISO CR network is considered, where the multi-antenna SU-Tx simultaneously transmits confidential information and energy to the SU-Rx and ER, respectively. This communication is established by sharing the spectrum that has been assigned for signal transmission between a PU-Tx and a PU-Rx. It is assumed that the SU-Tx is equipped with three ($N_t = 3$) antennas, while the PU-Rx,
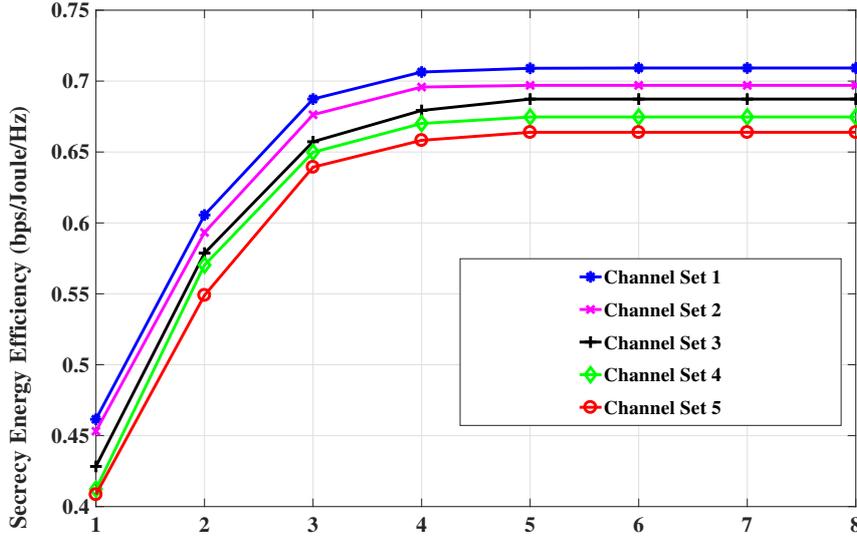
Fig. 5.2 Convergence of the proposed algorithm with perfect CSI assumption in Algorithm 1 for different sets of channels.

SU-Rx and ER each has a single antenna. All the channel coefficients are generated by CSCG with zero mean and unit variance.[1] The maximum interference leakage to the PU-Rx is set to be 0 dBW. Furthermore, the energy conversion ratio is assumed to be $\zeta_{eh} = 0.5$. The noise variances at the SU-Rx and ER are set to $\sigma_{su}^2 = 1$ and $\sigma_{er}^2 = 1$, respectively. The convergence tolerances ε, ς and ρ defined in (5.16), (5.31) and (5.42), respectively, are assumed to be 10^{-3}.

Note that in majority of the simulation scenarios outlined here, the difference between the transmit power and the EH requirement is large, and thus, may raise concerns on the validity of this assumption ($\omega_s = -20$ dBW). However, this assumption being is reasonable, for the following two reasons: 1) similar assumptions (large gap between transmit power and EH requirements) have been widely made in the literature, for instance in [198, 199, 84]; and 2) for the secondary system in CR network, the ER could be wireless sensors or IoT devices. For these devices, only a power in the order of milliwatts is sufficient [60].

First, the convergence of the proposed algorithms are evaluated by randomly generating different five sets of channels for each CSI assumption (perfect CSI, statistical CSI on the ER channel and imperfect CSI. Figure 5.2 presents the convergence of the achieved SEE with perfect CSI obtained by Algorithm 1. Here, the target secrecy rate (R_d), the transmit power consumption (P_{tx}) and the EH requirement (ω_s) have been set to 0.5 bps/Hz, 20 dBW and $\omega_s = -20$ dBW, respectively. Then the convergence of achieved SEE with statistical CSI on

---

[1]Similar assumptions have been employed in the literature [195, 169]. In practical communication scenarios, this assumption may be implemented in wireless sensors and mobile wearable devices [196, 197].
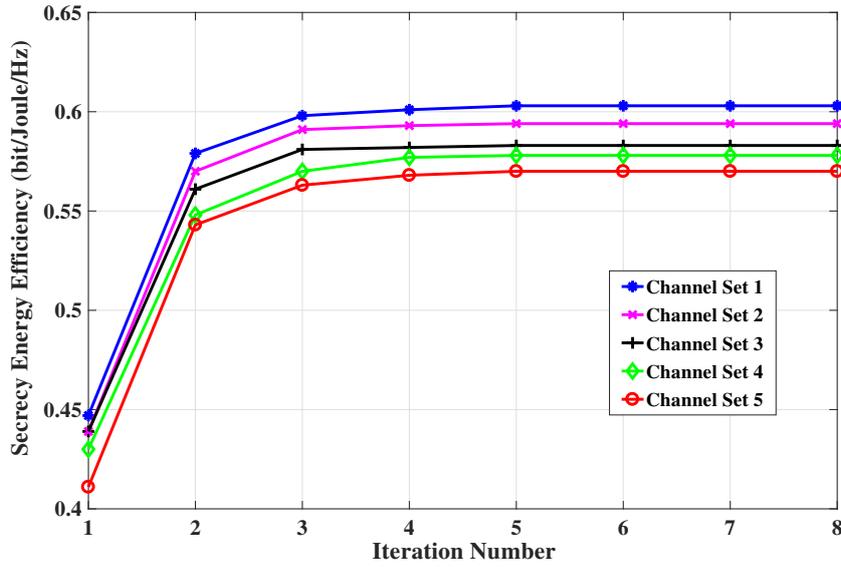
Fig. 5.3 Convergence of the proposed algorithm with the assumption of statistical CSI on the ER channel in Algorithm 2 for different sets of channels.

the ER's channel, and the convergence of the achieved SEE with imperfect CSI are showed in Fig. 5.3 and 5.4. Here, the system parameters are: target secrecy rate $R_d = 0.5$ bps/Hz, the transmit power consumption $P_{\text{tx}} = 20$ dBW and the EH requirement $\omega_s = -20$ dBW. As can be observed, the maximum SEE and the convergence of the proposed algorithms can be attained with a limited number of iterations.

Figure 5.5 shows the achieved SEE with different target secrecy rates, and EH requirements under the assumption of perfect CSI. As seen in Fig. 5.5, the optimal SEE decreases as the target rate increases. Note that zero SEE means that the original SEE maximization problem is infeasible; this is due to unachievable target secrecy rate constraint. Furthermore, it is possible to achieve a better SEE performance with a reduced EH requirement. Note that the same SEE performance can be achieved with different available transmit power as in Fig. 5.5 provided the SEE maximization problem is feasible with the same rest of the constraints. Hence, increasing the transmit power consumption cannot provide a better SEE. However, it can achieve a higher secrecy rate.

The performance comparisons between the proposed schemes and similar schemes available in the literature is also performed. Fig. 5.6 provides comparisons of the achievable SEE of three different transmission schemes under the perfect CSI assumption: SEE maximization, power minimization and secrecy rate maximization. The power minimization problem is formulated as

$$\min_{\mathbf{Q}_s} \ \text{tr}(\mathbf{Q}_s)$$
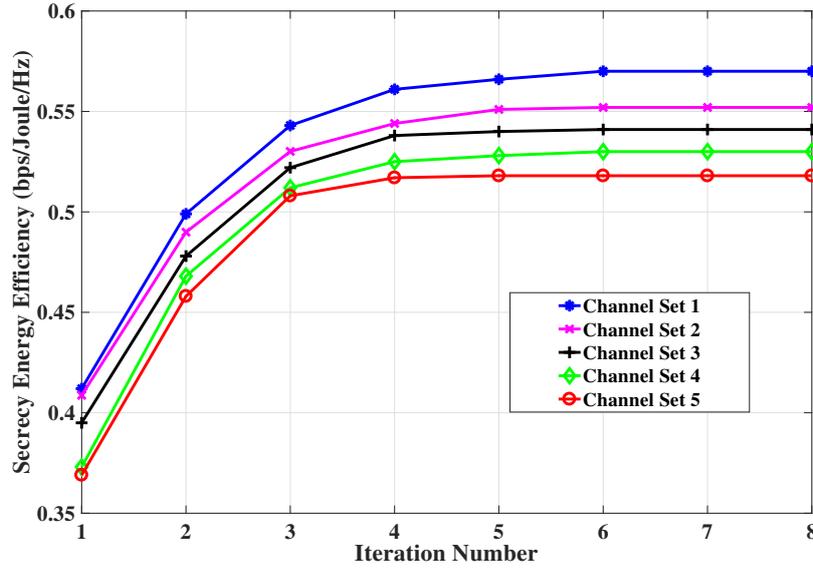
Fig. 5.4 Convergence of the proposed algorithm with norm-error bounded imperfect CSI assumption in Algorithm 3 for different sets of channels.
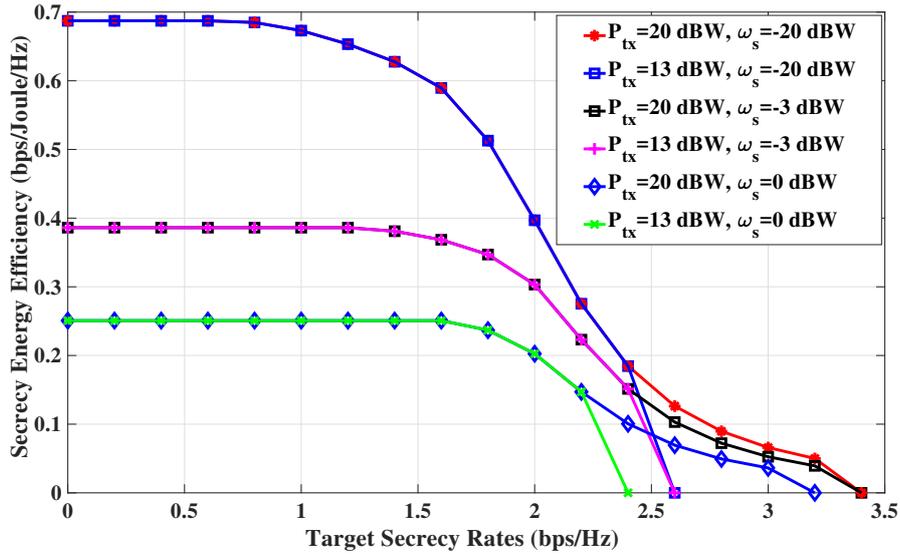
Fig. 5.5 The achieved SEE with different target secrecy rates, transmit power constraints and EH requirements under the perfect CSI assumption.

$$s.t. \ (5.9b)\text{-}(5.9d), \ \mathbf{Q}_s \succeq \mathbf{0}, \text{tr}(\mathbf{Q}_s) \leq P_{\text{tx}}, \tag{5.53}$$

and the secrecy rate maximization problem is written as

$$\max_{\mathbf{Q}_s} \ \log_2\left(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{q}\mathbf{q}^H\mathbf{h}_s\right) - \log_2\left(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{q}\mathbf{q}^H\mathbf{h}_e\right)$$
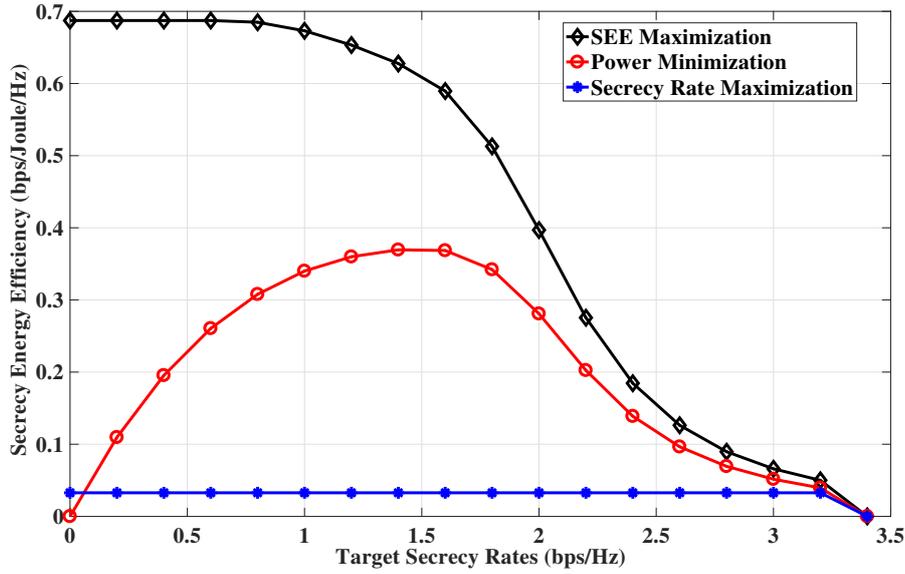
Fig. 5.6 The achieved SEE for different transmission schemes under the assumption of perfect CSI: SEE maximization, power minimization and secrecy rate maximization.

$$s.t. \ (5.9b)\text{-}(5.9d), \ \mathbf{Q}_s \succeq \mathbf{0}, \text{tr}(\mathbf{Q}_s) \leq P_{\text{tx}}. \tag{5.54}$$

In these simulation results, the maximum available transmit power and EH requirement are assumed to be 20 dBW and -20 dBW, respectively. As expected, the proposed scheme for SEE maximization outperforms the other two schemes in terms of achieved SEE. Furthermore, the achievable SEE remains constant in the secrecy rate maximization scheme with different target secrecy rate values within its feasible region. This trend is due to the fact that the secrecy rate maximization scheme always achieves a better constant secrecy rate than the target secrecy rate in all cases with a feasible target secrecy rate. In other words, the target secrecy rate constraint becomes inactive with the available transmit power, which is more than the required to achieve the target secrecy rate. Furthermore, zero SEE means that it is not feasible to achieve the target secrecy rate with the available transmit power.

Then the performance of the proposed scheme with statistical CSI on the channel of ER is performed. Figure 5.7 shows the achieved SEE with different target secrecy rates and EH requirements. The channel covariance matrix of ER is modelled as $\mathbf{G}_e = a^2\mathbf{I}$. The parameters $\alpha, \beta,$ and γ define the outage probabilities for SEE, secrecy rate and EH requirement, respectively, and are assumed to be the same with all constraints, i.e., $\alpha = \beta = \gamma = 0.1$. The parameter a^2 is introduced to control the ER's channel variance. Furthermore, the performance of the scheme with the perfect CSI is also evaluated to draw a performance comparison. As shown in Fig. 5.7, the scheme with perfect CSI achieves the best performance and the optimal achievable SEE decreases as the target secrecy rate
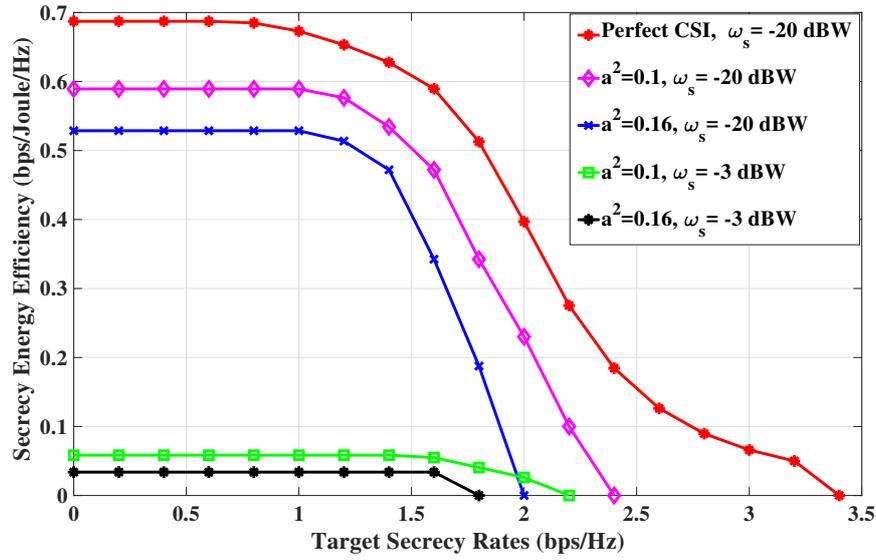
Fig. 5.7 The achieved SEE with different target secrecy rates, channel covariances a^2 and EH requirements with the assumption of statistical CSI of ER channel.
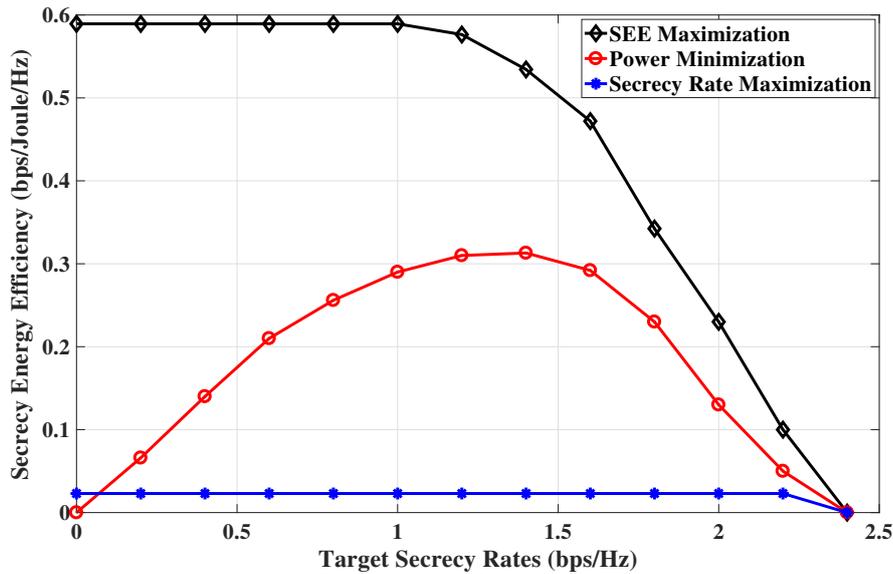
Fig. 5.8 The achieved SEE with different schemes under the assumption of statistical CSI of ER channel: SEE maximization, power minimization and secrecy rate maximization.

increases for all sets of the assumptions. Note that with a higher EH requirement and a^2, the proposed scheme achieves less SEE due to more required transmit power to meet the EH constraints and deal with the channel variance.
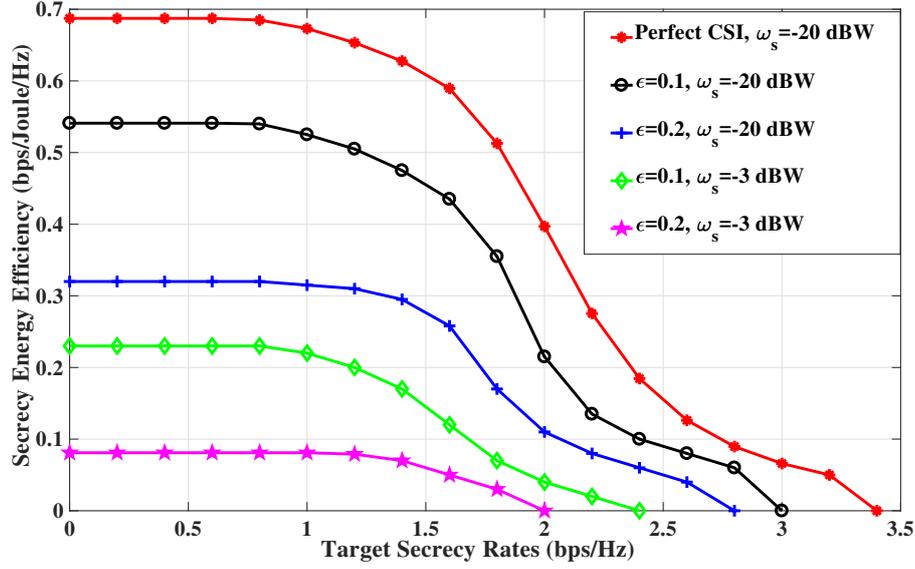
Fig. 5.9 The achieved SEE with different target secrecy rates, norm-error bounds and EH requirements under the assumption of imperfect CSI.

The achieved SEE of different schemes is evaluated under the assumption of statistical CSI on the ER's channel. Figure 5.8 presents the achievable SEE of three schemes: SEE maximization, power minimization and secrecy rate maximization. Here, it is assumed that the maximum available transmit power and EH requirement are set to be 20 dBW and -20 dBW, respectively. The power minimization problem is defined as

$$\min_{\mathbf{Q}_s}\ \mathrm{tr}(\mathbf{Q}_s)$$
$$s.t.\ (5.29b)\text{-}(5.29d),\qquad\qquad(5.55)$$

and the secrecy rate maximization problem is written as

$$\max_{\mathbf{Q}_s}\ \log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)-\log_2(1-\frac{1}{\sigma_{su}^2}\mathrm{tr}(\mathbf{G}_e\mathbf{Q}_s)\ln\alpha)$$
$$s.t.\ (5.29b)\text{-}(5.29d).\qquad\qquad(5.56)$$

As expected, the proposed SEE maximization scheme outperforms the other two schemes in terms of the achieved SEE. Furthermore, the SEE performance obtained with the secrecy rate maximization scheme is not affected by different target secrecy rates in its feasible secrecy rate region. The reason is that in secrecy rate maximization scheme, all the available transmit power is consumed to achieve the maximum secrecy rate. Thus, the achievable SEE would remain a constant with a given available transmit power.
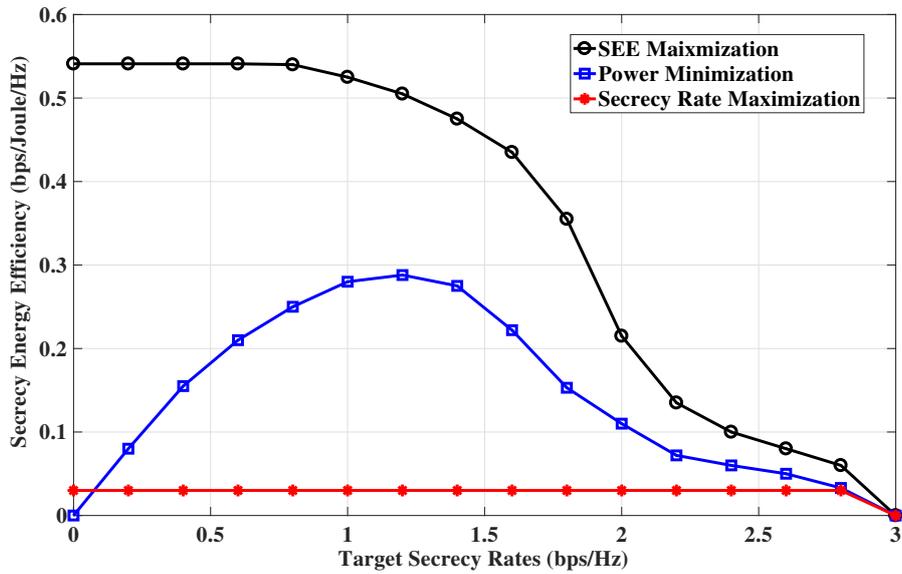
Fig. 5.10 The achieved SEE with different schemes under the assumption imperfect CSI: SEE maximization, power minimization and secrecy rate maximization.
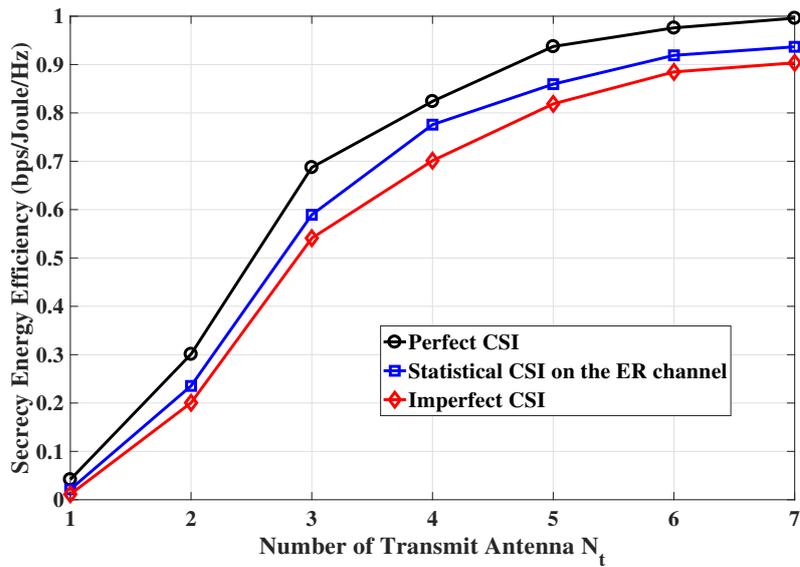
Fig. 5.11 The achieved SEE with different number of transmit antennas with perfect CSI, statistical CSI on the ER channel and imperfect CSI assumptions.

