Performance of Multi-antenna Wireless Systems with Channel Estimation Error

by

Nadia Jamal

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Wireless services and applications have become extremely popular and widely employed over the past decades. This, in turn, has led to a dramatic increase in the number of wireless users who demand reliable services with high data rates. But such services are very challenging to provide due to radio channel impairments including multipath fading and co-channel interference. In this regard, the use of multiple antennas in wireless systems was proposed recently which has rapidly received great attention. Multi-antenna technology is shown to have powerful capabilities to improve reliability via spatial diversity and to increase data rates via spatial multiplexing as compared with traditional single-antenna systems. Furthermore, by exploiting additional spatial dimensions, transmit beamforming techniques can be used to manage co-channel interference in such systems.

In a rich scattering environment, multiple antennas that are located sufficiently far apart at a transmitter experience independent fading with high probability. Therefore, the transmitter can send redundant versions of the same data stream over these independent channels to improve reliability. In particular, if the transmitter has access to perfect channel state information (CSI), it can set the beamforming weights such that the received signals from different transmit antennas combine constructively at some intended receiver(s) and destructively at some unintended receiver(s) so that no co-channel interference is generated.

Spatial multiplexing is another powerful multi-antenna transmission technique which aids in enhancing data rates without increasing bandwidth or transmit power. Multiple parallel and independent channels can be established between a transmitter and a receiver that both use multiple antennas in a rich scattering environment. Therefore, multiple independent streams of data can be simultaneously sent over these channels within the bandwidth of operation. This, in turn, enhances the data rate by a multiplicative factor equal to the number of the independent streams. Water-filling is a strategy that achieves the maximum data rate in such multiple-input multiple-output (MIMO) systems when perfect CSI is available at both the transmitter and the receiver.

In practice, CSI can be obtained at the receiver by the use of training sequences and its accuracy can be increased by carefully selecting sequences with good auto-correlation

properties. The transmitter can acquire CSI by using the channel reciprocity principle in wireless systems or by relying on a feedback path to convey the CSI from the receiver. Due to practical limitations such as rate-limited feedback links and the delay involved in such procedures, perfect CSI can be very challenging to obtain at the transmitter side. This motivates the need to evaluate the effect of imperfect CSI at the transmitter (CSIT) on the performance of transmit diversity and beamforming in multiple-input single-output (MISO) systems and water-filling power allocation in MIMO systems.

In this thesis, transmit diversity and beamforming are studied in a MISO system with an n-antenna transmitter, an intended single-antenna receiver, and some unintended single-antenna receivers. Two scenarios are considered, namely, null-steering beamforming and ϵ-threshold beamforming in which the allowable interference threshold at the unintended receivers is zero and $\epsilon > 0$, respectively. With perfect CSIT, null-steering beamforming can successfully nullify interference at m unintended receivers, where $m < n$, and achieve a nonzero received power at the intended receiver with a mean value that grows linearly with $n - m$ and is directly proportional to the power of the line-of-sight component between the transmitter and the intended receiver. With imperfect CSIT, null-steering beamforming based on erroneous channel estimates results in a nonzero interference at the unintended receivers with a mean value that is interestingly independent of n. Also, it is shown that a moderate line-of-sight component can significantly reduce the effect of estimation error on the performance of the intended link.

Intuitively, the allowance of a small nonzero interference at the unintended receivers, as in ϵ-threshold beamforming, should improve the received power at the intended receiver. The analysis in this thesis shows that this enhancement is marginal and not worthwhile, notably in the case of imperfect CSIT. Therefore, there is no significant loss in the performance of the intended link if the transmitter performs null-steering beamforming instead. In fact, the transmitter can employ additional antennas to improve the performance of the intended link without generating significant extra interference on the unintended receivers.

Furthermore, in this thesis, the effect of channel estimation error on the performance of water-filling power allocation in a MIMO system is explored when the transmitter and

the receiver both have n antennas. At low signal to noise ratios (SNR), the gap between water-filling throughput with perfect CSIT and the throughput corresponding to equal-power allocation with no CSIT is large asymptotically. It is thus interesting and worthwhile to evaluate how water-filling based on erroneous channel estimates may result in a throughput that falls between these two extremes. In this regard, it is first shown that, at low SNR, the normalized (by $1/n$) water-filling throughput with imperfect CSIT converges to a non-random value denoted by R, almost surely as n increases. Denoting C^{P} as the asymptotic normalized water-filling throughput with perfect CSIT and using it as a baseline for comparison, we then compare R with C^{P} and find that for moderate channel estimation errors, water-filling can still achieve significant normalized throughputs that are close to C^{P}. Furthermore, when the quality of channel estimation is very low, water-filling is shown asymptotically to achieve the same throughput as equal power allocation in the low SNR regime.