Next the performance of the proposed scheme is evaluated under the assumption of the norm-bounded channel uncertainties in section 5.5. Figure 5.9 illustrates the achieved SEE with different target secrecy rates and EH requirements. The channel bounds of the uncertainties are assumed to be the same, i.e., $\epsilon_s = \epsilon_p = \epsilon_e = \epsilon$. The maximum available transmit power is set to be 20 dBW. As seen in Fig. 5.9, the optimal SEE decreases as the
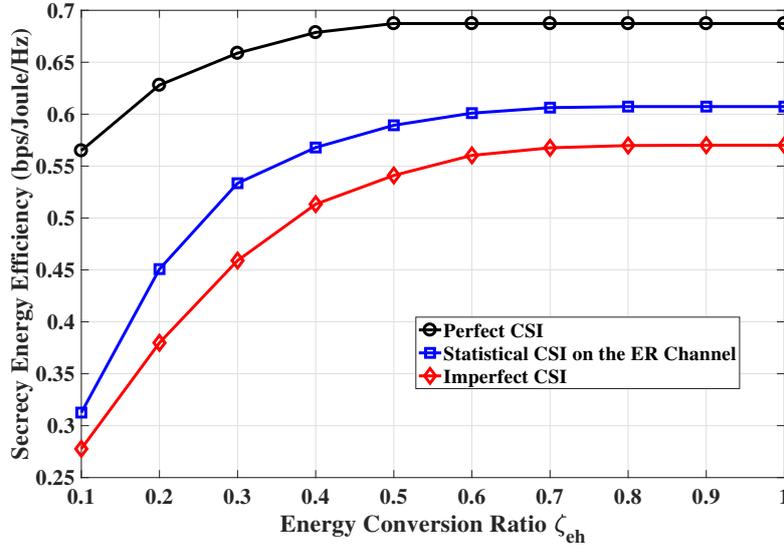
Fig. 5.12 The achieved SEE with different energy conversion ratio ζ_{eh} with perfect CSI, statistical CSI on the ER channel and imperfect CSI assumptions.

target rate increases. The scheme with perfect CSI can be seen as a special case of the robust SEE maximization (with a zero error bound on the channel uncertainties, i.e., $\epsilon = 0$). Similar to the previous results, zero SEE refers to an infeasible problem with a given target secrecy rate constraint. Furthermore, the optimal SEE has a better performance with a smaller EH requirement. Moreover, a better SEE can be realized for a feasible problem with a given target secrecy rate constraint and with a smaller error bound ϵ. This is due to the fact that the SU-Tx requires more transmit power to deal with a larger channel uncertainty bound and to meet all set of constraints.

Then the performance comparison of the achieved SEE with different transmission schemes under norm-error bounded channel uncertainties is provided in Fig. 5.10. Here, three different schemes are considered: SEE maximization, secrecy rate maximization and power minimization. Furthermore, the maximum available transmit power and the EH requirement are set to be 20 dBW and -20 dBW, respectively. The power minimization problem for this CSI assumption is defined in problem (5.52). The secrecy rate maximization problem is written as

$$\max_{\mathbf{W}, \tau, \varrho} \log_2(\varrho)$$
$$s.t. \ (5.51\text{c}) - (5.51\text{f}), \ (5.48\text{f}),$$
$$\vartheta_2 \geq 0, \vartheta_3 \geq 0, \vartheta_4 \geq 0, \vartheta_5 \geq 0. \tag{5.57}$$

As seen, the proposed SEE maximization scheme outperforms the other two schemes in terms of achieved SEE. Furthermore, the achievable SEE is not affected by the target secrecy rates and remains constant. This is due to the fact that all available transmit power is used to achieve the maximum secrecy rate in the secrecy maximization scheme, which is similar to that of the previous simulation results of different schemes under perfect and statistical CSI assumptions.

The achieved SEE with different number of transmit antennas for all CSI assumptions is presented in Fig. 5.11. The system parameters for all the cases are set to be as follows: target secrecy rate $R_d = 0.5$ bps/Hz, the transmit power consumption $P_{tx} = 20$ dBW and the EH requirement $\omega_s = -20$ dBW. Furthermore, the channel variance of ER for the assumption of statistical CSI on the ER's channel is set to be $a^2 = 0.1$ and the channel uncertainty bound is assumed to be $\epsilon = 0.1$ for the scheme with imperfect CSI, respectively. As shown in this figure, larger number of antennas at the SU-Tx provides better performance in terms of the SEE performance. This is due to the fact that larger number of transmit antennas brings more degree of freedom, which results in higher achievable secrecy rate without any additional transmit power consumption.

Finally, the SEE performance with different energy conversion ratio is provided for all three CSI assumptions in Fig. 5.12. Here, the target secrecy rate R_d, the transmit power consumption P_{tx} and the EH requirement ω_s are set to be 0.5 bps/Hz, 20 dBW and -20 dBW, respectively. Furthermore, the channel variance of ER for the scheme with statistical CSI on the ER channel and the channel uncertainty bound for the scheme with imperfect CSI are assumed to be $a^2 = 0.1$ and $\epsilon = 0.1$, respectively. As seen in Fig. 5.12, the achieved SEE increases as the energy conversion ratio ζ_{eh} increases for all three CSI assumptions. The reason is that the SU-Tx requires more power to meet the EH requirement at the ER with a small energy conversion ratio ζ_{eh} (i.e., $\zeta_{eh} = 0.1$).

5.7 Summary

In this chapter, the SEE maximization problems were considered for an underlay MISO CR network in the presence of energy harvesting. First, a transmit beamforming design is presented to meet the required secrecy rate at the SU-Rx while satisfying the interference leakage constraint on the PU-Rx and the EH requirement on the ER. The original problem was not convex due to the non-linear fractional objective function. To overcome this non-convexity issue, the original problem was reformulated into a convex one by exploiting SDR, non-linear fractional and DC programming. Then the beamforming design was extended to address

the problems of imperfect CSI at the transmitter. In particular, two robust designs were proposed with the assumptions of statistical CSI on the ER's channel and error bounded channel uncertainties, respectively. For the case of statistical CSI assumption, the outage probability constraints were firstly presented in a closed-form expression. By exploiting SDR, non-linear fractional and DC programming, the original problem was efficiently solved through an iterative approach. Next, the robust design with bounded channel uncertainties was formulated as a series of SDP by exploiting the same techniques as in the previous robust design. Then, the S-Procedure was used to convert this problem into a convex one. Simulation results were provided (1) to validate the convergence of the proposed algorithms, (2) to show the impacts of different parameters on the achieved SEE, including the maximum available transmit power, the EH requirement, the number of transmit antennas and the energy conversion ratio, and (3) to demonstrate the superior performance of the proposed SEE maximization scheme over power minimization and secrecy rate maximization schemes available in the literature.

5.8   Appendix

5.8.1   Proof of Theorem 5.1

Let $\mathbf{Q}_s^*$ be the optimal solution of problem (5.9), and S be the set including all feasible $\mathbf{Q}_s$ under the constraints defined in (5.9b)-(5.9e). Then, one can derive

$$\lambda^* = \frac{R_s(\mathbf{Q}_s^*)}{P_t(\mathbf{Q}_s^*)} \geq \frac{R_s(\mathbf{Q}_s)}{P_t(\mathbf{Q}_s)}, \forall \, \mathbf{Q}_s \in S. \tag{5.58}$$

Hence,

$$R_s(\mathbf{Q}_s) - \lambda^* P_t(\mathbf{Q}_s) \leq 0, \forall \, \mathbf{Q}_s \in S, \tag{5.59}$$

$$R_s(\mathbf{Q}_s^*) - \lambda^* P_t(\mathbf{Q}_s^*) = 0. \tag{5.60}$$

From the inequality in (5.59) $\max_{\mathbf{Q}_s}\{R_s(\mathbf{Q}_s) - \lambda^* P_t(\mathbf{Q}_s)\} = 0$ can be derived, and from (5.60) it is obvious that the maximum value is achieved with $\mathbf{Q}_s^*$. The sufficient condition has been satisfied, which completes the first part of the proof.

Next, the second part of the proof with the necessary condition is considered. Let $\mathbf{Q}_s^*$ be the solution for $\max_{\mathbf{Q}_s}\{R_s(\mathbf{Q}_s) - \lambda^* P_t(\mathbf{Q}_s)\} = 0$ and $R_s(\mathbf{Q}_s^*) - \lambda^* P_t(\mathbf{Q}_s^*) = 0$. This implies

the following:

$$R_s(\mathbf{Q}_s) - \lambda^* P_t(\mathbf{Q}_s) \leq 0, \forall \, \mathbf{Q}_s \in S, \tag{5.61}$$

$$R_s(\mathbf{Q}_s^*) - \lambda^* P_t(\mathbf{Q}_s^*) = 0. \tag{5.62}$$

From (5.61), it is obvious $\lambda^* \geq \frac{R_s(\mathbf{Q}_s)}{P_t(\mathbf{Q}_s)}$ for all $\mathbf{Q}_s \in S$, and λ^* is the maximum objective value of the problem defined in (5.9). This completes the proof of Theorem 5.1. ■

5.8.2 Proof of Proposition 5.1

First, the following Lagrangian function of the optimization problem defined in (5.19) is considered:

$$\mathcal{L}(\mathbf{Q}_s, \mathbf{Z}, \mu, \iota, \kappa, \varpi) = -\{\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - f(\mathbf{Q}_s, \mathbf{Q}_s^k) - \lambda[\mathrm{tr}(\mathbf{Q}_s) + P_c]\} - \mathrm{tr}(\mathbf{Z}\mathbf{Q}_s)$$

$$+ \mu(\mathrm{tr}(\mathbf{Q}_s) - P_{\mathrm{tx}}) - \iota[\log_2(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - \log_2(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e) - R_d] - \kappa[\zeta_{eh}(\mathbf{h}_e\mathbf{Q}_s\mathbf{h}_e$$

$$+ \sigma_{er}^2) - \omega_s] + \varpi(\mathbf{h}_p^H\mathbf{Q}_s\mathbf{h}_p - P_f), \tag{5.63}$$

where $\mathbf{Q}_s \in \mathbb{H}_+^{N_t}$, $\mathbf{Z} \in \mathbb{H}_+^{N_t}$, $\mu \in \mathbb{R}_+$, $\iota \in \mathbb{R}_+$, $\kappa \in \mathbb{R}_+$ and $\varpi \in \mathbb{R}_+$ are the Lagrangian multipliers associated with the constraints in the problem defined in (5.19). Then, the corresponding KKT conditions is derived as follows [152]:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{Q}_s} = -(\iota + 1)\left[\frac{\frac{1}{\sigma_{su}^2}\mathbf{h}_s\mathbf{h}_s^H}{(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)\ln 2}\right] + (\lambda + \mu)\mathbf{I} + \iota\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e)\ln 2}\right] -$$

$$\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e)\ln 2}\right] - \kappa\zeta_{eh}\mathbf{h}_e\mathbf{h}_e^H - \varpi\mathbf{h}_p\mathbf{h}_p^H - \mathbf{Z} = \mathbf{0}, \tag{5.64}$$

$$\mathbf{Z}\mathbf{Q}_s = \mathbf{0}, \mathbf{Z} \succeq \mathbf{0}. \tag{5.65}$$

Based on these KKT conditions, the following equality holds:

$$-(\iota + 1)\left[\frac{\frac{1}{\sigma_{su}^2}\mathbf{h}_s\mathbf{h}_s^H}{(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)\ln 2}\right] + (\lambda + \mu)\mathbf{I} - \varpi\mathbf{h}_p\mathbf{h}_p^H + \iota\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e)\ln 2}\right]$$

$$-\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e)\ln 2}\right] - \kappa\zeta_{eh}\mathbf{h}_e\mathbf{h}_e^H = \mathbf{Z}, \tag{5.66}$$

$$\Rightarrow \left\{-(\iota + 1)\left[\frac{\frac{1}{\sigma_{su}^2}\mathbf{h}_s\mathbf{h}_s^H}{(1 + \frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)\ln 2}\right] + (\lambda + \mu)\mathbf{I} + \iota\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1 + \frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e)\ln 2}\right] -$$

$$\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e)\ln 2}\right] - \kappa\zeta_{eh}\mathbf{h}_e\mathbf{h}_e^H - \varpi\mathbf{h}_p\mathbf{h}_p^H\Bigg\}\mathbf{Q}_s = \mathbf{0}, \tag{5.67}$$

$$\Rightarrow\Bigg\{(\lambda+\mu)\mathbf{I}+\iota\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e)\ln 2}\right] - \left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e)\ln 2}\right] - \kappa\zeta_{eh}\mathbf{h}_e\mathbf{h}_e^H$$

$$- \varpi\mathbf{h}_p\mathbf{h}_p^H\Bigg\}\mathbf{Q}_s = (\iota+1)\left[\frac{\frac{1}{\sigma_{su}^2}\mathbf{h}_s\mathbf{h}_s^H}{(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)\ln 2}\right]\mathbf{Q}_s, \tag{5.68}$$

$$\Rightarrow\mathbf{Q}_s = \Bigg\{(\iota+1)\left[\frac{\frac{1}{\sigma_{su}^2}\mathbf{h}_s\mathbf{h}_s^H}{(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)\ln 2}\right]\Bigg\}\Bigg\{(\lambda+\mu)\mathbf{I}+\iota\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e)\ln 2}\right] -$$

$$\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e)\ln 2}\right] - \kappa\zeta_{eh}\mathbf{h}_e\mathbf{h}_e^H - \varpi\mathbf{h}_p\mathbf{h}_p^H\Bigg\}^{-1}\mathbf{Q}_s. \tag{5.69}$$

Hence, the following rank relation holds:

$$\text{rank}(\mathbf{Q}_s) = \text{rank}\Bigg\{\Bigg\{(\iota+1)\left[\frac{\frac{1}{\sigma_{su}^2}\mathbf{h}_s\mathbf{h}_s^H}{(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)\ln 2}\right]\Bigg\}\Bigg\{\iota\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s\mathbf{h}_e)\ln 2}\right] -$$

$$\left[\frac{\frac{1}{\sigma_{er}^2}\mathbf{h}_e\mathbf{h}_e^H}{(1+\frac{1}{\sigma_{er}^2}\mathbf{h}_e^H\mathbf{Q}_s^k\mathbf{h}_e)\ln 2}\right] - \kappa\zeta_{eh}\mathbf{h}_e\mathbf{h}_e^H - \varpi\mathbf{h}_p\mathbf{h}_p^H + (\lambda+\mu)\mathbf{I}\Bigg\}^{-1}\mathbf{Q}_s\Bigg\}$$

$$\leq \text{rank}\left[\frac{\frac{1}{\sigma_{su}^2}\mathbf{h}_s\mathbf{h}_s^H}{(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s)\ln 2}\right] \leq 1. \tag{5.70}$$

By eliminating the trivial solution $\mathbf{Q}_s = \mathbf{0}$, one can obtain $\text{rank}(\mathbf{Q}_s) = 1$, which completes the proof of Proposition 5.1. ∎

5.8.3   Proof of Proposition 5.2

The proof of Proposition 5.2 is similar to that of Proposition 5.1 provided in the previous chapter. First, the following Lagrangian function of the problem defined in (5.34) is considered:

$$\mathcal{L}(\mathbf{Q}_s, \mathbf{Z}, \mu, \iota, \kappa, \varpi) = -\{\log_2(1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s) - f_s(\mathbf{Q}_s, \mathbf{Q}_s^k) - \lambda[\text{tr}(\mathbf{Q}_s)+P_c]\} - \text{tr}(\mathbf{Z}\mathbf{Q}_s)$$

$$+\mu(\text{tr}(\mathbf{Q}_s)-P_{\text{tx}}) - \iota[1+\frac{1}{\sigma_{su}^2}\mathbf{h}_s^H\mathbf{Q}_s\mathbf{h}_s - 2^{R_d} + \frac{2^{R_d}}{\sigma_{er}^2}\text{tr}(\mathbf{G}_e\mathbf{Q}_s)\ln\beta] + \kappa[\zeta_{eh}(\text{tr}(\mathbf{G}_e\mathbf{Q}_s)\ln(1-\gamma)$$

$$- \sigma_{er}^2) + \omega_s] + \varpi(\mathbf{h}_p^H\mathbf{Q}_s\mathbf{h}_p - P_f), \tag{5.71}$$

where $\mathbf{Q}_s \in \mathbb{H}_+^{N_t}$, $\mathbf{Z} \in \mathbb{H}_+^{N_t}$, $\mu \in \mathbb{R}_+$, $\iota \in \mathbb{R}_+$, $\kappa \in \mathbb{R}_+$ and $\varpi \in \mathbb{R}_+$ are the Lagrangian multipliers associated with problem (5.34). Then, the corresponding KKT conditions [152] are derived as follows:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{Q}_s} = - \left[\frac{\frac{1}{\sigma_{su}^2} \mathbf{h}_s \mathbf{h}_s^H}{(1 + \frac{1}{\sigma_{su}^2} \mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s) \ln 2} \right] + (\lambda + \mu)\mathbf{I} - \varpi \mathbf{h}_p \mathbf{h}_p^H + \left[\frac{-\frac{1}{\sigma_{er}^2} \mathbf{G}_e \ln \alpha}{(1 - \frac{1}{\sigma_{er}^2} \mathrm{tr}(\mathbf{G}_e \mathbf{Q}_s^k) \ln \alpha) \ln 2} \right]$$

$$+ \kappa \zeta_{eh} \mathbf{h}_e \mathbf{h}_e^H - \iota \left[\frac{1}{\sigma_{su}^2} \mathbf{h}_s \mathbf{h}_s^H + \frac{2^{R_d}}{\sigma_{er}^2} \mathrm{tr}(\mathbf{G}_e) \ln \beta \right] - \mathbf{Z} = \mathbf{0}$$

$$\mathbf{Z}\mathbf{Q}_s = \mathbf{0}, \ \mathbf{Z} \succeq \mathbf{0}. \tag{5.72}$$

Following the same set of mathematical manipulations in (5.66)-(5.69), the following rank relation holds:

$$\mathrm{rank}(\mathbf{Q}_s) = \mathrm{rank}\left\{ \left\{ \left[\frac{\frac{1}{\sigma_{su}^2} \mathbf{h}_s \mathbf{h}_s^H}{(1 + \frac{1}{\sigma_{su}^2} \mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s) \ln 2} \right] \right\} \left\{ (\lambda + \mu)\mathbf{I} + \left[\frac{-\frac{1}{\sigma_{er}^2} \mathbf{G}_e \ln \alpha}{(1 - \frac{1}{\sigma_{er}^2} \mathrm{tr}(\mathbf{G}_e \mathbf{Q}_s^k) \ln \alpha) \ln 2} \right] \right. \right.$$

$$\left. \left. + \kappa \zeta_{eh} \mathbf{h}_e \mathbf{h}_e^H - \varpi \mathbf{h}_p \mathbf{h}_p^H - \iota \left[\frac{1}{\sigma_{su}^2} \mathbf{h}_s \mathbf{h}_s^H + \frac{2^{R_d}}{\sigma_{er}^2} \mathrm{tr}(\mathbf{G}_e) \ln \beta \right] \right\}^{-1} \mathbf{Q}_s \right\}$$

$$\leq \mathrm{rank}\left[\frac{\frac{1}{\sigma_{su}^2} \mathbf{h}_s \mathbf{h}_s^H}{(1 + \frac{1}{\sigma_{su}^2} \mathbf{h}_s^H \mathbf{Q}_s \mathbf{h}_s) \ln 2} \right] \leq 1. \tag{5.73}$$

By eliminating the trivial solution $\mathbf{Q}_s = \mathbf{0}$, one can obtain $\mathrm{rank}(\mathbf{Q}_s) = 1$, which completes the proof of Proposition 5.2. ∎

5.8.4 Proof of Proposition 5.3

First, suppose that the problem defined in (5.46) is solved with the optimal value Λ^*. Now, the following power minimization problem is considered:

$$\min_{\mathbf{Q}_s, \tau} \mathrm{tr}(\mathbf{Q}_s)$$

$$s.t. \ 1 + \frac{1}{\sigma_{su}^2} (\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s)^H \mathbf{Q}_s (\hat{\mathbf{h}}_s + \hat{\mathbf{e}}_s) \geq \tau \Lambda^*,$$

$$(5.43a), (5.43b), (5.39c), (5.39d),$$

$$\lambda (\mathrm{tr}(\mathbf{Q}_s) + P_c) \leq \nu, \ \mathrm{tr}(\mathbf{Q}_s) \leq P_{\mathrm{tx}}, \ \mathbf{Q}_s \succeq \mathbf{0}. \tag{5.74}$$

By applying the S-Procedure [47], the above problem can be transformed into the following SDP problem:

$$\min_{\mathbf{Q}_s,\tau} \mathrm{tr}(\mathbf{Q}_s)$$

$$s.t.\ \mathbf{T}_1(\mathbf{Q}_s,\chi_1) \succeq \mathbf{0}, \mathbf{T}_2(\mathbf{Q}_s,\chi_2) \succeq \mathbf{0}, \mathbf{T}_3(\mathbf{Q}_s,\chi_3) \succeq \mathbf{0},$$

$$\mathbf{T}_4(\mathbf{Q}_s,\chi_4) \succeq \mathbf{0}, \mathbf{T}_5(\mathbf{Q}_s,\chi_5) \succeq \mathbf{0},$$

$$\lambda(\mathrm{tr}(\mathbf{Q}_s) + P_c) \leq \nu, \mathrm{tr}(\mathbf{Q}_s) \leq P_{\mathrm{tx}}, \mathbf{Q}_s \succeq \mathbf{0}, \tag{5.75}$$

where

$$\mathbf{T}_1(\mathbf{Q}_s,\chi_1) = \begin{bmatrix} \chi_1\mathbf{I} + \frac{1}{\sigma_{su}^2}\mathbf{Q}_s & \frac{1}{\sigma_{su}^2}\mathbf{Q}_s\hat{\mathbf{h}}_s \\ \frac{1}{\sigma_{su}^2}\hat{\mathbf{h}}_s^H\mathbf{Q}_s & \frac{1}{\sigma_{su}^2}\hat{\mathbf{h}}_s^H\mathbf{Q}_s\hat{\mathbf{h}}_s + 1 - \tau\Lambda^* - \chi_1\epsilon_s^2 \end{bmatrix}, \tag{5.76}$$

$$\mathbf{T}_2(\mathbf{Q}_s,\chi_2) = \begin{bmatrix} \chi_2\mathbf{I} + \frac{1}{\sigma_{su}^2}\mathbf{Q}_s & \frac{1}{\sigma_{su}^2}\mathbf{Q}_s\hat{\mathbf{h}}_s \\ \frac{1}{\sigma_{su}^2}\hat{\mathbf{h}}_s^H\mathbf{Q}_s & \frac{1}{\sigma_{su}^2}\hat{\mathbf{h}}_s^H\mathbf{Q}_s\hat{\mathbf{h}}_s + 1 - \tau 2^{R_d} - \chi_2\epsilon_s^2 \end{bmatrix}, \tag{5.77}$$

$$\mathbf{T}_3(\mathbf{Q}_s,\chi_3) = \begin{bmatrix} \chi_3\mathbf{I} - \frac{1}{\sigma_{er}^2}\mathbf{Q}_s & -\frac{1}{\sigma_{er}^2}\mathbf{Q}_s\hat{\mathbf{h}}_e \\ -\frac{1}{\sigma_{er}^2}\hat{\mathbf{h}}_e^H\mathbf{Q}_s & \tau - 1 - \frac{1}{\sigma_{er}^2}\hat{\mathbf{h}}_e^H\mathbf{Q}_s\hat{\mathbf{h}}_e - \chi_3\epsilon_e^2 \end{bmatrix}, \tag{5.78}$$

$$\mathbf{T}_4(\mathbf{Q}_s,\chi_4) = \begin{bmatrix} \chi_4\mathbf{I} + \zeta_{eh}\mathbf{Q}_s & \zeta_{eh}\mathbf{Q}_s\hat{\mathbf{h}}_e \\ \zeta_{eh}\hat{\mathbf{h}}_e^H\mathbf{Q}_s & \zeta_{eh}\hat{\mathbf{h}}_e^H\mathbf{Q}_s\hat{\mathbf{h}}_e + \zeta_{eh}\sigma_{er}^2 - \omega_s - \chi_4\epsilon_e^2 \end{bmatrix}, \tag{5.79}$$

$$\mathbf{T}_5(\mathbf{Q}_s,\chi_5) = \begin{bmatrix} \chi_5\mathbf{I} - \mathbf{Q}_s & -\mathbf{Q}_s\hat{\mathbf{h}}_p \\ -\hat{\mathbf{h}}_p^H\mathbf{Q}_s & P_f - \hat{\mathbf{h}}_p^H\mathbf{Q}_s\hat{\mathbf{h}}_p - \chi_5\epsilon_p^2 \end{bmatrix}. \tag{5.80}$$

Next, the Lagrangian function of the problem defined in (5.75) can be written as

$$\mathcal{L}(\mathbf{X}) = \mathrm{tr}(\mathbf{Q}_s) - \mathrm{tr}(\mathbf{Q}_s\mathbf{Z}) - \sum_{i=1}^{5}\mathrm{tr}(\mathbf{T}_i(\mathbf{Q}_s,\chi_i)\mathbf{A}_i) - \sum_{i=1}^{5}\chi_i m_i, \tag{5.81}$$

where $\mathbf{X}$ denotes the set of all primal and dual variables: $\mathbf{X} = \{\mathbf{Q}, \mathbf{Z}, \tau, \mathbf{A}, \chi, \mathbf{m}\}$ with $\mathbf{A} = \{\mathbf{A}_i\}_{i=1}^{5}, \chi = [\chi_1, ..., \chi_5]^T$ and $\mathbf{m} = [m_1, ..., m_5]^T$; $\mathbf{A}_i \in \mathbb{H}_+^{N_t+1}$, $\mathbf{Z} \in \mathbb{H}_+^{N_t}$ and $m_i \in \mathbb{R}_+$ are the Lagrangian multipliers associated with the problem in (5.75). For simple expressions, $\mathbf{T}_i(\mathbf{Q}_s,\chi_i)$ is rewritten as

$$\mathbf{T}_1(\mathbf{Q}_s,\chi_1) = \mathbf{S}_1(\mathbf{Q}_s,\chi_1) + \frac{1}{\sigma_{su}^2}\mathbf{V}_s^H\mathbf{Q}_s\mathbf{V}_s, \tag{5.82}$$

$$\mathbf{T}_2(\mathbf{Q}_s,\chi_2) = \mathbf{S}_2(\mathbf{Q}_s,\chi_2) + \frac{1}{\sigma_{su}^2}\mathbf{V}_s^H\mathbf{Q}_s\mathbf{V}_s, \tag{5.83}$$

$$\mathbf{T}_3(\mathbf{Q}_s,\chi_3) = \mathbf{S}_3(\mathbf{Q}_s,\chi_3) - \frac{1}{\sigma_{er}^2}\mathbf{V}_e^H\mathbf{Q}_s\mathbf{V}_e, \tag{5.84}$$

$$\mathbf{T}_4(\mathbf{Q}_s, \chi_4) = \mathbf{S}_4(\mathbf{Q}_s, \chi_4) + \zeta_{eh}\mathbf{V}_e^H\mathbf{Q}_s\mathbf{V}_e, \tag{5.85}$$

$$\mathbf{T}_5(\mathbf{Q}_s, \chi_5) = \mathbf{S}_5(\mathbf{Q}_s, \chi_5) - \mathbf{V}_p^H\mathbf{Q}_s\mathbf{V}_p, \tag{5.86}$$

where

$$\mathbf{S}_1(\mathbf{Q}_s, \chi_1) = \begin{bmatrix} \chi_1\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 1\tau\Lambda^* - \chi_1\epsilon_s^2 \end{bmatrix},$$

$$\mathbf{S}_2(\mathbf{Q}_s, \chi_2) = \begin{bmatrix} \chi_2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 1 - \tau 2^{R_d} - \chi_2\epsilon_s^2 \end{bmatrix},$$

$$\mathbf{S}_3(\mathbf{Q}_s, \chi_3) = \begin{bmatrix} \chi_3\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tau - 1 - \chi_3\epsilon_e^2 \end{bmatrix},$$

$$\mathbf{S}_4(\mathbf{Q}_s, \chi_4) = \begin{bmatrix} \chi_4\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \zeta_{eh}\sigma_{er}^2 - \omega_s - \chi_4\epsilon_e^2 \end{bmatrix},$$

$$\mathbf{S}_5(\mathbf{Q}_s, \chi_5) = \begin{bmatrix} \chi_5\mathbf{I} & \mathbf{0} \\ \mathbf{0} & P_f - \chi_5\epsilon_p^2 \end{bmatrix},$$

$$\mathbf{V}_s = [\mathbf{I}_{N_T} \; \hat{\mathbf{h}}_s], \mathbf{V}_e = [\mathbf{I}_{N_T} \; \hat{\mathbf{h}}_e], \mathbf{V}_p = [\mathbf{I}_{N_T} \; \hat{\mathbf{h}}_p]. \tag{5.87}$$

Substituting (5.82)-(5.86) into (5.81), a different expression for the Lagrangian function is obtained as follows:

$$\mathcal{L}(\mathbf{X}) = \text{tr}(\mathbf{Q}_s) - \text{tr}(\mathbf{Q}_s\mathbf{Z}) - \text{tr}(\frac{1}{\sigma_{su}^2}\mathbf{Q}_s\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H) - \text{tr}(\frac{1}{\sigma_{su}^2}\mathbf{Q}_s\mathbf{V}_s\mathbf{A}_2\mathbf{V}_s^H)$$
$$+ \text{tr}(\frac{1}{\sigma_{er}^2}\mathbf{Q}_s\mathbf{V}_e\mathbf{A}_3\mathbf{V}_e^H) - \text{tr}(\zeta_{eh}\mathbf{Q}_s\mathbf{V}_e\mathbf{A}_4\mathbf{V}_e^H) + \text{tr}(\mathbf{Q}_s\mathbf{V}_p\mathbf{A}_5\mathbf{V}_p^H) + \varphi(\tau, \mathbf{A}, \chi, \mathbf{m}), \tag{5.88}$$

where $\varphi(\tau, \mathbf{A}, \chi, \mathbf{m})$ collects all the terms that are not related to $\mathbf{Q}_s$ and $\mathbf{Z}$, which are independent of this proof. Then, the KKT conditions [152] are defined as:

$$\frac{\partial\mathcal{L}}{\partial\mathbf{Q}_s} = \mathbf{I} - \mathbf{Z} - \frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H - \frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_2\mathbf{V}_s^H + \frac{1}{\sigma_{er}^2}\mathbf{V}_e\mathbf{A}_3\mathbf{V}_e^H - \zeta_{eh}\mathbf{V}_e\mathbf{A}_4\mathbf{V}_e^H + \mathbf{V}_p\mathbf{A}_5\mathbf{V}_p^H$$
$$= \mathbf{0}, \tag{5.89}$$

$$\mathbf{T}_1(\mathbf{Q}_s, \chi_1)\mathbf{A}_1 = \mathbf{0}, \tag{5.90}$$

$$\mathbf{Q}_s\mathbf{Z} = \mathbf{0}, \mathbf{Z} \succeq \mathbf{0}, \mathbf{A}_i \succeq \mathbf{0}, i \in [1, 2, 3, 4, 5]. \tag{5.91}$$

Premultiplying the two sides of (5.89) and using the fact that $\mathbf{Q}_s\mathbf{Z} = \mathbf{0}$, one can obtain

$$\mathbf{Q}_s\frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H = \mathbf{Q}_s\left(\mathbf{I} - \frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_2\mathbf{V}_s^H + \frac{1}{\sigma_{er}^2}\mathbf{V}_e\mathbf{A}_3\mathbf{V}_e^H - \zeta_{eh}\mathbf{V}_e\mathbf{A}_4\mathbf{V}_e^H + \mathbf{V}_p\mathbf{A}_5\mathbf{V}_p^H\right). \tag{5.92}$$

The following rank relationship holds

$$\text{rank}(\mathbf{Q}_s) = \text{rank}\left(\mathbf{Q}_s\left(\mathbf{I} - \frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_2\mathbf{V}_s^H + \frac{1}{\sigma_{er}^2}\mathbf{V}_e\mathbf{A}_3\mathbf{V}_e^H - \zeta_{eh}\mathbf{V}_e\mathbf{A}_4\mathbf{V}_e^H + \mathbf{V}_p\mathbf{A}_5\mathbf{V}_p^H\right)\right)$$

$$= \text{rank}(\mathbf{Q}_s\frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H) \leq \min\ \{\text{rank}(\frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H), \text{rank}(\mathbf{Q}_s)\}. \tag{5.93}$$

If it can be proven that $\text{rank}(\frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H) = 1$, then it is obvious that $\text{rank}(\mathbf{Q}_s) \leq 1$. Hence, in the remaining part of this proof, the rank of $\frac{1}{\sigma_{su}^2}\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H$ is focused. By substituting (5.82) into (5.90), the following equation can be obtained

$$\mathbf{S}_1(\mathbf{Q}_s, \chi_1)\mathbf{A}_1 + \frac{1}{\sigma_{su}^2}\mathbf{V}_s^H\mathbf{Q}_s\mathbf{V}_s\mathbf{A}_1 = \mathbf{0}. \tag{5.94}$$

Furthermore, it follows by postmultiplying by $\mathbf{V}_s^H$ that

$$\mathbf{S}_1(\mathbf{Q}_s, \chi_1)\mathbf{A}_1\mathbf{V}_s^H + \frac{1}{\sigma_{su}^2}\mathbf{V}_s^H\mathbf{Q}_s\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H = \mathbf{0}. \tag{5.95}$$

One can easily verify that

$$[\mathbf{I}_{N_T}\ \mathbf{0}]\mathbf{S}_1(\mathbf{Q}_s, \chi_1) = \chi_1[\mathbf{I}_{N_T}\ \mathbf{0}] = \chi_1(\mathbf{V}_s - [\mathbf{0}_{N_T}\ \hat{\mathbf{h}}_{\mathbf{s}}]),\ [\mathbf{I}_{N_T}\ \mathbf{0}]\mathbf{V}_s^H = \mathbf{I}_{N_T}. \tag{5.96}$$

By premulitiplying both sides of (5.95) by $[\mathbf{I}_{N_T}\ \mathbf{0}]$, the following relation can be obtained

$$\chi_1(\mathbf{V}_s - [\mathbf{0}_{N_T}\ \hat{\mathbf{h}}_{\mathbf{s}}])\mathbf{A}_1\mathbf{V}_s^H + \frac{1}{\sigma_{su}^2}\mathbf{Q}_s\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H = \mathbf{0}$$

$$\Leftrightarrow (\chi_1\mathbf{I}_{N_T} + \frac{1}{\sigma_{su}^2}\mathbf{Q}_s)\mathbf{V}_s\mathbf{A}_1\mathbf{V}_s^H = \chi_1[\mathbf{0}_{N_T}\ \hat{\mathbf{h}}_{\mathbf{s}}]\mathbf{A}_1\mathbf{V}_s^H. \tag{5.97}$$

Lemma 5.4: [200]): If a block Hermitian matrix

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 \\ \mathbf{P}_3 & \mathbf{P}_4 \end{bmatrix} \succeq \mathbf{0},$$

then the main diagonal matrices $\mathbf{P}_1$ and $\mathbf{P}_4$ are always positive semidefinite matrices.

Based on Lemma 5.4 and $\mathbf{T}_1(\mathbf{Q}_s, \chi_1) \succeq \mathbf{0}$, it is obvious that $\chi_1 \mathbf{I}_{N_T} + \frac{1}{\sigma_{su}^2} \mathbf{Q}_s$ is a positive definite matrix. Since multiplying both sides by a non-singular matrix does not change the rank of the original matrix, the following rank relationship holds:

$$\text{rank}(\mathbf{V}_s \mathbf{A}_1 \mathbf{V}_s^H) = \text{rank}\left((\chi_1 \mathbf{I}_{N_T} + \frac{1}{\sigma_{su}^2} \mathbf{Q}_s) \mathbf{V}_s \mathbf{A}_1 \mathbf{V}_s^H \right) = \text{rank}(\chi_1 [\mathbf{0}_{N_T} \ \hat{\mathbf{h}}_{\mathbf{s}}] \mathbf{A}_1 \mathbf{V}_s^H)$$
$$= \text{rank}([\mathbf{0}_{N_T} \ \hat{\mathbf{h}}_{\mathbf{s}}]) \leq 1. \tag{5.98}$$

Combining (5.93) and (5.98), the following inequality can be derived

$$\text{rank}(\mathbf{Q}_s) \leq \text{rank}(\mathbf{V}_s \mathbf{A}_1 \mathbf{V}_s^H) \leq 1. \tag{5.99}$$

Similar to the previous proofs, by eliminating the trivial solution $\mathbf{Q}_s = \mathbf{0}$, $\text{rank}(\mathbf{Q}_s) = 1$ can be obtained, which completes the proof of Proposition 5.3. ∎

Chapter 6

Exploiting Machine Learning for Secure Transmission in an Underlay Cognitive Radio Network

6.1   Introduction

In secure transmission designs, different optimization approaches with various approximations techniques have been widely exploited to solve a numerous resource allocation problems [83, 33, 55, 49, 201]. However, these techniques often have been developed based on iterative approaches to yield either the optimal or the sub-optimal solutions. The computational complexities associated with these conventional optimization techniques are not affordable in low powered devices in IoT or not suitable for applications with the requirements of ultra reliability and low latency in future wireless networks. Furthermore, these conventional optimization techniques pose different challenges in delay sensitive systems as the dynamic nature of real-time parameters requires frequent updates in very short time [137]. This introduces different stringent delay requirements in updating those design parameters which is impossible to meet by conventional optimization approaches.

In this chapter, a secure transmission in an underlay CR network is considered. This secure network consists of one PU-Tx, one PU-Rx, one SU-Tx, one SU-Rx and one EVE. All of these terminals are equipped with a single antenna. The aim is to enhance the quality of secure communications between SU-Tx and SU-Rx in the presence of an EVE. In particular, the optimal power allocation is determined to maximize the achievable secrecy rate under the constraints of total transmit power of the SU-Tx and the interference leakage to the PU-Rx.
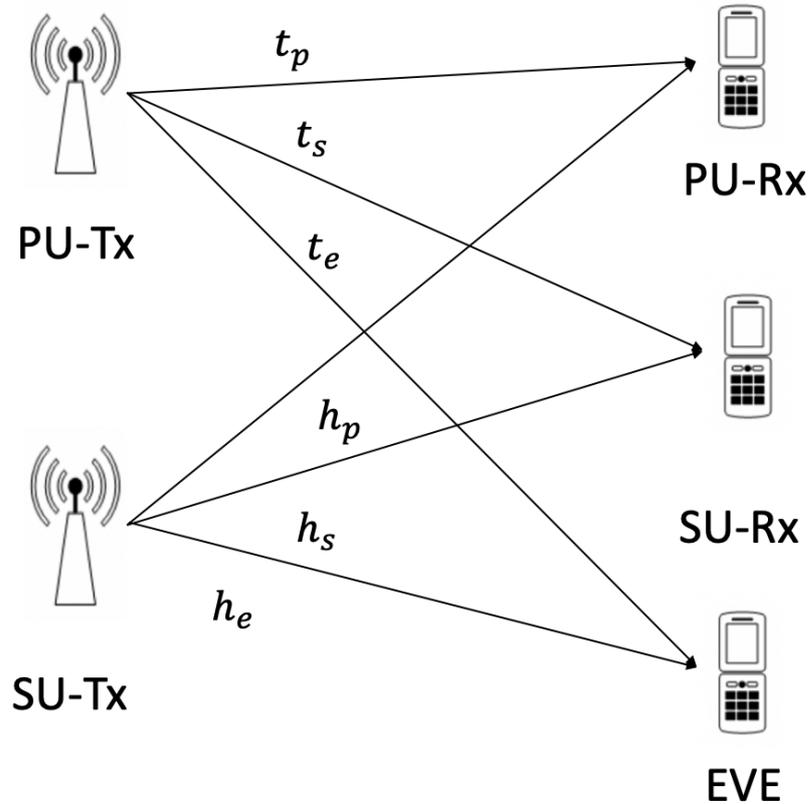
Fig. 6.1 A CR network with one PU-Tx, one PU-Rx, one SU-Tx, one SU-Rx and one EVE. All the terminals are equipped with a single antenna.

Then, a NN based framework is proposed to learn the relationship between the input and the output parameters of this secure transmission system. It is showed that the proposed NN can be utilized to solve different types of resource allocation problems. For example, the perfect and imperfect CSI scenarios require different conventional convex optimization frameworks to obtain the optimal solutions, however, the same NN model has the capability to handle both robust and non-robust designs.