Acknowledgements

I would like to express my sincere appreciation to the people who made my stay at Waterloo a rewarding experience. I am very grateful to my PhD. supervisor, Professor Patrick Mitran, for his constant support, encouragement, and patience. I also wish to thank the members of my dissertation committee: Professors Stark Draper, Henry Wolkowicz, Oleg Michailovich, and Ravi Mazumdar for taking the time to read my thesis and for providing me with valuable comments and suggestions.

During the past four years, I have been in constant support of my husband and love of my life, Maziar Moradi, and my beloved parents, Hassan Jamal and Moloud Rayatparvar. Their invaluable and relentless support, encouragement, and love are without a doubt the most important reasons for my success.

Dedication

To my love, Maziar

and to my dear parents

Table of Contents

List of Tables

List of Figures

xiii

Table 1: Table for Notations

Notation	Meaning	Page first defined
n	number of transmit (transmit and receive) antennas in MISO (MIMO) system	16
P	total transmit power	16
ϵ	interference threshold at unintended receiver	16
m	number of unintended receivers	17
K	ratio of the line-of sight component's power to the scattering component's power	17
G^{null}	maximum received power at the intended receiver in null-steering beamforming	17
$G^{\text{null, Im}}$	actual received power at the intended receiver in null-steering beamforming	18
$I^{\text{null, Im}}$	actual interference power at the unintended receivers in null-steering beamforming	18
G^{th}	maximum received power at the intended receiver in ϵ-threshold beamforming	19
$G^{\text{th, Im}}$	actual received power at the intended receiver in ϵ-threshold beamforming	19
$I^{\text{th, Im}}$	actual interference power at the unintended receiver in ϵ-threshold beamforming	19
I^{N}	asymptotic normalized throughput of equal power allocation	20
C^{P}	asymptotic normalized water-filling throughput with perfect CSIT	21
R	asymptotic normalized water-filling throughput with imperfect CSIT	22
γ	signal to estimation error ratio (SER)	36

Chapter 1

Introduction

1.1 Wireless Communication Systems

Wireless technology is undoubtedly one of the significant revolutionary advancements in recent decades as it has enabled the transfer of information over both short and long distances without the use of wires (see Fig. 1.1). Long-range wireless services such as satellite television, satellite radio, and satellite telephony in remote areas are only possible using wireless technology. Mobile communications is another increasingly popular service that wireless technology can facilitate. One of the many examples of mobile communications is the well-known and widely-employed cellular system with more than 6.8 billion subscribers worldwide in 2013 [1]. Remote measurement, wireless sensing, traffic control, cordless telephony, mobile satellite communications, and WiFi are only a few examples of the services that owe their existence to wireless communication technology.

Wireless systems and services have strict requirements in terms of data rate and reliability. Data rate is measured by the number of bits transferred via the communication link per unit of time and is generally desired to be as high as possible. Reliability, on the other hand, is a parameter showing how reliable a communication link is in maintaining a connection or providing a satisfactory quality of service. It is often quantified by the frame error rate (the ratio of the number of corrupted data frames over the total num-
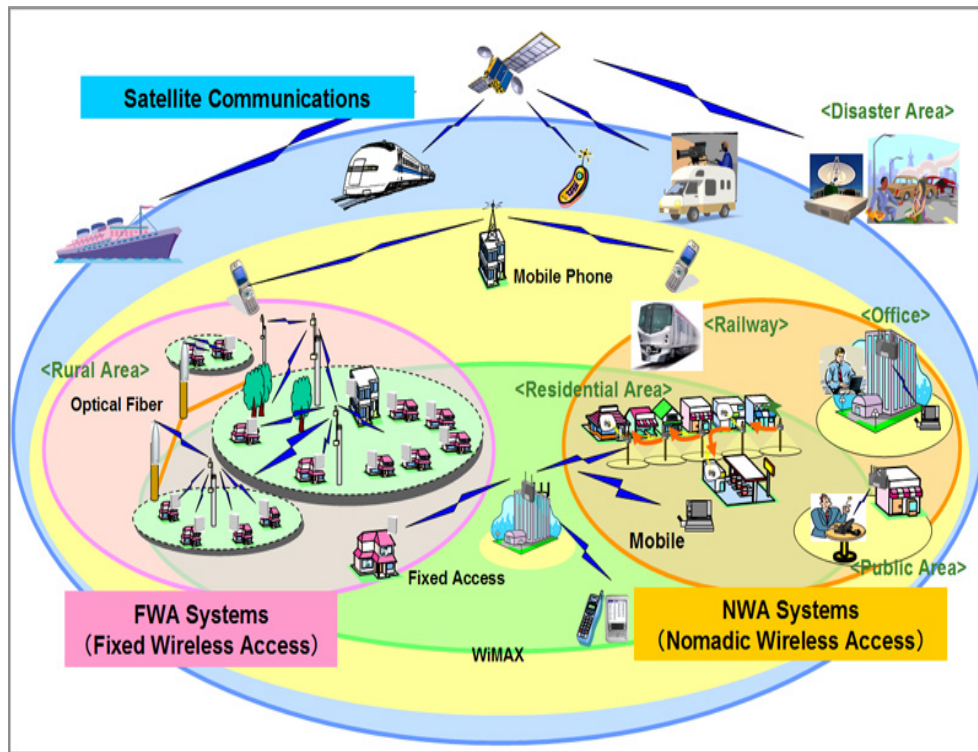
Figure 1.1: Long-range and short-range wireless communication services; courtesy of [2]