6.2   System Model

A CR network is considered with five terminals as shown in Fig. 6.1 : one PU-Tx, one SU-Tx, one SU-Rx, one PU-Rx and one EVE. It is assumed that all of these terminals are equipped with single antenna. The SU-Tx intends to send a confidential message to the SU-Rx while ensuring the interference leakage to the PU-Rx is less than a predefined threshold. At the same time the EVE attempts intercepting the information sent by the SU-Tx

to the SU-Rx. The channels between the PU-Tx and the PU-Rx, the SU-Rx as well as the EVE are represented by t_p, t_s and t_e, respectively, whereas the channels between the SU-Tx and the PU-Rx, the SU-Rx, and the EVE are denoted by h_p, h_s and h_e, respectively. The received signal at the SU-Rx and the EVE can be expressed, respectively, as

$$y_s = \sqrt{P_s}h_s x_s + \sqrt{P_p}t_s x_p + n_s, \tag{6.1}$$

$$y_e = \sqrt{P_s}h_e x_s + \sqrt{P_p}t_e x_p + n_e, \tag{6.2}$$

where $x_s(\mathbb{E}\{|x_s|^2\} = 1)$ and $x_p(\mathbb{E}\{|x_p|^2\} = 1)$ are the symbols that sent from the SU-Tx to the SU-Rx and PU-Tx to the PU-Rx, respectively. The noise at the SU-Rx and the EVE are denoted by $n_s(\mathbb{E}\{|n_s|^2\} = \sigma_s^2)$ and $n_e(\mathbb{E}\{|n_e|^2\} = \sigma_e^2)$, respectively. Furthermore, P_s and P_p denote the transmit power of the SU-Tx and the PU-Tx, respectively. Based on these system parameters, the SINR at the SU-Rx and the EVE are defined as

$$\gamma_s = \frac{P_s|h_s|^2}{P_p|t_s|^2 + \sigma_s^2}, \tag{6.3}$$

$$\gamma_e = \frac{P_s|h_e|^2}{P_p|t_e|^2 + \sigma_e^2}. \tag{6.4}$$

The achievable secrecy rate at the SU-Rx can be written as [11]

$$R_s = [\log_2(1 + \gamma_s) - \log_2(1 + \gamma_e)]^+. \tag{6.5}$$

The interference leakage to the PU-Rx can be expressed as

$$P_{in} = P_s|h_p|^2. \tag{6.6}$$

With these definitions, the secrecy rate maximization problem can be formulated as

$$\begin{aligned}
\max_{P_s} \quad & R_s \\
\text{s.t.} \quad & P_s|h_p|^2 \leq q, \\
& P_s \leq P_t, P_s \geq 0,
\end{aligned} \tag{6.7}$$

where q is the maximum interference leakage to the PU-Rx, and P_t is the maximum transmit power available at the SU-Tx. In the following sections, two approaches are presented to solve this problem: Conventional optimization and NN based approach.

6.3   Conventional Optimization based Design

In this section, conventional convex optimization approaches are presented to solve the secrecy rate maximization problem defined in (6.7) by taking into account the scenarios of having both perfect and imperfect CSI at the SU-Tx.

6.3.1   Design with Perfect CSI

In this subsection, the conventional convex optimization based approach is provided to solve the problem defined in (6.7) with perfect CSI assumption. The original problem (6.7) is not convex in its original form due to the non-convex objective function. Based on the monotonicity of logarithmic functions, the original problem in (6.7) can be reformulated as

$$
\max_{P_s} \quad \frac{1 + \frac{P_s |h_s|^2}{P_p |t_s|^2 + \sigma_s^2}}{1 + \frac{P_s |h_e|^2}{P_p |t_e|^2 + \sigma_e^2}}
$$
$$
s.t. \; P_s |h_p|^2 \leq q,
$$
$$
P_s \leq P_t, P_s \geq 0. \tag{6.8}
$$

The above problem still remains as non-convex due to the fractional objective function and therefore it cannot be directly solved using existing convex optimization tools. To circumvent this non-convexity issue, the original problem is converted into a two-level optimization problem, namely, outer problem and inner problem. The outer problem can be written w.r.t. a new scalar variable t as

$$
\max_{t \geq 0} \; \frac{f(t)}{1 + t}, \tag{6.9}
$$

whereas the inner problem can be expressed for a given t as

$$
f(t) = \max_{P_s} \; 1 + \frac{P_s |h_s|^2}{P_p |t_s|^2 + \sigma_s^2}
$$
$$
s.t. \; P_s |h_p|^2 \leq q,
$$
$$
\frac{P_s |h_e|^2}{P_p |t_e|^2 + \sigma_e^2} \leq t,
$$
$$
P_s \leq P_t, \; P_s \geq 0. \tag{6.10}
$$

The outer problem in (6.9) is convex for a given t and can be solved by using standard interior-point methods. Since the inner problem in (6.10) is a convex problem, the outer problem in (6.9) is a quasi-convex optimization problem w.r.t. variable t. Therefore, an one dimensional search is employed to obtain the optimal t^* [16]. The proposed one dimensional search algorithm is summarized in Algorithm 4.

Algorithm 4 One dimensional search based on bisection method

1: Initialize $t \in [0, t_{max}]$, $c = (\sqrt{5} - 1)/2$, $a = 0$, $b = t_{max}$, $t_1 = (1 - c)b$, $t_2 = cb$;
2: Compute $f(t_1)$, $f(t_2)$;
3: **repeat**
4:     If $\frac{1+f(t_1)}{1+t_1} > \frac{1+f(t_2)}{1+t_2}$, $b = t_2$, $t_2 = t_1$, $f(t_2) = f(t_1)$, $t_1 = b - c(b - a)$ and update $f(t_1)$;
5:     Else, $a = t_1$, $t_1 = t_2$, $f(t_1) = f(t_2)$, $t_2 = a + c(b - a)$, and update $f(t_2)$;
6: **until** $|b - a| \leq \epsilon$, where ϵ is a predefined threshold to terminate the algorithm.

6.3.2 Design with Imperfect CSI

In this subsection, a tractable approach is developed to solve the secrecy tare maximization problem with imperfect CSI. This robust problem is reformulated into a tractable one by exploiting Charnes-Cooper transformation [194] and S-Procedure [47]. The actual channel coefficients can be modelled with corresponding channel uncertainties as follows:

$$h_s = \hat{h}_s + e_s, \ h_e = \hat{h}_e + e_e, \ h_p = \hat{h}_p + e_p, \tag{6.11}$$

where $\hat{h}_s$, $\hat{h}_e$ and $\hat{h}_p$ are the channel coefficients estimated by the SU-Tx. Furthermore, the symbols e_s, e_e and e_p represent the channel uncertainties. These channel uncertainties are assumed to be bounded by a predefined ellipsoids as follows:

$$|e_s| = |h_s - \hat{h}_s| \leq \epsilon_s, \tag{6.12}$$

$$|e_e| = |h_e - \hat{h}_e| \leq \epsilon_e, \tag{6.13}$$

$$|e_p| = |h_p - \hat{h}_p| \leq \epsilon_p, \tag{6.14}$$

where $\epsilon_s \geq 0$, $\epsilon_e \geq 0$ and $\epsilon_p \geq 0$ are the error bounds. Based on these bounded channel uncertainties and the monotonicity of log functions, the robust secrecy rate maximization

problem can be reformulated into the following robust optimization framework:

$$\max_{P_s} \quad \frac{1 + \frac{P_s|\hat{h}_s + e_s|^2}{P_p|t_s|^2 + \sigma_s^2}}{1 + \frac{P_s|\hat{h}_e + e_e|^2}{P_p|t_e|^2 + \sigma_e^2}}$$

$$s.t. \quad P_s|\hat{h}_p + e_p|^2 \le q,$$

$$P_s \le P_t, \ P_s \ge 0. \tag{6.15}$$

First, the Charnes-Cooper transformation is introduced [194] as

$$\overline{P}_s = \frac{P_s}{t}, \tag{6.16}$$

to recast the problem defined in (6.15) as

$$\max_{\overline{P}_s, \, t} \quad t + \frac{\overline{P}_s|\hat{h}_s + e_s|^2}{P_p|t_s|^2 + \sigma_s^2}$$

$$s.t. \quad t + \frac{\overline{P}_s|\hat{h}_e + e_e|^2}{P_p|t_e|^2 + \sigma_e^2} \le 1,$$

$$\overline{P}_s|\hat{h}_p + e_p|^2 \le tq,$$

$$\overline{P}_s \le tP_t, \ \overline{P}_s \ge 0. \tag{6.17}$$

The problem defined in (6.17) can be rewritten by introducing a new slack variable τ and defining it in the epigraph form as

$$\max_{\overline{P}_s, \, t, \, \tau} \quad \tau \tag{6.18a}$$

$$s.t. \quad t + \frac{\overline{P}_s|\hat{h}_s + e_s|^2}{P_p|t_s|^2 + \sigma_s^2} \ge \tau, \tag{6.18b}$$

$$t + \frac{\overline{P}_s|\hat{h}_e + e_e|^2}{P_p|t_e|^2 + \sigma_e^2} \le 1, \tag{6.18c}$$

$$\overline{P}_s|\hat{h}_p + e_p|^2 \le tq, \tag{6.18d}$$

$$\overline{P}_s \le tP_t, \ \overline{P}_s \ge 0. \tag{6.18e}$$

The above problem is still intractable due to the infinite number of the uncertainty sets in the constraints (6.18b)-(6.18d). Hence, (6.12) and (6.18b) are rewritten as

$$|e_s|^2 - \epsilon_s^2 \le 0, \tag{6.19}$$

$$\tau - t - \frac{\overline{P}_s |\hat{h}_s + e_s|^2}{P_p |t_s|^2 + \sigma_s^2} \le 0. \tag{6.20}$$

By considering S-Procedure [47], the constraint (6.18b) can be reformulated as the following constraints:

$$\begin{bmatrix} \lambda_1 + \frac{\overline{P}_s}{P_p |t_s|^2 + \sigma_s^2} & \frac{\overline{P}_s \hat{h}_s}{P_p |t_s|^2 + \sigma_s^2} \\ \frac{\overline{P}_s \hat{h}_s}{P_p |t_s|^2 + \sigma_s^2} & \frac{\overline{P}_s |\hat{h}_s|^2}{P_p |t_s|^2 + \sigma_s^2} + t - \tau - \lambda_1 \epsilon_s^2 \end{bmatrix} \succeq \mathbf{0},$$
$$\lambda_1 \ge 0. \tag{6.21}$$

Similarly, the constraints (6.18c) and (6.18d) can be recast respectively as

$$\begin{bmatrix} \lambda_2 - \frac{\overline{P}_s}{P_p |t_e|^2 + \sigma_s^2} & -\frac{\overline{P}_s \hat{h}_e}{P_p |t_e|^2 + \sigma_s^2} \\ -\frac{\overline{P}_s \hat{h}_e}{P_p |t_e|^2 + \sigma_s^2} & 1 - \frac{\overline{P}_s |\hat{h}_s|^2}{P_p |t_e|^2 + \sigma_e^2} - t - \lambda_2 \epsilon_e^2 \end{bmatrix} \succeq \mathbf{0},$$
$$\lambda_2 \ge 0, \tag{6.22}$$

and

$$\begin{bmatrix} \lambda_3 - \overline{P}_s & -\overline{P}_s \hat{h}_p \\ -\overline{P}_s \hat{h}_p & tq - \overline{P}_s |\hat{h}_p|^2 + \sigma_s^2 - \lambda_3 \epsilon_p^2 \end{bmatrix} \succeq \mathbf{0},$$
$$\lambda_3 \ge 0. \tag{6.23}$$

Therefore, the problem in (6.18) can be rewritten into the following equivalent form:

$$\begin{aligned} \max_{\overline{P}_s, \, t, \, \tau} \quad & \tau \\ s.t. \quad & (6.21)\text{-}(6.23), \\ & \overline{P}_s \le t P_t, \; \overline{P}_s \ge 0. \end{aligned} \tag{6.24}$$

The above problem is convex and therefore it can be solved efficiently by convex optimization tool box [152].

6.4   Power Allocation Framework based on NN

In this section, a NN based power allocation scheme is presented. In this approach, the secrecy rate maximization problem is treated as an unknown non-linear mapping, and a NN is trained to learn the relationship between the input and the output parameters. The motivation
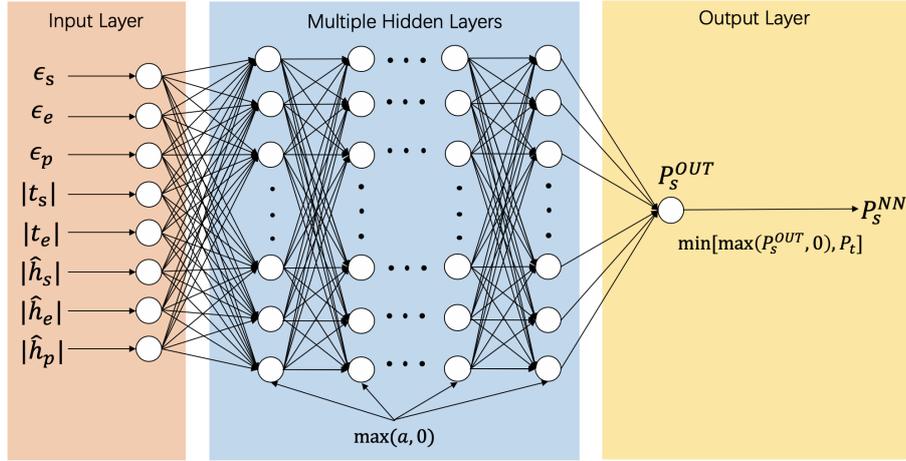
Fig. 6.2 The structure of the proposed NN.

is to utilize the high computational efficiency of the NN in its testing stage to design a time and computation efficient real-time power allocation scheme which can be applied to solve the power allocation problem with both perfect and imperfect CSI. As shown in Fig. 6.2, the proposed NN consists of three layers: the input layer, multiple hidden layers and the output layer. In particular, the $|\hat{h}_s|$, $|\hat{h}_p|$, $|\hat{h}_e|$, $|t_s|$, $|t_e|$, ϵ_s, ϵ_e and ϵ_p as the inputs and P_s^* are chosen as the output of the training data of the proposed NN, respectively. Note that the perfect CSI scheme becomes as a special case of imperfect CSI scheme by setting the inputs for perfect CSI scheme as to be $|\hat{h}_s| = |h_s|$, $|\hat{h}_p| = |h_p|$, $|\hat{h}_e| = |h_e|$, and $\epsilon_s = \epsilon_e = \epsilon_p = 0$. The mapping between the input and the output parameters can be expressed as

$$P_s^* = f(|\hat{h}_s|, |\hat{h}_p|, |\hat{h}_e|, |t_s|, |t_e|, \epsilon_s, \epsilon_e, \epsilon_p). \tag{6.25}$$

Starting from the input and then passing them through the NN and calculating the actual output straightforwardly, which is referred as feed-forward. Furthermore, the calculation flow follows in the natural forward direction from the input layer to the hidden layers and finally to the output layer. This process can be expressed mathematically as

$$\mathbf{z}^{(l+1)} = \mathbf{W}^{(l)}\mathbf{a}^{(l)} + \mathbf{b}^{(l)}, \tag{6.26}$$

$$\mathbf{a}^{(l+1)} = g(\mathbf{z}^{(l+1)}), \tag{6.27}$$

where $\mathbf{z}^{(l+1)}$ is the linear transformation of given inputs at the $l+1$-th layer, whereas $\mathbf{a}^{(l+1)}$ is output activation value of the $l+1$-th layer. The $g(\mathbf{z})$ denotes any activation function, in this work, the ReLU function is chosen as the activation function, which can be expressed as

$g(x) = \max\{0, x\}$. The $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ are the weight matrix and the bias vector for the l-th layer, respectively. Suppose there is a N-layer NN, the mapping between the inputs and the output parameters can be expressed as

$$y = f(\mathbf{S}, \mathbf{W}, \mathbf{b}), \tag{6.28}$$

where $\mathbf{S} = [|\hat{h}_s|, |\hat{h}_p|, |\hat{h}_e|, |t_s|, |t_e|, \epsilon_s, \epsilon_e, \epsilon_p]$. The goal is to determine the weights $\mathbf{W}(= [\mathbf{W}^{(1)}, ..., \mathbf{W}^{(N-1)}])$ and the bias $\mathbf{b} = [\mathbf{b}^{(1)}, ..., \mathbf{b}^{(N-1)}])$ such that both functions in (6.25) and (6.28) yield a similar output for the same set of inputs.

Proposition 6.1: In order to have a similar outputs from both (6.25) and (6.28), the following normalized loss function should be minimized:

$$J(\mathbf{W}, \mathbf{b}) = \frac{1}{M} \sum_{m=1}^{M} (y_m - P_{s,m}^*)^2, \tag{6.29}$$

where M is the number of training data set, y_m and $P_{s,m}^*$ are the m-th output of the NN and the optimal transmit power obtained by conventional optimization scheme, respectively.

Proof: Please refer to Section 6.7.1.   ■

The back-propagation based gradient descent algorithm can be iteratively used to update the weights matrices $\mathbf{W}$ and the bias vectors $\mathbf{b}$.

Proposition 6.2: Based on back-propagation and gradient descent algorithm, the weight matrix and the bias vector for the l-th layer $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ can be updated respectively by

$$\mathbf{W}^{(l)} = \mathbf{W}^{(l)} - \frac{\alpha}{M} \sum_{m=1}^{M} [\boldsymbol{\delta}_m^{(l+1)} (\mathbf{a}_m^{(l)})^T], \tag{6.30}$$

$$\mathbf{b}^{(l)} = \mathbf{b}^{(l)} - \frac{\alpha}{M} \sum_{m=1}^{M} \boldsymbol{\delta}_m^{(l+1)}, \tag{6.31}$$

where α is the learning rate and $\boldsymbol{\delta}_m^{(l+1)}$ is defined as $\boldsymbol{\delta}_m^{(l+1)} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}_m^{(l+1)}}$.

Proof: Please refer to Section 6.7.2.   ■

In a NN, over-fitting is the result of a model that is very closely to or precisely aligned with a specific set of data [202], which occurs when the model learns the training data set along with noises [203]. Over-fitting leads the model not to be able to fit additional data

or reliably predict future observations [202]. Regularization is an approach to reduce the well-known over-fitting problem of any types of machine learning model [204, 205]. To overcome this over-fitting problems, the L_1 and L_2 regularizations are most widely utilized techniques in the literature [206, 207].

The regularization term is added to the loss function to reduce the sum of absolute value of the weights in L_1 regularization method, where the loss function can be written as

$$J(\mathbf{W}, \mathbf{b}) = \frac{1}{M} \sum_{m=1}^{M} (y_m - P_{s,m}^*)^2 + \frac{\lambda}{2M} \sum_{l=1}^{N-1} ||\mathbf{W}^{(l)}||_1, \qquad (6.32)$$

where λ is the regularization parameter. Following the similar derivation of Proposition 2, the weights for L_1 regularization can be updated as

$$\mathbf{W}^{(l)} = \mathbf{W}^{(l)} - \frac{\alpha}{M} \sum_{m=1}^{M} [\boldsymbol{\delta}_m^{(l+1)} (\mathbf{a}_m^{(l)})^T] - \frac{\alpha\lambda}{M} sgn(\mathbf{W}^{(l)}), \qquad (6.33)$$

The bias $\mathbf{b}^{(l)}$ can be updated by using the same equation provided in (6.31).

In L_2 regularization method, the sum of squares of the weights are reduced by adding the regularization term to the loss function, which can be mathematically expressed as

$$J(\mathbf{W}, \mathbf{b}) = \frac{1}{M} \sum_{m=1}^{M} (y_m - P_{s,m}^*)^2 + \frac{\lambda}{2M} \sum_{l=1}^{N-1} ||\mathbf{W}^{(l)}||_2^2, \qquad (6.34)$$

where λ is the regularization parameter. Following the similar derivation of Proposition 2, the weights for L_1 regularization can be updated as

$$\mathbf{W}^{(l)} = (1 - \frac{\alpha\lambda}{M})\mathbf{W}^{(l)} - \frac{\alpha}{M} \sum_{m=1}^{M} [\boldsymbol{\delta}_m^{(l+1)} (\mathbf{a}_m^{(l)})^T], \qquad (6.35)$$

The bias $\mathbf{b}^{(l)}$ for L_2 regularization can be updated by using the equation provided in (6.31).

The development of the proposed NN scheme can be divided into three steps: (1) Obtaining the training data set by solving the secrecy rate maximization problem through an one dimensional search based conventional optimization approach; (2) developing a NN based algorithm to learn the relationship between the input and output parameters of this secure transmission system; (3) after completing the training process, evaluating the performance of

the trained NN over the conventional optimization algorithm. The details of these steps are provided in Algorithm 5.

Algorithm 5 The NN approach

Preparing process:

1: Obtain the training data set by utilizing the conventional approaches in Section 6.3: The optimal transmit power P_s^* for corresponding channel coefficients $|\hat{h}_s|, |\hat{h}_p|, |\hat{h}_e|, |t_s|, |t_e|$ and channel error bounds $\epsilon_s, \epsilon_e, \epsilon_p$;

Training process:

1: Initialize the weights matrices $\mathbf{W}$, the bias vectors $\mathbf{b}$ and the learning rate α;
2: Divide the training data set into I mini-batches, the size of each mini-batch is M;
3: **For each batch**: Input the training set $\mathbf{S} = [\mathbf{S}_1, ..., \mathbf{S}_M]$ and $\mathbf{y} = [y_1, ..., y_M]$, where $\mathbf{S}_m = [|\hat{h}_{s,m}|, |\hat{h}_{p,m}|, |\hat{h}_{e,m}|, |t_{s,m}|, |t_{e,m}|, \epsilon_{s,m}, \epsilon_{e,m}, \epsilon_{p,m}]$;
4: For NN without any regularization, update the weights matrices $\mathbf{W}$, and the bias vectors $\mathbf{b}$ by minimizing the loss function defined in (6.29) using the back-propagation based gradient descent method that provided in (6.30) and (6.31);
5: For NN with L_1 regularization, update the weights matrices $\mathbf{W}$, and the bias vectors $\mathbf{b}$ by utilizing the back-propagation based gradient descent method that provided in (6.33) and (6.31), which are based on minimizing the loss function defined in (6.32) ;
6: For NN with L_2 regularization, update the weights matrices $\mathbf{W}$, and the bias vectors $\mathbf{b}$ by minimizing the loss function defined in (6.34) using the back-propagation based gradient descent method that provided in (6.35) and (6.31);
7: **End for**;
8: Save the trained NN;

Testing process:

1: Generate the channel coefficients for the test data set $\mathbf{S}_{test}$;
2: Feed $\mathbf{S}_{test}$ as the input parameters and determine the output results based on the trained NN;

6.5   Simulation Results

In this section, numerical results are presented to demonstrate the superior performance of the proposed NN schemes. The training and test processes are performed by using a computer with Intel Core i7-9750H processor and 16GB random access memory. The training data set is obtained by utilizing the conventional optimization scheme provided in Section 6.3 with 6×10^5 different random channel realizations. The training data set is split into two subsets of data: 5×10^5 for training data set and 10^5 for validation data set. In the training process, the mini-batch gradient descent algorithm is employed, where the batch size is chosen to be 10. Furthermore, the test data set is obtained by using 3000 channel realizations. The maximum available transmit power of SU-Tx is set to 100 mW. The transmit power of PU-Tx is assumed to be 60 mW, whereas all the noise power are set to

Fig. 6.3 The mean square error between the power allocations obtained by the conventional scheme and NN approach without regularization against different number of training steps.

be 0.001. The channels $\hat{h}_s, \hat{h}_p, \hat{h}_e, t_s$ and t_e are all generated by $\hat{h}_i = \chi_i \sqrt{d_i^{-\Phi_2}}$, $i = s, e, p$ and $t_j = \chi_j \sqrt{c_j^{-\Phi_2}}$, $j = s, e$, where $\chi_i \sim \mathcal{CN}(0, 1)$, $\chi_j \sim \mathcal{CN}(0, 1)$, d_i is the distance between the SU-Tx and the i-th user and c_j denotes distance between the PU-Tx and the i-th user. The parameter $\Phi_2 = 1.7$ denotes the path loss exponent. The distances between the transmitters and the corresponding receivers are assumed to be $d_s = 10$m, $d_e = 20$m, $d_p = 10$m, $c_s = 20$m and $c_e = 20$m. Furthermore, it is assumed that the NN has two hidden layers with one hundred neurons in each layer. The learning rate α is set to be 10^{-4} and the regularization parameter λ is assumed to be 5×10^{-4}.

First, the mean square error obtained by NN schemes without regularization, with L_1 regularization and L_2 regularization against the number of training steps are showed, respectively, in Figs. 6.3-6.5. For a better presentation, samples are taken for every 100 points from the whole training steps. It is obvious that the mean square error decreases and approaches to almost zero as the number of iterations increases. This is due to the fact that the weights $\mathbf{W}$ and the bias $\mathbf{b}$ of the proposed NN are iteratively updated by using the mini-batch gradient descent algorithm. Furthermore, the mean square error of the validation data set for the three schemes are also provided, respectively, in Figs. 6.3-6.5. As seen in these figures, the mean square errors first decreases and then remain constant, which confirms that training process does not over-fit the NN for all three different regularization cases.

Next, Fig. 6.6 presents the performance comparison in terms of optimal transmit power obtained by using the conventional optimization scheme and the proposed NN scheme against
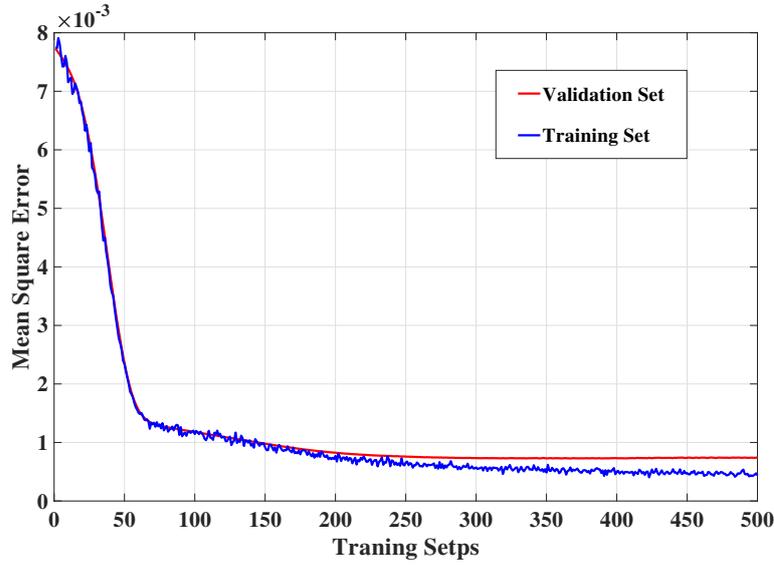
Fig. 6.4 The mean square error between the power allocations obtained by the conventional scheme and NN approach with L_1 regularization against different number of training steps.

different number of training steps. Similar to Figs. 6.3-6.5, the results are obtained by sampling the results from every 100 points of the whole training steps. As seen in this figure, the output transmit power of the proposed NN scheme is increasingly getting close to the optimal transmit power obtained from the conventional scheme with the growth of training steps. The reason is that the weights $\mathbf{W}$ and the bias $\mathbf{b}$ of the proposed NN are continuously updated in the training process to achieve the a better minimum mean square error. Note that the output power of the proposed NN may be negative or larger than the available transmit power, since the training errors between the NN output power and the optimal power obtained by conventional optimization scheme cannot be completely eliminated. In order to incorporate the power constraints ($0 \leq P_s \leq P_t$), we choose $P_s^{NN} = \min(\max(P_s^{OUT}, 0), P_t)$ as the SU-Tx transmit power of the proposed NN scheme in the following simulation results.

Next, Fig. 6.7 presents the achievable secrecy rates of the SU-Rx versus the interference leakage tolerance of the PU-Rx obtained by both conventional optimization and the proposed NN schemes with perfect CSI assumption. The maximum available transmit power of the SU-Tx is assumed to be 100 mW. It can be seen that the achievable secrecy rate increases with the interference leakage tolerance for all schemes. In addition, all the three different regularization based NN schemes can achieve a similar performance with the conventional optimization approach. Note that there is a performance gap between the conventional scheme and the three NN schemes, and this is due to the training errors between the output power and the desired optimal power.
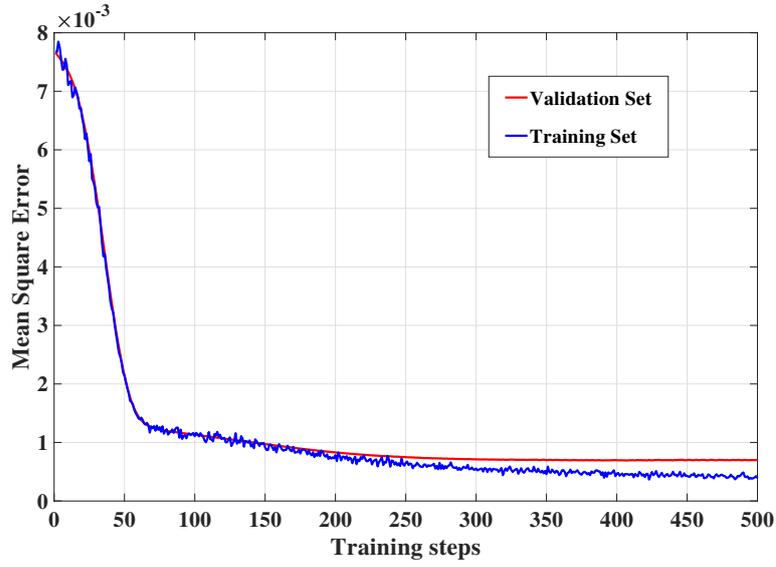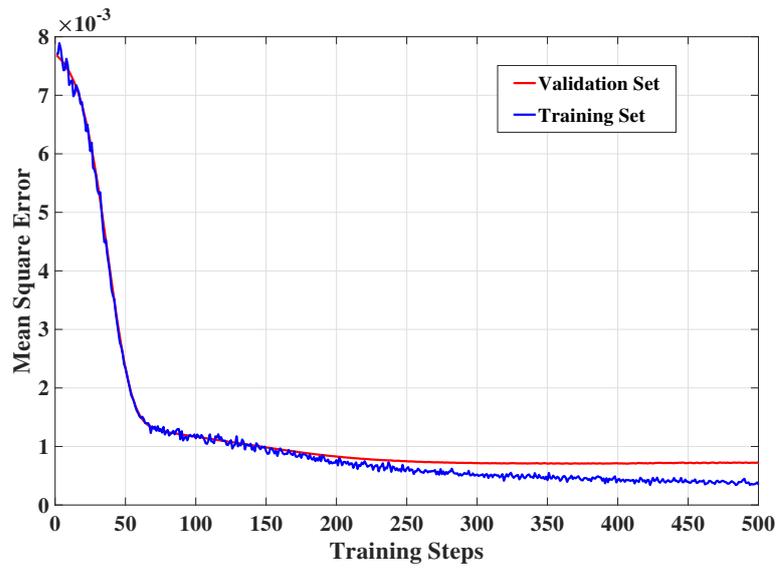
Fig. 6.5 The mean square error between the power allocations obtained by the conventional scheme and NN approach with L_2 regularization against different number of training steps.
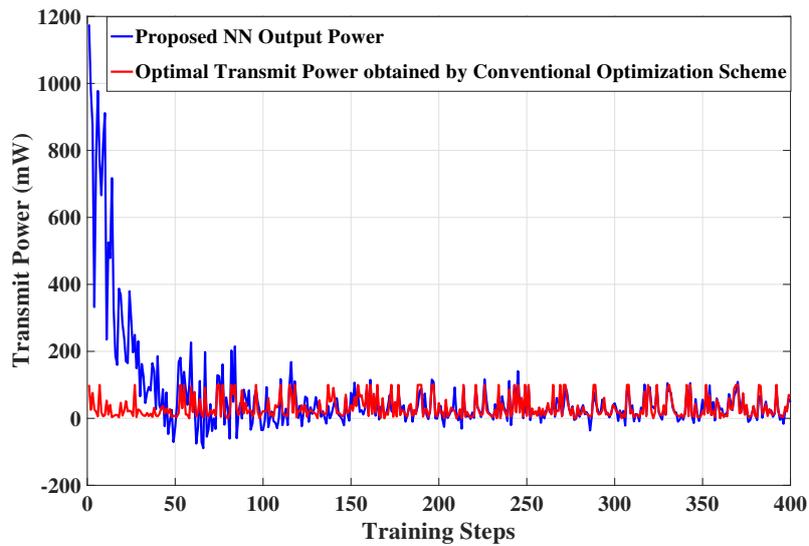
Fig. 6.6 The performance comparison in terms of the optimal transmit power obtained by using the conventional optimization scheme and the proposed NN based scheme against different number of training steps.
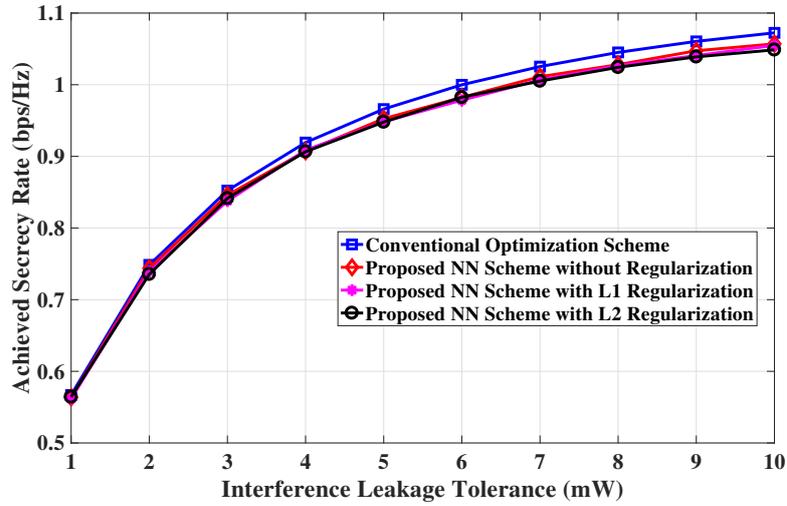
Fig. 6.7 The achievable secrecy rates against different interference tolerance of the PU-Rx obtained by both conventional optimization approach and the proposed NN framework under the perfect CSI assumption.
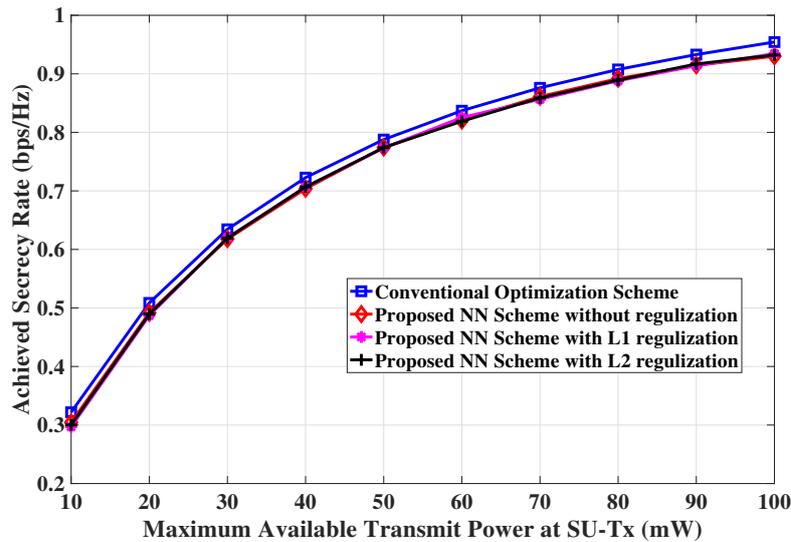
Fig. 6.8 The achievable secrecy rates against different maximum transmit power of the SU-Tx obtained by conventional optimization scheme and the proposed NN framework under the perfect CSI assumption.

Next the achievable secrecy rates of conventional optimization and proposed NN schemes against different available transmit power with perfect CSI are evaluated. Fig. 6.8 presents the achievable secrecy rates of SU-Rx of both conventional optimization and the proposed NN schemes. The interference leakage tolerance is set to be 6 mW. It can be seen that the achievable secrecy rate increases as the transmit power increases for all schemes. Similar to
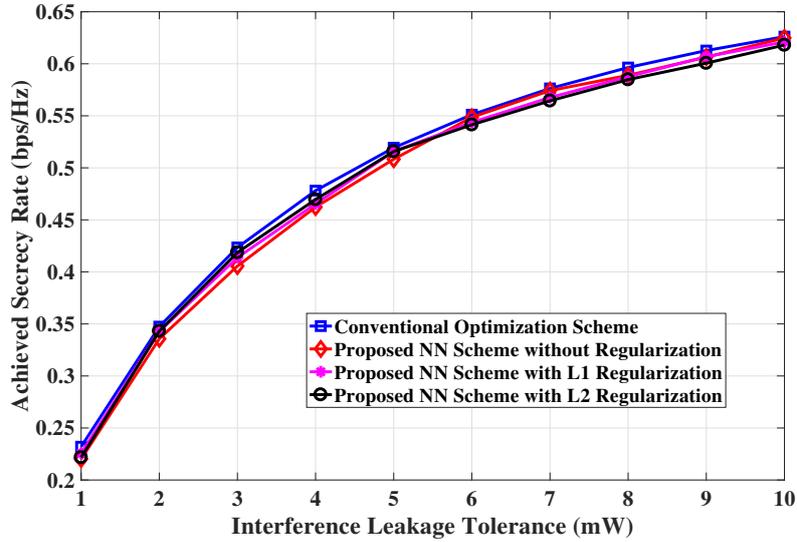
Fig. 6.9 The achievable secrecy rates against different interference tolerance of the PU-Rx obtained by both conventional optimization approach and the proposed NN framework under the imperfect CSI assumption.

Fig. 6.7, the proposed NN schemes also shows a similar performance as the conventional optimization approach.

Next, Fig. 6.9 presents the achievable secrecy rates versus different interference leakage tolerances at the PU-Rx obtained by both conventional optimization and the proposed NN schemes under imperfect CSI assumption. The channel error bound is assumed to be $\epsilon_s = \epsilon_e = \epsilon_p = 0.1$. The maximum available transmit power of SU-Tx is assumed to be 100 mW. As seen in Fig. 6.9, the achievable secrecy rate increases as the interference leakage tolerance increase for all schemes. Furthermore, all the three different regularization based NN schemes show similar performances compared to that of the conventional optimization approach.

Next the achievable secrecy rates of conventional optimization and the proposed NN schemes with different available transmit power at SU-Tx. Fig. 6.10 presents the achievable secrecy rates of SU-Rx of both schemes. The interference leakage tolerance is set to be 6 mW. As seen in Fig. 6.10, the achievable secrecy rate increases with the interference leakage tolerance increasing. Similar to previous results, the proposed NN schemes also provide a similar performance as the conventional convex optimization approach.

The achievable secrecy rates of conventional optimization and proposed NN schemes with different channel error bounds is provided in Fig. 6.11. The maximum available transmit
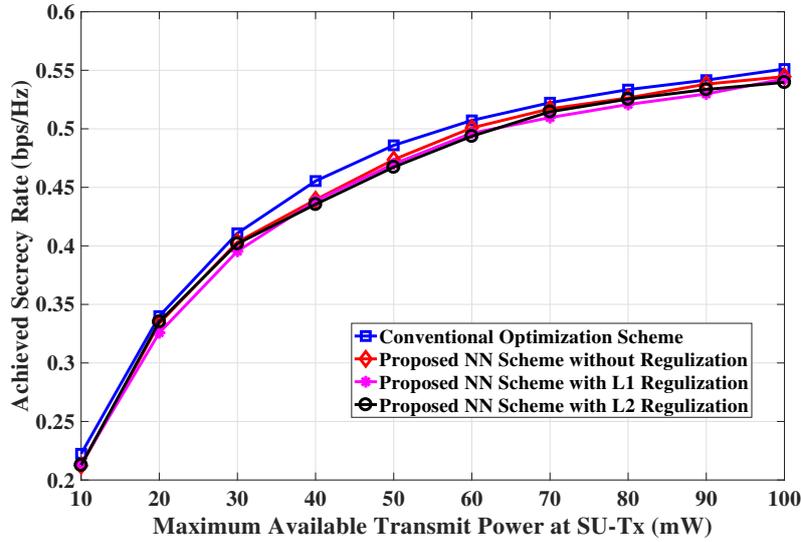
Fig. 6.10 The achievable secrecy rates against different maximum transmit power of the SU-Tx obtained by conventional optimization scheme and the proposed NN framework under the imperfect CSI assumption.

power at SU-Tx is set to be 100 mW and the interference leakage tolerance at PU-Rx is assumed to be 6 mW. All the channel error bounds are assumed to be the same for each point in this figure, i.e., $\epsilon_s = \epsilon_e = \epsilon_p$. As seen in this figure, the achieved secrecy rate decreases as the channel error bound increases for all the schemes. Furthermore, as observed in previous set of simulation results that proposed NN scheme can achieve the similar performances in comparison with conventional convex optimization scheme.

In Fig. 6.12, the achieved secrecy rate (left axis) and computation time (right axis) against different number of hidden layers for the NN scheme without any regularization are presented. The maximum available transmit power at the SU-Tx and the interference leakage tolerance at the PU-Rx are assumed to be 100 mW and 9 mW, respectively. All the channel error bounds are set to be 0 to represent the perfect CSI scenarios. As shown in this figure, the differences of performances among different number of hidden layers are within a range of 1%. However, the computation time for the testing set increases as the number of hidden layers increase. In other words, introducing more hidden layers cannot lead to much performance improvement, and these additional layers will introduce much more computation complexity to the NN.

Next the statistical results are presented in Figs. 6.13 and 6.14 to evaluate the interference leakage tolerance satisfaction at the PU-Rx of the proposed NN scheme. These statistical results are calculated by combing the results of test data of both perfect and imperfect CSI
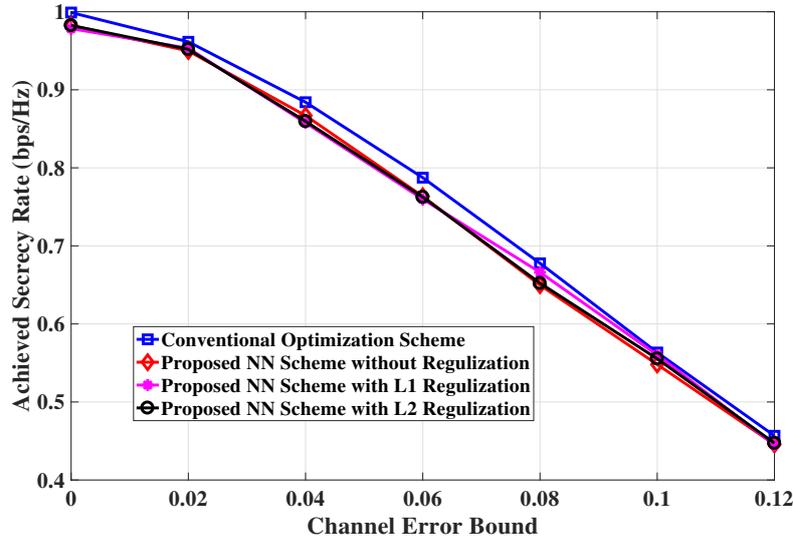
Fig. 6.11 The achievable secrecy rates against different channel error bounds obtained by conventional optimization scheme and the proposed NN framework.

scenarios. In Fig. 6.13, the interference leakage tolerance is set to 6 mW, while the maximum available transmit power at the SU-Tx is assumed to be 100 mW in Fig. 6.14. Fig. 6.13 provides the interference leakage tolerance satisfaction against different maximum transmit power, whereas Fig. 6.14 presents that against different interference leakage tolerances. As shown in these figures, more than 93% of the test results can meet the interference leakage constraint at the PU-Rx.

Table 6.1 The achieved secrecy rates of both schemes against interference leakage tolerances.

Perfect CSI:					
Interference leakage tolerance (mW)	NN scheme without regularization (bps/Hz)	NN scheme with L_1 regularization (bps/Hz)	NN scheme with L_2 regularization (bps/Hz)	Conventional scheme (bps/Hz)	Minimum ratio (%)
1	0.5636	0.5618	0.5642	0.5679	98.93
2	0.7438	0.7393	0.7355	0.7486	98.25
3	0.846	0.8377	0.8414	0.8523	98.29

Imperfect CSI:					
Interference leakage tolerance (mW)	NN scheme without regularization (bps/Hz)	NN scheme with L_1 regularization (bps/Hz)	NN scheme with L_2 regularization (bps/Hz)	Conventional scheme (bps/Hz)	Minimum ratio (%)
1	0.2202	0.2256	0.2216	0.2320	94.91
2	0.3357	0.3427	0.3431	0.3474	96.63
3	0.4056	0.4138	0.4186	0.4235	95.77

Table 6.1 provides the secrecy rate performance of all the schemes with perfect and imperfect CSI against different interference leakage tolerances, similar to the results presented
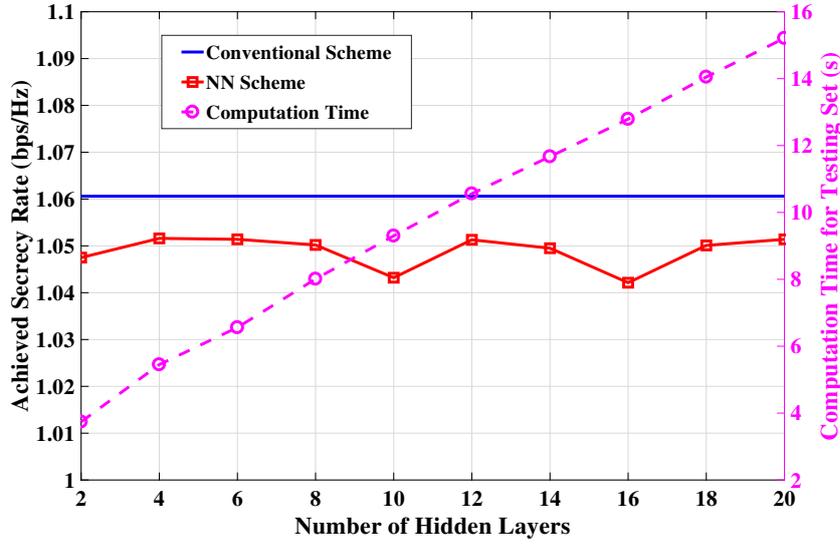
Fig. 6.12 The achievable secrecy rates (left axis) and computation time for the training set (right axis) against different number of hidden layers.