ber of transmitted data frames). It is known that there is a tradeoff between reliability and data rate, i.e., increasing the data rate leads to higher frame error rate (less reliable service) and decreasing the data rate can result in a more reliable service. Since the transmission medium for wireless communications is the radio channel between the transmitter and the receiver, impairments such as small-scale and large-scale propagation effects along with variations of the channel with time make it very challenging for wireless systems to provide reliable and high-data-rate services.

Multipath propagation is a small-scale propagation effect in wireless channels that occurs when the transmitted signal reaches the receiver via multiple different propagation paths. This phenomenon is shown in Fig 1.2 for three paths only. The direct path between the transmitter and the receiver is known as the line-of-sight path. The indirect paths
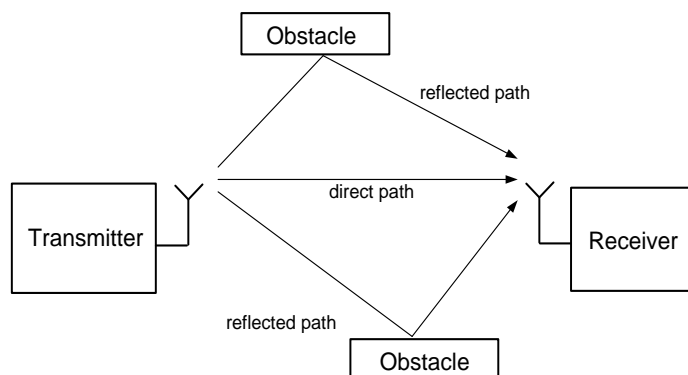
Figure 1.2: Multipath propagation

are caused, e.g., by reflections of the radio waves from surrounding obstacles such as buildings, cars, and mountains. The multipath components have distinct propagation delays, amplitudes, and phase shifts with respect to each other. Such components add up at the receiver destructively (which causes fading) or constructively, and thus generate a combined signal that is shown to have a Rayleigh distributed magnitude and a uniformly distributed phase (this statistical model is confirmed by practical measurements as well [3, 4]). In the case that a strong line-of-sight component exists between the transmitter and the receiver, a Rician distribution is a more accurate model for the received signal's magnitude. Multipath fading may lead to severe degradation in the performance of wireless communication systems. For example, if the channel between a transmitter and a receiver is in a deep fade, i.e., multipath components add up destructively at the receiver, the received signal's power significantly drops which results in lower signal to noise ratios (SNR) and, in turn, higher frame error rates.

Furthermore, due to the broadcast nature of the communication medium, wireless communication systems that simultaneously operate in the same frequency band and are used in the same geographical region, may impose co-channel interference on one another. To avoid such performance-degrading interference, distinct frequency bands (out of the limited

frequency spectrum) are allocated to different wireless systems as regulated by the International Telecommunications Union (ITU). Obviously, designing wireless systems that satisfy the rapidly growing demand for reliable and high-data-rate communications becomes more challenging when the system can only operate in such scarce frequency bands with limited transmit power.