Table 6.2 The required computational time for both schemes against interference leakage tolerances.

Perfect CSI:

Interference leakage tolerance (mW)	NN scheme without regularization (s)	NN scheme with L_1 regularization (s)	NN scheme with L_2 regularization (s)	Conventional scheme (s)	Maximum ratio (%)
1	3.59	4.40	4.38	558.71	0.79
2	3.69	4.58	4.47	584.68	0.78
3	3.77	4.45	4.46	573.38	0.78

Imperfect CSI:

Interference leakage tolerance (mW)	NN scheme without regularization (s)	NN scheme with L_1 regularization (s)	NN scheme with L_2 regularization (s)	Conventional scheme (s)	Maximum ratio (%)
1	3.72	4.67	4.22	651.18	0.65
2	3.71	4.79	4.39	639.32	0.75
3	3.62	4.74	4.29	647.28	0.73

in Figs. 6.7 and 6.9. Table 6.2 presents the comparison of the required computation time of the four schemes against different interference leakage tolerances. Similarly, Table 6.3 presents the achieved the secrecy rate performance of all schemes against different available transmit power. Table 6.4 provides the required computation time of all schemes against different maximum transmit power. To draw a performance comparison of achieved secrecy rate, the minimum ratio is calculated, which is determined by dividing the minimum achieved secrecy rate among the three NN schemes by that of the conventional scheme. Similarly, for the comparison of computational time, the maximum ratio is calculated, which can be
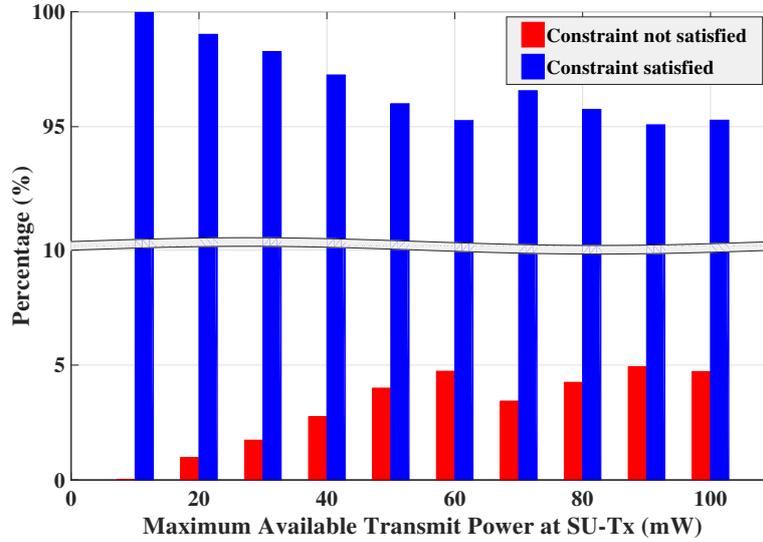
Fig. 6.13 Distributions of the interference leakage satisfactions against different maximum available transmit power at SU-Tx.

Table 6.3 The achieved secrecy rates of both schemes against maximum transmit power.

Perfect CSI:					
Maximum available transmit power (mW)	NN scheme without regularization (bps/Hz)	NN scheme with L_1 regularization (bps/Hz)	NN scheme with L_2 regularization (bps/Hz)	Conventional scheme (bps/Hz)	Minimum ratio (%)
10	0.5636	0.5618	0.5642	0.5679	98.93
20	0.7438	0.7393	0.7355	0.7486	98.25
30	0.846	0.8377	0.8414	0.8523	98.29

Imperfect CSI:					
Maximum available transmit power (mW)	NN scheme without regularization (bps/Hz)	NN scheme with L_1 regularization (bps/Hz)	NN scheme with L_2 regularization (bps/Hz)	Conventional scheme (bps/Hz)	Minimum ratio (%)
10	0.2123	0.2138	0.2127	0.2224	95.46
20	0.3338	0.3258	0.3351	0.3400	95.82
30	0.4033	0.3954	0.4019	0.4108	96.25

obtained by dividing the maximum computation of the three NN schemes by that of the conventional scheme. Note that the testing process for all schemes are performed on the same computer. In all the results provided in these tables, the achievable secrecy rates are obtained by averaging the results over test data set with 3000 channel realizations, while the computational time is the total computation time of 3000 channel realizations. From these results, it is obvious that the proposed NN schemes achieve at least 94% of the optimal performance of the conventional scheme, while significantly reducing the required computation time. In particular, the proposed NN based schemes require only less than 1% of the time that required for the conventional optimization scheme. This is due to the fact that
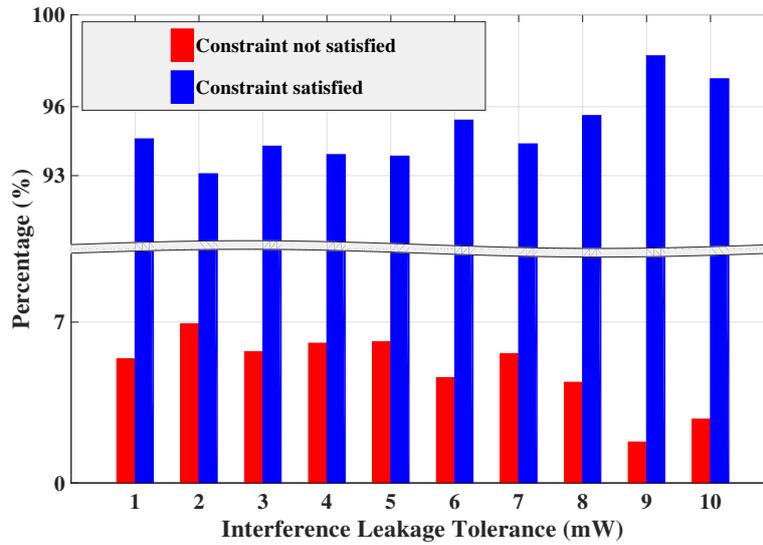
Fig. 6.14 Distributions of the interference leakage satisfactions against different interference leakages at PU-Rx.

Table 6.4 The required computational time for both schemes against maximum transmit power.

Perfect CSI:

Maximum available transmit power (mW)	NN scheme without regularization (s)	NN scheme with L_1 regularization (s	NN scheme with L_2 regularization (s)	Conventional scheme (s)	Minimum ratio (%)
10	4.21	5.03	5.34	578.06	0.92
20	3.90	4.65	4.77	574.38	0.83
30	3.63	4.39	4.58	566.87	0.81

Imperfect CSI:

Maximum available transmit power (mW)	NN scheme without regularization (s)	NN scheme with L_1 regularization (s)	NN scheme with L_2 regularization (s)	Conventional scheme (s)	Minimum ratio (%)
10	3.87	4.62	4.76	658.36	0.72
20	4.07	4.52	4.44	654.87	0.69
30	3.93	4.82	4.68	661.43	0.73

the conventional optimization based solutions of the perfect CSI assumption are obtained through an iterative approach and sub-gradient algorithms, while the conventional scheme for the imperfect CSI assumption only require sub-gradient algorithms. These algorithms and approach would requires a much higher computation time. In the NN based schemes, once the weights are determined, it is should be able to compute the solution with a reasonable complexity within a short time compared to that of the conventional approach.

6.6 Summary

In this chapter, a NN based approach for the power allocation design is proposed to maximize the secrecy rate in a CR network under transmit power and interference leakage constraints. It is showed that the developed NN algorithm has the capability to solve the power allocation problem with both perfect and imperfect CSI whereas it required to develop both robust and non-robust optimization frameworks in the conventional approaches. First, the conventional optimization scheme for perfect CSI scenario was developed based on a one dimensional search, while that for the imperfect assumption was developed based on Charnes-Cooper transformation and S-Procedure approach. Then, the NN based scheme was proposed where a relationship between the input and output parameters is established by determining an approximated function. The training set to establish the relationship between inputs and output was obtained through the conventional optimization approaches and the NN has been trained to determine the weights of the connections in the network. When the NN has been trained, the performance was evaluated with a test set in terms of achieved secrecy rates and required computational time. Numeric results demonstrated that the proposed NN scheme can achieve more than 94% of the secrecy rate performance with less than 1% computation time and more than 93% satisfaction of interference leakage constraints compared those of the conventional approaches. Simulation results were provided to demonstrate the effectiveness of the proposed NN based approach over the bench mark conventional optimization schemes.

6.7 Appendix

6.7.1 Proof of Proposition 6.1

In order to achieve a similar performance through the functions provided in (6.25) and (6.28), we should maximize the following likelihood function:

$$L(\mathbf{W}, \mathbf{b}) = \prod_{m=1}^{M} p_m(y|x; \mathbf{W}, \mathbf{b}) = \prod_{m=1}^{M} \exp\left(-\frac{(f_m(\mathbf{S}, \mathbf{W}, \mathbf{b}) - P_{s,m}^*)^2}{2\sigma^2} \right). \qquad (6.36)$$

By utilizing the monotonicity of logarithmic function, the logarithmic likelihood function can be expressed as

$$\log L(\mathbf{W}, \mathbf{b}) = \log \prod_{m=1}^{M} \exp\left(-\frac{(f_m(\mathbf{S}, \mathbf{W}, \mathbf{b}) - P_{s,m}^*)^2}{2\sigma^2}\right)$$

$$= M \log\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2\sigma^2}\sum_{m=1}^{M}(f_m(\mathbf{S}, \mathbf{W}, \mathbf{b}) - P_{s,m}^*)^2. \tag{6.37}$$

Since $M \log \frac{1}{\sqrt{2\pi}\sigma}$ and $\frac{1}{2\sigma^2}$ are constants, maximizing the likelihood function is equivalent to minimizing the following loss function:

$$J(\mathbf{W}, \mathbf{b}) = \sum_{m=1}^{M}(y_m - P_{s,m}^*)^2. \tag{6.38}$$

Furthermore, this loss function can be normalized without loss of generality as follows:

$$J(\mathbf{W}, \mathbf{b}) = \frac{1}{M}\sum_{m=1}^{M}(y_m - P_{s,m}^*)^2, \tag{6.39}$$

which completes the proof of Proposition 6.1. ∎

6.7.2   Proof of Proposition 6.2

First, we provide the following basic chain rule:

$$\frac{\partial h(g)}{\partial z} = \frac{\partial h}{\partial g}\frac{\partial g}{\partial z}. \tag{6.40}$$

From the feed-forward process, we have

$$\mathbf{z}^{(l+1)} = \mathbf{W}^{(l)}\mathbf{a}^{(l)} + \mathbf{b}^{(l)}, \tag{6.41}$$

$$\mathbf{a}^{(l+1)} = g(\mathbf{z}^{(l+1)}), \tag{6.42}$$

where $\mathbf{z}^{(l+1)}$ is the linear transformation of a given set of input parameters at the $l + 1$-th layer, whereas $\mathbf{a}^{(l+1)}$ is the output activation value of the $l + 1$-th layer. The function $g(\mathbf{z})$ represents any activation function. By assuming that $J(\mathbf{W}, \mathbf{b})$ is the loss function of the NN,

we have the following based on the chain rule defined in (6.40):

$$\frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{W}^{(l)}} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}^{(l+1)}} \frac{\partial \mathbf{z}^{(l+1)}}{\partial \mathbf{W}^{(l)}}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{\delta}_m^{(l+1)} (\mathbf{a}_m^{(l)})^T, \tag{6.43}$$

$$\frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{b}^{(l)}} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}^{(l+1)}} \frac{\partial \mathbf{z}^{(l+1)}}{\partial \mathbf{b}^{(l)}}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{\delta}_m^{(l+1)}. \tag{6.44}$$

Since we can calculate $\mathbf{a}_m^{(l)}$ from feed-forward process, then, the $\boldsymbol{\delta}_m^{(l)}$ can be derived as follows. Based on the chain rule, we have

$$\boldsymbol{\delta}_m^{(l)} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}_m^{(l)}} = \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{z}_m^{(l+1)}} \frac{\partial \mathbf{z}_m^{(l+1)}}{\partial \mathbf{a}_m^{(l)}} \frac{\partial \mathbf{a}_m^{(l)}}{\partial \mathbf{z}_m^{(l)}}$$

$$= [(\mathbf{W}^{(l)})^T \boldsymbol{\delta}_m^{(l+1)}] \cdot g'(\mathbf{z}_m^{(l)}). \tag{6.45}$$

Starting from the output layer, we can calculate $\boldsymbol{\delta}^{(l)}$ back forward layer by layer until the input layer. Finally, considering the gradient descent method, the weights matrix $\mathbf{W}^{(l)}$ and the bias vector $\mathbf{b}^{(l)}$ for the l-th layer can be updated respectively as follows:

$$\mathbf{W}^{(l)} = \mathbf{W}^{(l)} - \frac{\alpha}{M} \sum_{m=1}^{M} [\boldsymbol{\delta}_m^{(l+1)} (\mathbf{a}_m^{(l)})^T], \tag{6.46}$$

$$\mathbf{b}^{(l)} = \mathbf{b}^{(l)} - \frac{\alpha}{M} \sum_{m=1}^{M} \boldsymbol{\delta}_m^{(l+1)}, \tag{6.47}$$

which completes the proof of of Proposition 6.2. ∎

Chapter 7

Conclusions and Future Work

The exponential growth of mobile services and applications has brought increasing demands and major challenges in terms of security in wireless information transmission. The conventional security approaches are implemented in upper layers, which mainly rely on the complexity of some intractable mathematical problems. Although these techniques are proven to be uncrackable, there still exists potential risks in the processes of secret key management and distribution. As a promising solution to tackle these issues and further improve the security in mobile communications, physical layer security has been recently recognized as a promising new design paradigm to provide security in wireless networks in addition to the existing conventional cryptographic methods, where physical layer dynamics of fading channels are exploited to establish secured wireless links. On the perspective of physical layer security, confidential transmission can be guaranteed if the SNR or SINR of the legitimate channel is larger than that of the channel of the EVEs. In order to further improve the performance of physical layer security, AN was employed to mask the secret information between the legitimate terminals and confuse the EVEs. Furthermore, the computational complexities associated with convex optimization techniques are not affordable in low power devices or not suitable for ultra low latency future wireless networks. Hence, machine learning techniques have received a huge research interests for physical layer security wireless communications in recent years, which have the capability to deal with the real-time parameters that require frequent updates in very short time frame.

The unprecedented growth in mobile data traffic in recent years brings different challenges on the aspects surrounding the energy consumption in wireless networks. The energy consumption, regardless of the underpinning technologies, has a knock-on effect on the environment we live in, carbon footprint, global warming, and thus unanticipated financial

consequences. Wireless EH is an emerging technology, which facilitates the mobile devices to collect energy from external energy sources without any wired connections. In general, conventional EH methods harvest energy from natural sources, such as wind, solar and waves, such an extended approach is impractical for at least two reasons. First, these external energy sources are unreliable and heavily correlated to the environments conditions. Secondly, practically it is infeasible to exploit them to the mobile devices, owing to the size limitations of harvesting devices and the unstable power output caused by different geographical conditions. As such, wireless EH facilitates practical design and implementation especially in mobile devices.

This thesis has developed different secure transmission techniques based on the concepts of physical layer security for different communication scenarios. Various techniques have been employed, such as convex optimization, AN, PS, SWIPT and machine learning, for the purposes of improving security performance, extending the battery lifetime and dealing with the real-time processing.

Chapter 4 has investigated a beamforming design for downlink transmission of a MISO system in the presence of purely unknown EVEs, where each legitimate user employs a PS based SWIPT technique. Since the transmitter did not have any knowledge of the CSI of the EVEs, the AN approach was employed to establishing secure communication. Based on the assumption of imperfect CSI of legitimate users at the transmitter, two robust design approaches for the joint beamforming and PS ratio have been proposed to maximize the AN power under both energy EH and SINR requirements at each legitimate user. These two problems have been transformed into two convex forms and solved by using convex optimization software. The performance of the proposed designs were evaluated in terms of the maximum SINR among EVEs. Simulation results were provided to demonstrate the effectiveness and superior performance of the proposed robust designs while comparing with that of no AN aided designs, and to presented that AN technique can be utilized to mask the confidential information and significantly reduce the achieved SINR at EVEs. Then, it was showed that the achieved SINR at EVEs can be reduced by introducing more transmit antennas. This is due to the fact that introducing larger number of transmit antennas providing additional degrees of freedom in the beamforming design. Furthermore, the performance comparisons between linear and non-linear EH model were provided, which indicated that non-linear EH model is more accurate under the scenario of continuous EH requirements.

Then, in order to achieve energy efficient secure transmission, different SEE optimization problems were studied in Chapter 5 for a MISO underlay CR network, in the presence of an ER. These energy efficient designs were developed with different assumptions of CSI

at the transmitter, namely perfect CSI, statistical CSI and imperfect CSI with bounded channel uncertainties. The beamforming was designed to maximize the SEE while satisfying all relevant constraints. The original problems were non-convex and their solutions are intractable. To overcome this non-convexity issue, the original problem into was reformulated into a convex one by exploiting SDR, non-linear fractional and DC programming. The beamforming design is then extended to address the problems of imperfect CSI at the transmitter. In particular, two robust designs were proposed with the assumptions of statistical CSI on the ER's channel and error bounded channel uncertainties, respectively. For the case of statistical CSI assumption, the outage probability constraints was first expressed in closed form expressions. By exploiting SDR, non-linear fractional and DC programming, this problem was efficiently solved through an iterative approach. Next, the robust design with bounded channel uncertainties is studied. By exploiting the same techniques as in the previous robust design, the problem is reformulated into a series of SDP. Then, S-procedure is utilized to convert this problem into a convex one. The performance of the proposed designs were evaluated in terms of the achieved SEE. Numerical results, when compared with those obtained with existing techniques in the literature, showed the proposed SEE maximization designs outperform the secrecy rate maximization and power minimization schemes. These results also validated the convergence of the proposed iterative algorithms and demonstrated the impacts of different parameters on the achieved SEE. Furthermore, the methods to construct rank-one optimal covariance matrices for all the three designs were provided, which confirmed the optimality of the proposed SEE maximization based beamforming designs.

The algorithms developed based on the conventional optimization approaches often require iterative techniques with a considerable number of iterations, which poses challenges for real-time applications with stringent delay requirements in future wireless networks. To meet the unprecedented requirements of future ultra-reliable low-latency networks, Chapter 6 has studied two different approaches to maximize the secrecy rate of the SU-Rx under the constraints of total transmit power of SU-Tx and the interference leakage of the PU-Rx, namely, conventional optimization framework and NN based algorithm. It was showed that the developed NN algorithm has the capability to solve the power allocation problem with both perfect and imperfect CSI whereas it required to develop both robust and non-robust optimization frameworks in the conventional approaches. First, the conventional optimization scheme for perfect CSI scenario was developed based on a one dimensional search, while that for the imperfect assumption was developed based on Charnes-Cooper transformation and S-Procedure approach. Then, the NN based scheme was proposed where a relationship between the input and output parameters is established by determining an approximated

function. The performance of the proposed NN based scheme was evaluated with a test data set in terms of achieved secrecy rates and required computational times. Simulation results were provided to demonstrate that the proposed NN was able to achieve more than 94% secrecy rate performance and less than 1% computation time in comparing with the bench mark conventional optimization scheme. The results also showed that at least 93% CSI combinations among the testing data set in the NN scheme could satisfy the interference leakage tolerance constraints. Furthermore, the significant reduction in computational time confirmed its suitability to employ in future real-time wireless applications.

7.1   Future Work

In this thesis, the research work of employing different techniques for physical layer security has been presented. However, several limitations of the research work in this thesis have been recognized: (1) The cooperative jamming has not been considered for secure transmission in the presence of unknown EVEs; (2) The SEE designs in this thesis only considered single user scenario; (3) The computational complexities associated with conventional optimization techniques are not affordable in low powered devices in IoT or not suitable for applications with ultra reliability and low latency in future wireless networks; (4) Conventionally, NNs work in real domain [208], which introduces to handle resource allocation problems for multiple-antenna systems. Therefore, the potential extensions of the research in this thesis are presented in following subsections.

7.1.1   Cooperative Jamming embedded Designs

Cooperative jamming also has the capability to mask the confidential information and reduce the SINR of unknown EVEs. The fundamental idea of this approach is introducing a cooperative jammer to the system, which can send jamming signals to confuse the unknown EVEs.Therefore, the potential extension direction of the work in Chapter 4 will be cooperative jamming based robust designs with unknown EVEs. Furthermore, combing AN and cooperative jamming is also an interesting topic in future work. However, the problems will be more complicated when utilizing cooperative jamming technique. Hence, it is emerging to find effective ways to solve these problems.
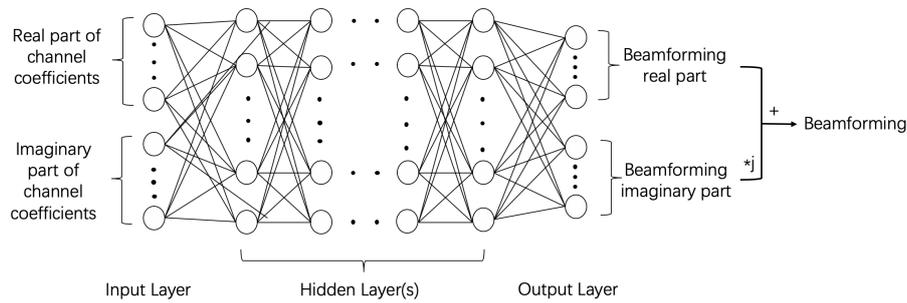
Fig. 7.1 The NN model for multiple antenna wireless communication systems.

7.1.2   SEE Designs

In this thesis, SEE maximization designs have been studied, where it is assumed that there only exists one user of each user group (i.e., PU-Rx, SU-Tx and ER). However, this assumption might not always happen in practical scenarios. The future extension direction of the work in Chapter 5 will focus on maximize SEE in multiple user CR networks. Therefore, it is important to study effective algorithms to deal with the more complex SEE maximization problems. Furthermore, the SEE another possible future work will be focused on SEE designs for other wireless communication scenarios, i.e., relay assisted networks and unmanned aerial vehicle based networks.

7.1.3   Machine Leaning based Resource Allocation

Another potential area of future research work will be paid attention to machine leaning based resource allocation for physical layer security. Conventional NNs employ real derivations to determine their weights and bias through minimizing loss functions. However, there might arise a problem that the complex derivations exist if and only if they satisfy the Cauchy-Riemann equations. In the complex domain, the functions that satisfy these equations are called holomorphic functions, otherwise, they are called non-holomorphic functions [209]. This stringent condition brings challenges for employing NNs to learn to optimize in multiple antenna wireless communication scenarios, since NNs need to deal with different parameters in complex domains for these scenarios. There exist two potential solutions for utilizing NNs in multiple antenna wireless communication systems.

The first method is to separate both complex input and output parameters into two parts: the real part and imaginary part [208]. As shown in Fig. 7.1, both parts of the input channel coefficients and output beamforming of the NN are connected to the real and

imaginary components of the corresponding complex parameters. Then, the NN can perform its calculations in real domain. The performance of the NN based approach will be evaluated with a test data set once the NN is trained with appropriate training data set. For the test procedure, one more step needs to be performed: constructing the complex beamforming vector by appropriately combining the imaginary part of the output and the corresponding real part of the output.

The second approach to handle the complex parameters is to employ the Wirtinger calculus to deal with non-holomorphic functions in complex domain [209]. The basic idea of Wirtinger calculus is to perform the differentiable mapping for non-holomorphic functions in complex domain by introducing conjugate coordinates. In other words, this mapping is viewed from pseudo-real perspective in terms of conjugate coordinates. The back propagation based complex gradient descent algorithm for NNs has been investigated in [209], which enables NNs to handle all computations directly on complex variables without separating the data into two parts.

Another interesting machine learning technique that can be utilized in wireless communications is reinforcement learning, which is implemented based on the Markov decision processes. Recently, applications of reinforcement learning in wireless communication has attracted significant research interests [34, 210, 211]. Several designs have been proposed by applying reinforcement learning to solve various problems in wireless communications, including resource allocation problems in V2V communications [138], IoT networks [212, 213], CR networks [214] and heterogeneous networks [215]. Moreover, Q-learning, which is a special form of reinforcement learning also received much research attentions [34]. The main idea of Q-learning is to determine the optimal action in an unknown system [34]. In other words, Q-learning is a model free learning method. Furthermore, Q-learning can be applied in heterogeneous networks to solve both power allocation and interference control problems in femto/small cells [216]. Reinforcement learning approaches, including Q-learning, can be exploited for resource allocation problems for physical layer security. For example, the objective secrecy rate function can be chosen as the rewards, the constraints can be assumed to be the states and thus the feasible region of the constraints are the actions. Then, the some iterative computations are performed to determined the optimal actions that maximize the rewards.

List of Symbols

$(\cdot)^*$ Complex Conjugate

$(\cdot)^{-1}$ Inverse

$(\cdot)^H$ Hermitian Transpose

$(\cdot)^T$ Transpose

$[x]^+$ $\max\{x, 0\}$

$\Im\{\cdot\}$ Imaginary Part

$\log(x)$ The Logarithm of x to Base 10

$\mathbb{E}\{\cdot\}$ Statistical Expectation

$\mathbf{A} \succeq \mathbf{0}$ $\mathbf{A}$ is a Positive Semidefinite Matrix

$\mathcal{CN}(\mu, \sigma^2)$ Complex Gaussian Random Variable with Mean μ and Variance σ^2

$\mathrm{dom}\, f(x)$ Domain of the Function $f(x)$

$\mathbf{I}$ Identity Matrix

$\mathcal{L}(\cdot)$ Lagrange Dual Function

$|\, x\, |$ Modulus of Complex Number x

$\|\cdot\|_1$ 1-Norm

$\|\cdot\|_2$ Euclidean Norm

ln   Natural Logarithm

tr$(\cdot)$ Trace Operator

vec$(\mathbf{A})$ Vectorization of $\mathbf{A}$

$Re\{\cdot\}$ Real Part

List of Abbreviations

1G	First Generation
2G	Second Generation
3G	Third Generation
4G	Fourth Generation
5G	Fifth Generation
AES	Advanced Encryption Standard
AI	Artificial Intelligence
AN	Artificial Noise
AWGN	Additive White Gaussian Noise
CR	Cognitive Radio
CSI	Channel State Information
D2D	Device-to-Device
DC	Difference of Concave
DES	Data Encryption Standard
EE	Energy Efficiency
EH	Energy Harvesting
ER	Energy Receiver

EVE	Eavesdropper
GPS	Global Positioning System
GSVD	Generalized Singular Value Decomposition
IoT	Internet of Things
IR	Information Receiver
KKT	Karush Kuhn Tucker
LMI	Linear Matrix Inequality
LP	Linear Programming
MIMO	Multiple-input Multiple-ouput
MISO	Multiple-input Single-ouput
NN	Neural Network
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PS	Power Splitting
PU-Rx	Primary Receiver
PU-Tx	Primary Transmitter
QCQP	Quadratically Constrained Quadratic Programming
QoE	Quality of Experience
QoS	Quality of Service
QP	Quadratic Programming
ReLU	Rectified Linear Unit
RF	Radio Frequency
RSA	Rivest–Shamir–Adleman
SDP	Semidefinite Programming

SDR	Semidefinite Relaxation
SEE	Secrecy Energy Efficiency
SINR	Signal to Noise plus Interference Ratio
SISO	Single-input Single-output
SNR	Signal to Noise Ratio
SOC	Second-Order Cone
SOCP	Second-Order Cone Programming
SON	Self-Organizing Networks
SU-Rx	Secondary Reciever
SU-Tx	Secondary Transmitter
SWIPT	Simultaneous Wireless Information Power Transfer
TS	Time Splitting
w.r.t.	With respect to
WiFi	Wireless Fidelity
WPT	Wireless Power Transfer
WSN	Wireless Sensor Network

References

[1] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.

[2] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[3] J. C. Maxwell, *A treatise on electricity and magnetism*. Oxford: Clarendon Press, 1873, vol. 1.

[4] A. A. Huurdeman, *The worldwide history of telecommunications*. Wiley Online Library, 2003.

[5] Carritech, "The evolution of mobile communications, from 1G to 5G," *http://www.carritech.com/news/evolution-mobile-communication-1g-5g/*.

[6] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Area Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[7] A. C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Beamforming design for full-duplex MIMO interference channels–QoS and energy-efficiency considerations," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4635–4651, Sep. 2016.

[8] R. Vannithamby and S. Talwar, *Towards 5G: Applications, requirements and candidate technologies*. John Wiley & Sons, 2017.

[9] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, "Physical layer security in wireless networks: a tutorial," *IEEE Commun. Mag.*, vol. 18, no. 2, pp. 66–74, Apr. 2011.

[10] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *ACM Commun.*, vol. 21, no. 2, pp. 120–126, 1978.

[11] M. Zhang, K. Cumanan, and A. Burr, "Secrecy rate maximization for MISO multicasting SWIPT system with power splitting scheme," in *Proc. IEEE SPAWC*, Jul., Edinburgh 2016, pp. 1–5.

[12] M. Zhang, K. Cumanan, and A. G. Burr, "Secure energy efficiency optimization for MISO cognitive radio network with energy harvesting," in *Proc. IEEE WCSP*, Nanjing, Oct. 2017.

[13] M. Zhang, K. Cumanan, J. Thiyagalingam, W. Wang, A. G. Burr, Z. Ding, and O. A. Dobre, "Energy efficiency optimization for secure transmission in MISO cognitive radio network with energy harvesting," *IEEE Access*, vol. 7, pp. 126 234–126 252, Sep. 2019.

[14] K. Cumanan, H. Xing, P. Xu, G. Zheng, X. Dai, A. Nallanathan, Z. Ding, and G. K. Karagiannidis, "Physical layer security jamming: Theoretical limits and practical designs in wireless networks," *IEEE Access*, vol. 5, pp. 3603–3611, Dec. 2016.

[15] K. Cumanan, Z. Ding, M. Xu, and H. V. Poor, "Secrecy rate optimization for secure multicast communications," *IEEE J. Sel. Signal Process.*, vol. 10, no. 8, pp. 1417–1432, Dec. 2016.

[16] K. Cumanan, G. C. Alexandropoulos, Z. Ding, and G. K. Karagiannidis, "Secure communications with cooperative jamming: Optimal power allocation and secrecy outage analysis," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7495–7505, Aug. 2017.

[17] P. Xu, K. Cumanan, Z. Ding, X. Dai, and K. K. Leung, "Group secret key generation in wireless networks: algorithms and rate optimization," *IEEE Trans. Inf. Forensics & Security*, vol. 11, no. 8, pp. 1831–1846, Aug. 2016.

[18] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Le Goff, "Robust outage secrecy rate optimizations for a MIMO secrecy channel," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 86–89, Feb. 2015.

[19] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Singal Process.*, vol. 58, no. 3, pp. 1875–1888, Mar. 2010.

[20] F. Jameel, S. Wyne, S. Junaid Nawaz, J. Ahmed, and K. Cumanan, "On the secrecy performance of SWIPT receiver architectures with multiple eavesdroppers," *Wireless Communications and Mobile Computing*, vol. 2018, Jun. 2018.

[21] W. Xiang, S. Le Goff, M. Johnston, and K. Cumanan, "Signal mapping for bit-interleaved coded modulation schemes to achieve secure communications," *IEEE Wireless Commun. Lett.*, vol. 4, no. 3, pp. 249–252, Jun. 2015.

[22] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1678–1690, May 2014.

[23] K. Cumanan, R. Zhang, and S. Lambotharan, "A new design paradigm for MIMO cognitive radio with primary user rate constraint," *IEEE Commun. Lett.*, vol. 16, no. 5, pp. 706–709, May 2012.

[24] K. Cumanan, Z. Ding, M. Xu, and H. V. Poor, "Secure multicast communications with private jammers," in *Proc. IEEE SPAWC*.   IEEE, Jul. Edinburgh 2016, pp. 1–6.

[25] C. Shannon, "Communication theory of secrecy systems," *Bell Sys. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct 1949.

[26] A. D. Wyner, "The wire-tap channel," *Bell Sys. Tech. J*, vol. 54, no. 8, pp. 1355–1387, Jan. 1975.

[27] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[28] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118–127, 2014.

[29] V. Raghunathan, S. Ganeriwal, and M. Srivastava, "Emerging techniques for long lived wireless sensor networks," *IEEE Commun. Mag.*, vol. 44, no. 4, pp. 108–114, 2006.

[30] L. R. Varshney, "Transporting information and energy simultaneously," *Proc. IEEE ISIT*, pp. 1612–1616, Toronto, Jul. 2008.

[31] P. Grover and A. Sahai, "Shannon meets Tesla: Wireless information and power transfer." *Proc. IEEE ISIT*, pp. 2363–2367, Austin, Jun. 2010.

[32] Z. Chu, T. A. Le, H. X. Nguyen, M. Karamanoglu, Z. Zhu, A. Nallanathan, E. Ever, and A. Yazici, "Robust design for MISO SWIPT system with artificial noise and cooperative jamming," in *Proc. IEEE GLOBECOM*, Singapore, Dec. 2017, pp. 1–6.

[33] Z. Chu, Z. Zhu, M. Johnston, and S. Y. Le Goff, "Simultaneous wireless information power transfer for MISO secrecy channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 6913–6925, Sep. 2016.

[34] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

[35] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT Press Cambridge, 2016, vol. 1.

[36] A. M. Turing, "Computing machinery and intelligence AM Turing," *Mind*, vol. 49, no. 236, pp. 433–460, Oct. 1950.

[37] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine learning*, vol. 50, no. 1-2, pp. 5–43, Jan. 2003.

[38] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.

[39] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Machine learning for wireless networks with artificial intelligence: A tutorial on neural networks," *IEEE Commun. Surveys & Tutorials*, Early Access 2019.

[40] J. W. Kalat, *Biological Psychology*. Nelson Education, 1995.

[41] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*. Martin Hagan, 2014.

[42] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys & Tutorials*, Mar. 2019.

[43] P. M. Domingos, "A few useful things to know about machine learning." *ACM Commun.*, vol. 55, no. 10, pp. 78–87, Spring 2012.

[44] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[45] W. Dinkelbach, "On nonlinear fractional programming," *Manag. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[46] T. P. Dinh and H. A. Le Thi, "Recent advances in DC programming and DCA," in *Trans. Comput. Intell. XIII*. Springer, 2014, pp. 1–37.

[47] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994, vol. 15.

[48] Z. Zhu, Z. Chu, F. Zhou, H. Niu, Z. Wang, and I. Lee, "Secure beamforming designs for secrecy MIMO SWIPT systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 424–427, Dec. 2017.

[49] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Y. Le Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1833–1847, Jul. 2014.

[50] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[51] Y. Zhang, *Physical layer security in wireless communications*. Crc Press, 2013.

[52] H. Xing, K.-K. Wong, Z. Chu, and A. Nallanathan, "To harvest and jam: A paradigm of self-sustaining friendly jammers for secure AF relaying," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6616–6631, Sep. 2015.

[53] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4599–4615, Apr. 2014.

[54] Q. Li, W.-K. Ma, and A. M.-C. So, "Robust artificial noise-aided transmit optimization for achieving secrecy and energy harvesting," in *Proc. IEEE ICASSP*, Florence, May 2014, pp. 1596–1600.

[55] Z. Chu, K. Cumanan, M. Xu, and Z. Ding, "Robust secrecy rate optimisations for multiuser multiple-input-single-output channel with device-to-device communications," *IET Commun.*, vol. 9, no. 3, pp. 396–403, Feb. 2015.

[56] M. Bloch and J. Barros, *Physical-layer security: from information theory to security engineering*. Cambridge University Press, 2011.

[57] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[58] ——, *Information theoretic security*. The Netherlands: Now Publishers, 2009.

[59] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys & Tutorials*, vol. 16, no. 3, pp. 1550–1573, Feb. 2014.

[60] D. W. K. Ng, T. Q. Duong, C. Zhong, and R. Schober, *Wireless Information and Power Transfer: Theory and Practice*. Wiley-IEEE Press, 2019.

[61] F. Zhou, X. Zhang, R. Q. Hu, A. Papathanassiou, and W. Meng, "Resource allocation based on deep neural networks for cognitive radio networks," in *Proc. IEEE ICCC*, Beijing, Aug. 2018, pp. 40–45.

[62] J. G. Proakis and M. Salehi, *Digital communications*. McGraw-hill New York, 2001, vol. 4.

[63] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Nov. 2013.

[64] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2082–2085, Sep. 2015.

[65] E. Boshkovska, A. Koelpin, D. W. K. Ng, N. Zlatanov, and R. Schober, "Robust beamforming for SWIPT systems with non-linear energy harvesting model," in *Proc. IEEE SPAWC*, Edinburgh, Jul. 2016, pp. 1–5.

[66] L. Ni, X. Da, H. Hu, Y. Huang, R. Xu, and M. Zhang, "Outage constrained robust transmit design for secure cognitive radio with practical energy harvesting," *IEEE Access*, vol. 6, pp. 71 444–71 454, Nov. 2018.

[67] L. Yang, M. O. Hasna, and I. S. Ansari, "Physical layer security for TAS/MRC systems with and without co-channel interference over η–μ fading channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12 421–12 426, Dec. 2018.

[68] Q. Li and L. Yang, "Beamforming for cooperative secure transmission in cognitive two-way relay networks," *IEEE Trans. Inf. Forensics & Security*, Early Access 2019.

[69] Q. Li and W.-K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.

[70] C. Jeong and I.-M. Kim, "Optimal power allocation for secure multicarrier relay systems," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5428–5442, 2011.

[71] C. Jeong, I.-M. Kim, and D. I. Kim, "Joint secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 310–325, Jan. 2012.

[72] G. Chen, Y. Gong, P. Xiao, and J. A. Chambers, "Dual antenna selection in secure cognitive radio networks," *IEEE Trans. Veh, Technol.*, vol. 65, no. 10, pp. 7993–8002, Oct. 2015.

[73] J. Huang and A. L. Swindlehurst, "Cooperative jamming for secure communications in MIMO relay networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4871–4884, Oct. 2011.

[74] G. Zheng, L.-C. Choo, and K.-K. Wong, "Optimal cooperative jamming to enhance physical layer security using relays," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1317–1322, Mar. 2011.

[75] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Y. Le Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1833–1847, May. 2015.

[76] Z. Chu, H. X. Nguyen, T. A. Le, M. Karamanoglu, E. Ever, and A. Yazici, "Secure wireless powered and cooperative jamming D2D communications," *IEEE Trans. Green Commun. & Networking*, vol. 2, no. 1, pp. 1–13, Oct. 2017.

[77] C. Wang, H.-M. Wang, X.-G. Xia, and C. Liu, "Uncoordinated jammer selection for securing SIMOME wiretap channels: A stochastic geometry approach," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2596–2612, May 2015.

[78] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, Jun. 2008.

[79] Q. Li and W.-K. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-eves secrecy rate maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2704–2717, May 2013.

[80] G. Chen, Z. Tian, Y. Gong, Z. Chen, and J. A. Chambers, "Max-ratio relay selection in secure buffer-aided cooperative wireless networks," *IEEE Trans. Inf. Forensics & Security*, vol. 9, no. 4, pp. 719–729, Apr. 2014.

[81] G. Chen, Y. Gong, P. Xiao, and J. A. Chambers, "Physical layer network security in the full-duplex relay system," *IEEE Trans. Inf. Forensics & Security*, vol. 10, no. 3, pp. 574–583, Mar. 2015.

[82] M. D. Yacoub, "The κ-μ distribution and the η-μ distribution," *IEEE Antennas and Propagation Magazine*, vol. 49, no. 1, pp. 68–81, 2007.

[83] M. Zhang, K. Cumanan, L. Ni, H. Hu, A. G. Burr, and Z. Ding, "Robust beamforming for AN aided MISO SWIPT system with unknown eavesdroppers and non-linear EH model," in *Proc. IEEE GLOBECOM Workshop*, Abu Dhabi, Dec. 2018.

[84] Z. Zhu, Z. Chu, N. Wang, S. Huang, Z. Wang, and I. Lee, "Beamforming and power splitting designs for AN-aided secure multi-user MIMO SWIPT systems," *IEEE Trans. Inf. Forensics & Security*, vol. 12, no. 12, pp. 2861–2874, Dec. 2017.

[85] L. Liu, R. Zhang, and K.-C. Chua, "Secrecy wireless information and power transfer with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1850–1863, Apr. 2014.

[86] Q. Li and W.-K. Ma, "Multicast secrecy rate maximization for MISO channels with multiple multi-antenna eavesdroppers," in *Proc. IEEE ICC*, Kyoto, Jun. 2011.

[87] C. Shen, T.-H. Chang, K.-Y. Wang, Z. Qiu, and C.-Y. Chi, "Distributed robust multicell coordinated beamforming with imperfect CSI: An ADMM approach," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2988–3003, Jun. 2012.

[88] S. K. Joshi, U. L. Wijewardhana, M. Codreanu, and M. Latva-aho, "Maximization of worst-case weighted sum-rate for MISO downlink systems with imperfect channel knowledge," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3671–3685, Oct. 2015.

[89] K. Cumanan, Z. Ding, Y. Rahulamathavan, M. M. Molu, and H.-H. Chen, "Robust MMSE beamforming for multiantenna relay networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 3900–3912, May 2017.

[90] Z. Chu, K. Cumanan, M. Xu, and Z. Ding, "Robust secrecy rate optimisations for multiuser multiple-input-single-output channel with device-to-device communications," *IET Communications*, vol. 9, no. 3, pp. 396–403, Feb. 2015.

[91] F. Alavi, K. Cumanan, Z. Ding, and A. G. Burr, "Robust beamforming techniques for non-orthogonal multiple access systems with bounded channel uncertainties," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2033–2036, Sep. 2017.

[92] F. Alavi, N. M. Yamchi, M. R. Javan, and K. Cumanan, "Limited feedback scheme for device-to-device communications in 5G cellular networks with reliability and cellular secrecy outage constraints," *IEEE Trans. Veh. Tech.*, vol. 66, no. 9, pp. 8072–8085, Sep. 2017.

[93] K. Cumanan, R. Krishna, V. Sharma, and S. Lambotharan, "Robust interference control techniques for cognitive radios using worst-case performance optimization," in *Proc. IEEE ISITA*. IEEE, Auckland, Dec. 2008, pp. 1–5.

[94] ——, "A robust beamforming based interference control technique and its performance for cognitive radios," in *Proc. IEEE ISITA*. IEEE, Auckland, Dec. 2008, pp. 9–13.

[95] S. Nasseri and M. R. Nakhai, "Robust interference management via outage-constrained downlink beamforming in multicell networks," in *Proc. IEEE GLOBECOM*, Atlanta, Dec. 2013, pp. 3470–3475.

[96] C. Shen, T.-H. Chang, K.-Y. Wang, Z. Qiu, and C.-Y. Chi, "Chance-constrained robust beamforming for multi-cell coordinated downlink," in *Proc. IEEE GLOBECOM*, Anaheim, Dec. 2012, pp. 4957–4962.

[97] X. He and Y.-C. Wu, "Tight probabilistic SINR constrained beamforming under channel uncertainties," *IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3490–3505, Jul. 2015.

[98] F. Alavi, K. Cumanan, Z. Dingl, and A. G. Burr, "Outage constraint based robust beamforming design for non-orthogonal multiple access in 5g cellular networks," in *Proc. IEEE PIMRC*. IEEE, Montreal, Nov. 2017, pp. 1–5.