1.2   Multi-antenna Wireless Systems

The use of multiple antennas in wireless systems has rapidly received considerable attention over the past decades in both academia and industry. Multi-antenna technology is shown to have powerful capabilities to improve reliability via spatial diversity and to increase data rates via spatial multiplexing as compared with traditional single-antenna systems. Furthermore, by exploiting additional spatial dimensions, transmit beamforming techniques can be performed to manage co-channel interference in wireless systems. Standards such as IEEE 802.11n WiFi, Long Term Evolution (LTE), and LTE-Advanced are examples of wireless systems that exploit multiple antennas [3].

Multiple antennas can be used at the transmitter, receiver, or both. Different antenna configurations are single-input single-output, multiple-input single-output (MISO), single-input multiple-output (SIMO), and multiple-input multiple-output (MIMO). Single-input single-output (SISO) is the traditional single-antenna system, while MISO has multiple transmit antennas and a single receive antenna, SIMO has a single transmit antenna and multiple receive antennas, and finally MIMO has multiple transmit antennas and multiple receive antennas.

1.2.1   Spatial Diversity

An efficient technique to overcome the negative effects of multipath fading and enhance the reliability of wireless channels is to exploit spatial diversity by employing multiple antennas at the transmitter, receiver, or both [3–5]. If the antennas are located sufficiently far

apart and the environment is rich scattering, it is very likely that the antennas experience independent fading and, in turn, the resultant channels have a low probability of being in a deep fade simultaneously. Thus, redundant versions of the same data stream can be sent/received over these multiple independently faded channels (or diversity branches) and can be combined appropriately to increase the received signal's power, and, in turn, decrease the frame error rate compared to a traditional single-antenna system.

Spatial diversity can be implemented as receiver diversity, transmit diversity, or both. Receiver diversity is applicable in SIMO systems [6], while transmit diversity can be used in MISO systems [7, 8]. For receiver diversity, various combining strategies are proposed such as equal-gain combining, selection combining, and maximal-ratio combining [8]. For instance, in maximal-ratio combining, the signal received in a diversity branch is rotated and weighted appropriately according to its corresponding channel gain. This process is performed so that all the received signals are combined coherently at the receiver and the diversity branches with stronger received signals are further amplified. Clearly, in order for this combining technique to work properly, perfect knowledge of the channel gains is required at the receiver.

Due to cost and space considerations, it is not feasible in practice to use multiple antennas at the receiver terminals. Thus, analyzing transmit diversity in MISO systems has become more practical and popular. In MISO systems, transmit diversity can be exploited with or without channel gain knowledge at the transmitter. In 2×1 MISO systems (2 antennas at the transmitter and a single antenna at the receiver) for example, Alamouti's scheme can be used to exploit transmit diversity over the space and time when the channel gains are unknown to the transmitter [9]. For a general $n \times 1$ MISO channel, space-time block codes and space-time trellis codes can be used at the transmitter to achieve full diversity [10, 11]. With perfect channel gain knowledge at the transmitter, transmit maximal-ratio combining can be employed, i.e., before transmission, the signal at each antenna is weighted appropriately according to its corresponding channel gain so that the received signal power is maximized [12, 13]. In MIMO systems, spatial diversity is utilized by combining the transmit and receiver diversity schemes mentioned above. In [14], a thorough overview of the benefits of spatial diversity in wireless systems is provided.

1.2.2 Interference Management via Beamforming

As stated earlier, co-channel interference is a major limiting factor in the performance of wireless systems. Using additional spatial dimensions offered in MISO wireless systems, the transmitter can perform an interference management technique known as beamforming to effectively reduce its co-channel interference to the unintended receivers operating in the same frequency bands in its vicinity. Such an interference management technique allows wireless systems to reuse frequency bands to increase spectral efficiency and thus provide higher data rates.

More specifically, taking advantage of the channel fading effect, the transmitter can set the beamforming weights such that the received signals from different transmit antennas combine destructively at the unintended receivers and constructively at the intended receiver. When beamforming is performed, the transmitter is required to know the exact channel gains in order to find the best beamforming vector that achieves the maximum signal power at the intended receiver and limits the interference at the unintended receivers.

1.2.3 Spatial Multiplexing

Spatial multiplexing is a powerful transmission technique that can be used in MIMO systems. This technique tremendously aids in enhancing data rates of wireless systems without increasing bandwidth or transmit power [15, 16].

In traditional single-antenna wireless systems, only one stream of data can be transferred at a time within the bandwidth of operation. If the transmitter and the receiver both use multiple antennas and the environment is rich scattering, multiple parallel and independent channels can be established between them. Therefore, multiple streams of data can be simultaneously sent over these independent channels within the bandwidth of operation. This, in turn, enhances the data rate by a multiplicative factor equal to the number of independent streams, which is equal to the minimum of the number of transmit antennas and the number of receive antennas.