[99] A. Mukherjee and A. L. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, Jan. 2011.

[100] L. Zhang, Y.-C. Liang, Y. Pei, and R. Zhang, "Robust beamforming design: From cognitive radio MISO channels to secrecy MISO channels," in *Proc. IEEE GLOBECOM*, Honolulu, Dec. 2009, pp. 1–5.

[101] S. Gerbracht, C. Scheunert, and E. A. Jorswieck, "Secrecy outage in MISO systems with partial channel information," *IEEE Trans. Inf. Forensics & Security,*, vol. 7, no. 2, pp. 704–716, Apr. 2012.

[102] L. Liu, R. Zhang, and K.-C. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 288–300, Jan. 2013.

[103] D. W. K. Ng and R. Schober, "Max-min fair wireless energy transfer for secure multiuser communication systems," in *Proc. IEEE ITW*, Hobart, Dec. 2014, pp. 326–330.

[104] R. Feng, Q. Li, Q. Zhang, and J. Qin, "Robust secure transmission in MISO simultaneous wireless information and power transfer system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 400–405, Jan. 2015.

[105] M. Tian, X. Huang, Q. Zhang, and J. Qin, "Robust AN-aided secure transmission scheme in MISO channels with simultaneous wireless information and power transfer," *IEEE, Signal Process. Lett.*, vol. 22, no. 6, pp. 723–727, Jun. 2015.

[106] M. Zhang, Y. Liu, and R. Zhang, "Artificial noise aided secrecy information and power transfer in OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 3085–3096, Apr. 2016.

[107] Z. Xiang and M. Tao, "Robust beamforming for wireless information and power transmission," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 372–375, Aug. 2012.

[108] A. M. Fouladgar and O. Simeone, "On the transfer of information and energy in multi-user systems," *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1733–1736, Nov. 2012.

[109] Q. Li and L. Yang, "Robust optimization for energy efficiency in MIMO two-way relay networks with SWIPT," *IEEE Sys. J.*, Early Access 2019.

[110] F. Zhou, Z. Chu, H. Sun, R. Q. Hu, and L. Hanzo, "Artificial noise aided secure cognitive beamforming for cooperative MISO-NOMA using SWIPT," *IEEE J. Sel. Area Commun.*, vol. 36, no. 4, pp. 918–931, Apr. 2018.

[111] C. R. Valenta and G. D. Durgin, "Harvesting wireless power: Survey of energy-harvester conversion efficiency in far-field, wireless power transfer systems," *IEEE Micro. Mag.*, vol. 15, no. 4, pp. 108–120, Jun. 2014.

[112] E. Boshkovska, R. Morsi, D. W. K. Ng, and R. Schober, "Power allocation and scheduling for SWIPT systems with non-linear energy harvesting model," in *Proc. IEEE ICC*, Kuala Lumpur, May 2016, pp. 1–6.

[113] Y. Lu, K. Xiong, P. Fan, T. Liu, and Z. Zhong, "SWIPT for MISO wiretap networks: Channel uncertainties and nonlinear energy harvesting features," in *Proc. IEEE GLOBECOM 2017*, Singapore, Dec. 2017, pp. 1–7.

[114] S. Wang, M. Xia, K. Huang, and Y.-C. Wu, "Wirelessly powered two-way communication with nonlinear energy harvesting model: Rate regions under fixed and mobile relay," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8190–8204, Dec. 2017.

[115] J. Hu, G. Zhang, W. Heng, and X. Li, "Optimal energy-efficient transmission in multiuser systems with hybrid energy harvesting transmitter," in *Proc. IEEE GLOBECOM.*, Washington, Dec. 2016, pp. 1–5.

[116] F. Alavi, K. Cumanan, M. Fozooni, Z. Ding, S. Lambotharan, and O. A. Dobre, "Robust energy-efficient design for MISO non-orthogonal multiple access systems," *IEEE Trans. Commun.*, Jul. 2019.

[117] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "Energy efficiency of the cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Green Commun. & Networking*, Jul. 2019.

[118] H. Al-Obiedollah, K. Cumanan, J. Thiyagalingam, A. G. Burr, Z. Ding, and O. A. Dobre, "Energy efficient beamforming design for MISO non-orthogonal multiple access systems," *IEEE Trans. Commun.*, Jun. 2019.

[119] A. Zappone, P.-H. Lin, and E. A. Jorswieck, "Energy-efficient secure communications in MISO-SE systems," in *Proc. IEEE ICASSP*, Florence, May 2014, pp. 1001–1005.

[120] Y. Yuan, Z. Chu, Z. Ding, K. Cumanan, and M. Johnston, "Joint relay beamforming and power splitting ratio optimization in a multi-antenna relay network," in *Proc. IEEE WCSP*, Hefei, Oct. 2014.

[121] Z. Chu, H. Xing, M. Johnston, and S. Le Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multi-antenna eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 283–297, Jan. 2016.

[122] P. Xu and K. Cumanan, "Optimal power allocation scheme for non-orthogonal multiple access with α-fairness," *IEEE J. Sel. Commun.*, vol. 35, no. 10, pp. 2357–2369, Oct. 2017.

[123] W. Mei, Z. Chen, and J. Fang, "Robust energy-efficient transmit design for MISOME wiretap channels," in *Proc. IEEE GlobalSIP*, Washington DC, Dec. 2016, pp. 981–985.

[124] W. Mei, Z. Chen, J. Fang, and B. Ning, "Energy efficiency optimization for MISO wiretap channels with statistical and imperfect CSI," in *Proc. IEEE SPAWC*, Sapporo, Jul. 2017, pp. 1–6.

[125] W. Mei, Z. Chen, and J. Fang, "Outage constrained robust energy efficiency optimization for MISO wiretap channels," in *Proc. IEEE VTC Spring*, Sydney, Jun. 2017, pp. 1–5.

[126] X. Wang, W. Feng, Y. Chen, Z. Chu, and N. Ge, "Energy-efficiency maximization for secure multiuser MIMO SWIPT systems with csi uncertainty," *IEEE Access*, vol. 6, pp. 2097–2109, Dec. 2017.

[127] W. Mei, L. Li, Z. Chen, and C. Huang, "Energy-efficient optimization for MISO Gaussian broadcast channel with integrated services," in *Proc. IEEE PIMRC*, Valencia, Sep. 2016, pp. 1–6.

[128] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *arXiv preprint arXiv:1803.04311*, 2018.

[129] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

[130] K. Zheng, Z. Yang, K. Zhang, P. Chatzimisios, K. Yang, and W. Xiang, "Big data-driven optimization for mobile networks toward 5G," *IEEE Network*, vol. 30, no. 1, pp. 44–51, Feb. 2016.

[131] J. Luo, J. Tang, D. K. So, G. Chen, K. Cumanan, and J. A. Chambers, "A deep learning-based approach to power minimization in multi-carrier NOMA with SWIPT," *IEEE Access*, vol. 7, pp. 17 450–17 460, Jan. 2019.

[132] D. D. Nguyen, H. X. Nguyen, and L. B. White, "Reinforcement learning with network-assisted feedback for heterogeneous rat selection," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6062–6076, Jun. 2017.

[133] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[134] R. Li, Z. Zhao, X. Zhou, G. Ding, Y. Chen, Z. Wang, and H. Zhang, "Intelligent 5G: When cellular networks meet artificial intelligence," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 175–183, Oct. 2017.

[135] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 4, pp. 2392–2431, Jul. 2017.

[136] X. Chen, J. Wu, Y. Cai, H. Zhang, and T. Chen, "Energy-efficiency oriented traffic offloading in wireless networks: A brief survey and a learning approach for heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 4, pp. 627–640, Jan. 2015.

[137] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for wireless resource management," in *Proc. IEEE SPAWC*, Sapporo, Jul. 2017, pp. 1–6.

[138] H. Ye and G. Y. Li, "Deep reinforcement learning for resource allocation in V2V communications," in *Proc. IEEE ICC*, Kansas, May 2018, pp. 1–6.

[139] M. A. Alsheikh, D. Niyato, S. Lin, H.-P. Tan, and Z. Han, "Mobile big data analytics using deep learning and Apache Spark," *IEEE Network*, vol. 30, no. 3, pp. 22–29, May 2016.

[140] Z. M. Fadlullah, F. Tang, B. Mao, N. Kato, O. Akashi, T. Inoue, and K. Mizutani, "State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2432–2455, May 2017.

[141] N. Ahad, J. Qadir, and N. Ahsan, "Neural networks in wireless networks: Techniques, applications and guidelines," *Elsevier J. Network & Computer App.*, vol. 68, pp. 1–27, Jun. 2016.

[142] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Commun. Surveys & Tutorials*, vol. 20, no. 4, pp. 2923–2960, Jun. 2018.

[143] N. D. Lane and P. Georgiev, "Can deep learning revolutionize mobile sensing?" in *Proc. ACM IWMCSA*, Feb. 2015, pp. 117–122.

[144] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Commun. Surveys & Tutorials*, vol. 20, no. 4, pp. 2595–2621, Jun. 2018.

[145] N. Samuel, T. Diskin, and A. Wiesel, "Deep MIMO detection," in *Proc. IEEE SPAWC*, Sapporo, Jul. 2017, pp. 1–5.

[146] M. A. Wijaya, K. Fukawa, and H. Suzuki, "Intercell-interference cancellation and neural network transmit power optimization for MIMO channels," in *Proc. IEEE VTC Fall*, Boston, Sep. 2015, pp. 1–5.

[147] J. Vieira, E. Leitinger, M. Sarajlic, X. Li, and F. Tufvesson, "Deep convolutional neural networks for massive MIMO fingerprint-based positioning," in *Proc. IEEE PIMRC*, Montreal, Oct. 2017, pp. 1–6.

[148] T. J. O'Shea, T. Erpek, and T. C. Clancy, "Deep learning based MIMO communications," *arXiv preprint arXiv:1707.07980*, 2017.

[149] M. Borgerding, P. Schniter, and S. Rangan, "AMP-inspired deep networks for sparse linear inverse problems," *IEEE Trans. Signal Process.*, vol. 65, no. 16, pp. 4293–4308, May 2017.

[150] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cognitive Commun. & Networking*, vol. 4, no. 3, pp. 433–445, May 2018.

[151] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognitive Commun. and Networking*, vol. 3, no. 4, pp. 563–575, Oct. 2017.

[152] S. Boyd and L. Vandenberghe, *Convex optimization*.   Cambridge University Press, 2004.

[153] Y. Zhang, H.-M. Wang, Q. Yang, and Z. Ding, "Secrecy sum rate maximization in non-orthogonal multiple access," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 930–933, Mar. 2016.

[154] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1426–1438, Aug. 2006.

[155] Z.-Q. Luo, "Applications of convex optimization in signal processing and digital communication," *Springer Math. Program.*, vol. 97, no. 1-2, pp. 177–207, Jul. 2003.

[156] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*.   SIAM, 1994, vol. 13.

[157] *Preparing for the future of artificial intelligence.* National Science and Technology Council Committee on Technology, Oct. 2016.

[158] J. Watt, R. Borhani, and A. K. Katsaggelos, *Machine learning refined: foundations, algorithms, and applications.* Cambridge University Press, 2016.

[159] A. Géron, *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems.* " O'Reilly Media, Inc.", 2017.

[160] R. S. Sutton, A. G. Barto *et al.*, *Introduction to reinforcement learning.* MIT press Cambridge, 1998, vol. 2, no. 4.

[161] J. Xiong, D. Ma, K.-K. Wong, and J. Wei, "Robust masked beamforming for MISO cognitive radio networks with unknown eavesdroppers," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 744–755, Feb. 2016.

[162] H. Ma, J. Cheng, and X. Wang, "Cooperative jamming aided robust beamforming for MISO channels with unknown eavesdroppers," in *IEEE GLOBECOM*, Singapore, Dec. 2017, pp. 1–5.

[163] Z. Zhu, Z. Chu, Z. Wang, and I. Lee, "Joint optimization of AN-aided beamforming and power splitting designs for MISO secrecy channel with SWIPT," in *Proc. IEEE ICC*, Kuala Lumpur, May 2016, pp. 1–6.

[164] H. Niu, D. Guo, Y. Huang, and B. Zhang, "Robust energy efficiency optimization for secure MIMO SWIPT systems with non-linear EH model," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2610–2613, Sep. 2017.

[165] Z.-Q. Luo, J. F. Sturm, and S. Zhang, "Multivariate nonnegative quadratic mappings," *SIAM J. Optim.*, vol. 14, no. 4, pp. 1140–1162, Jul. 2004.

[166] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.

[167] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Far-field multicast beamforming for uniform linear antenna arrays," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4916–4927, Oct. 2007.

[168] J. Guo and X. Zhu, "An improved analytical model for RF-DC conversion efficiency in microwave rectifiers," in *Proc IEEE MTT*, Montreal, Jun. 2012, pp. 1–3.

[169] Y. Pei, Y.-C. Liang, L. Zhang, K. C. Teh, and K. H. Li, "Secure communication over MISO cognitive radio channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, Apr. 2010.

[170] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[171] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer networks*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.

[172] K. Cumanan, Y. Rahulamathavan, and S. Lambotharan, "Linear transceiver design for multiuser spectrum sharing mimo networks," in *Proc. IEEE IWCR*. IEEE, Bangalore, Dec. 2010, pp. 1–4.

[173] Z. Xiong, K. Cumanan, Y. Rahulamathavan, and S. Lambotharan, "A robust sinr balancing technique for a multiuser MIMO cognitive radio network," in *Proc. IEEE SPAWC*. IEEE, Morocco, Jun. 2010, pp. 1–5.

[174] R. Krishna, K. Cumanan, Z. Xiong, and S. Lambotharan, "Cooperative relays for an underlay cognitive radio network," in *Proc. IEEE WCSP*. IEEE, Nanjing, Nov. 2009, pp. 1–4.

[175] K. Cumanan, R. Krishna, Z. Xiong, and S. Lambotharan, "Sinr balancing technique and its comparison to semidefinite programming based qos provision for cognitive radios," in *Proc. VTC Spring*. IEEE, Barcelona, Apr. 2009, pp. 1–5.

[176] I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, 2008.

[177] J. Mitola, "Cognitive radio. an integrated agent architecture for software defined radio." *Ph.D. Dissertation*, 2000.

[178] J. Mitola, G. Q. Maguire *et al.*, "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[179] K. Cumanan, R. Krishna, L. Musavian, and S. Lambotharan, "Joint beamforming and user maximization techniques for cognitive radio networks based on branch and bound method," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3082–3092, Oct. 2010.

[180] K. Cumanan, R. Krishna, Z. Xiong, and S. Lambotharan, "Multiuser spatial multiplexing techniques with constraints on interference temperature for cognitive radio networks," *IET Signal Process.*, vol. 4, no. 6, pp. 666–672, Dec. 2010.

[181] K. Cumanan, R. Krishna, V. Sharma, and S. Lambotharan, "Robust interference control techniques for multiuser cognitive radios using worst-case performance optimization," in *Proc. IEEE ACSSC*, Pacific Grove, Oct. 2008, pp. 378–382.

[182] Z. Xiong, C. Gong, L. Wu, K. Cumanan, and S. Lambotharan, "Capacity balancing for multiuser MIMO cognitive radio network," in *Proc. IEEE VTC Fall*. IEEE, San Franscisco, Sep. 2011, pp. 1–5.

[183] Y. Rahulamathavan, K. Cumanan, and S. Lambotharan, "An SINR balancing based beamforming technique for cognitive radio networks with mixed quality of service requirements," in *Proc. IEEE VTC Fall*. IEEE, San Franscisco, Sep. 2011, pp. 1–5.

[184] A. J. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective." *IEEE Proc.*, vol. 97, no. 5, pp. 894–914, 2009.

[185] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[186] D. W. K. Ng, E. S. Lo, and R. Schober, "Multiobjective resource allocation for secure communication in cognitive radio networks with wireless information and power transfer," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3166–3184, May 2016.

[187] E. Bedeer, O. A. Dobre, M. H. Ahmed, and K. E. Baddour, "A multiobjective optimization approach for optimal link adaptation of OFDM-based cognitive radio systems with imperfect spectrum sensing," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2339–2351, Feb. 2014.

[188] Y. Zeng, B. Clerckx, and R. Zhang, "Communications and signals design for wireless power transmission," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2264–2290, May 2017.

[189] F. Wang, J. Xu, X. Wang, and S. Cui, "Joint offloading and computing optimization in wireless powered mobile-edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1784–1797, Mar. 2017.

[190] E. Boshkovska, D. W. K. Ng, N. Zlatanov, A. Koelpin, and R. Schober, "Robust resource allocation for MIMO wireless powered communication networks based on a

non-linear EH model," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 1984–1999, May 2017.

[191] M. Stoopman, S. Keyrouz, H. Visser, K. Philips, and W. Serdijn, "A self-calibrating RF energy harvester generating 1V at- 26.3 dBm," in *Proc. IEEE VLSI Circuits*, Kyoto, Jun. 2013, pp. C226–C227.

[192] A. Mukherjee and A. L. Swindlehurst, "Detecting passive eavesdroppers in the MIMO wiretap channel," in *Proc. IEEE ICASSP*, Kyoto, Mar. 2012, pp. 2809–2812.

[193] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. SIAM, 2001, vol. 2.

[194] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Res. Logist. Quart.*, vol. 9, no. 3-4, pp. 181–186, Dec. 1962.

[195] C. Zhong, H. A. Suraweera, G. Zheng, I. Krikidis, and Z. Zhang, "Wireless information and power transfer with full duplex relaying," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3447–3461, Oct. 2014.

[196] H. J. Visser and R. J. Vullers, "RF energy harvesting and transport for wireless sensor network applications: Principles and requirements," *IEEE Proc.*, vol. 101, no. 6, pp. 1410–1423, Jun. 2013.

[197] T. Sun, X. Xie, and Z. Wang, *Wireless Power Transfer for Medical Microsystems*. Springer, 2013.

[198] L. Ni, X. Da, H. Hu, and M. Zhang, "Energy efficiency design for secure MISO cognitive radio network based on a nonlinear EH model," *Hindawi Math. Problems Eng.*, vol. 2018, May 2018.

[199] Z. Zhu, S. Huang, Z. Chu, F. Zhou, D. Zhang, and I. Lee, "Robust designs of beamforming and power splitting for distributed antenna systems with wireless energy harvesting," *IEEE Sys. J.*, vol. 13, no. 1, pp. 30–41, Mar. 2019.

[200] R. A. Horn, R. A. Horn, and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1990.

[201] Z. Chu, Z. Zhu, and J. Hussein, "Robust optimization for AN-aided transmission and power splitting for secure MISO SWIPT system," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1571–1574, Aug. 2016.

[202] D. J. Leinweber, "Stupid data miner tricks: overfitting the s&p 500," *J. Invest.*, vol. 16, no. 1, pp. 15–22, Spring 2007.

[203] D. Chicco, "Ten quick tips for machine learning in computational biology," *BioData mining*, vol. 10, no. 1, p. 35, Dec. 2017.

[204] Y. Bengio, F. Bastien, A. Bergeron, N. Boulanger-Lewandowski, T. Breuel, Y. Chherawala, M. Cisse, M. Côté, D. Erhan, J. Eustache, X. Glorot, X. Muller, S. P. Lebeuf, R. Pascanu, S. Rifai, F. Savard, and G. Sicard, "Deep learners benefit more from out-of-distribution examples," in *Proc. JMLR AISTATS*, Fort Lauderdale, Apr. 2011, pp. 164–172.

[205] F. Girosi, M. Jones, and T. Poggio, "Regularization theory and neural networks architectures," *Neural computation*, vol. 7, no. 2, pp. 219–269, Mar. 1995.

[206] L. Wang, M. D. Gordon, and J. Zhu, "Regularized least absolute deviations regression and an efficient algorithm for parameter tuning," in *Proc. IEEE ICDM*, Hong Kong, Dec. 2006, pp. 690–700.

[207] A. E. Hoerl and R. W. Kennard, "Ridge regression: biased estimation for nonorthogonal problems," *Technometrics*, vol. 42, no. 1, pp. 80–86, Feb. 2000.

[208] T.-H. Li, M. R. Khandaker, F. Tariq, K.-K. Wong, and R. T. Khan, "Learning the wireless V2I channels using deep neural networks," *arXiv preprint arXiv:1907.04831*, 2019.

[209] A. Hirose, *Complex-valued neural networks: Advances and applications.*   John Wiley & Sons, 2013, vol. 18.

[210] J. Tang, H. Tang, M. Cui, K. Cumanan, G. Chen, K. K. Wong, and J. A. Chambers, "Energy minimization in D2D-assisted cache-enabled internet of things: A deep reinforcement learning approach," *IEEE Trans. Industrial Informatics*, Nov.. 2019.

[211] J. Tang, H. Tang, K. Cumanan, S. Zhang, and Y. Zhou, "A reinforcement learning approach for D2D-assisted cache enabled hetnets," in *Proc. IEEE GLOBECOM*.   IEEE, Hawaii, USA, Dec. 2019.

[212] N. Jiang, Y. Deng, A. Nallanathan, and J. A. Chambers, "Reinforcement learning for real-time optimization in NB-IoT networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1424–1440, Jun. 2019.

[213] M. Chu, H. Li, X. Liao, and S. Cui, "Reinforcement learning-based multiaccess control and battery prediction with energy harvesting in IoT systems," *IEEE IoT J.*, vol. 6, no. 2, pp. 2009–2020, Apr. 2018.

[214] F. Shah-Mohammadi and A. Kwasinski, "Deep reinforcement learning approach to QoE-driven resource allocation for spectrum underlay in cognitive radio networks," in *Proc. IEEE ICC Workshop*, Kansas, May 2018, pp. 1–6.

[215] N. Zhao, Y.-C. Liang, D. Niyato, Y. Pei, and Y. Jiang, "Deep reinforcement learning for user association and resource allocation in heterogeneous networks," in *Proc. IEEE GLOBECOM*, Abu Dhabi, Dec. 2018, pp. 1–6.

[216] G. Alnwaimi, S. Vahid, and K. Moessner, "Dynamic heterogeneous learning games for opportunistic access in lte-based macro/femtocell deployments," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2294–2308, Apr. 2015.