Supposing that channel knowledge is only available at the receiver, the independent

streams can be separated at the receiver by means of interference cancellation algorithms such as maximum likelihood (ML) detection, linear zero-forcing (ZF) detection, minimum mean squared error (MMSE) detection, and successive interference cancellation (SIC) which were originally proposed in [15] for the BLAST (Bell-labs Layered Space Time Architecture) scheme. Even though the ML detector provides the best possible performance, its complexity is high. The linear techniques such as ZF detection and MMSE detection sacrifice some performance to reduce complexity. On the other hand, integer-forcing (IF) linear receivers [17, 18], offer closer performance to ML detection with a slightly more complex architecture compared to other linear techniques such as MMSE detection.

If the transmitter also has perfect knowledge of channel gains too, water-filling power allocation can be performed at the transmitter to achieve the maximum data rate [16, 19]. In this case, knowing the channel gains, the transmitter allocates power to each individual stream of data (eigen-direction) proportionally to its corresponding eigenvalue. In other words, more power is allocated to stronger eigen-directions while less power is dedicated to weaker eigen-directions.

As stated earlier, in MIMO systems spatial diversity schemes can be performed which provide diversity gain and in turn, achieve higher reliability. In addition, spatial multiplexing techniques can be employed which provide spatial multiplexing gain and achieve better data rates. In order to improve both data rate and reliability at the same time, spatial multiplexing and spatial diversity techniques can be performed together. However, there is a fundamental tradeoff between the diversity gain and the multiplexing gain (diversity-multiplexing tradeoff) because higher spatial multiplexing gain comes at the price of sacrificing diversity gain and vice versa [20–22].

1.3   Importance of Channel State Information (CSI)

In wireless communication systems, it is commonly required for the receiver to have very good knowledge of channel gains. In traditional single-antenna systems, channel gain knowledge is exploited at the receiver to reverse the effect of multipath fading before

performing data detection. In multi-antenna systems on the other hand, as discussed in previous sections, perfect knowledge of channel gains is crucial at the receiver to implement spatial diversity combining techniques such as in SIMO systems, and spatial multiplexing techniques in MIMO systems [23–25].

A popular approach to obtain channel gain knowledge at the receiver is to use training signals (pilot sequences) emitted by the transmitter [5]. Training signals are known by the receiver in advance. Therefore, channel gains can be estimated using the combined knowledge of the transmitted and the received signals. In order to have more accurate channel estimates, pilot sequences need to have good auto-correlation properties and their duration and location should be carefully selected [26,27]. Blind and semi-blind techniques can also be used for channel estimation where the former does not exploit training signals and performs estimation blindly with respect to the channel and noise characteristics, whereas the latter combines the use of training signals and blind based techniques [28,29].

In single-antenna systems, knowledge of channel gains at the transmitter can assist the transmitter to perform transmission rate control, power control, etc. In MISO and MIMO systems, in addition to these benefits, such knowledge can be used for spatial diversity combining, beamforming, and spatial multiplexing as discussed earlier. For example, when beamforming is performed, the transmitter is required to know the exact channel gains in order to find the best beamforming vector that achieves the maximum signal power at the intended receiver and limits the interference at the unintended receivers. Also note that, although much work on space-time coding requires no channel gain knowledge at the transmitter, if the transmitter has such knowledge, better throughputs can be achieved [30].

In order to estimate channel gains at the transmitter, generally, two main techniques can be used [5]. The first technique takes advantage of the channel reciprocity of the wireless medium in Time Division Duplex (TDD) transmission. In TDD, typically the same frequency band and antennas are used for forward and reverse links in different time slots. Techniques such as using pilot sequences can be applied to estimate the channel gain by the transmitter in the reverse link. If the duplexing time delay is not small enough, the estimated channel gains in the forward and reverse links are not equal as the channel gain

changes with time.

The second technique relies on a feedback link from the receiver to the transmitter to convey the channel estimates to the transmitter. This technique can be carried out when reciprocity principle cannot be used, as in Frequency Division Duplexing (FDD) transmission [5]. However, in practice, there are some obstacles in obtaining perfect channel gains at the transmitter in such cases. For example, explicitly feeding back the channel state information (CSI) is rate-limited. Since the CSI needs to be quantized at the receiver and sent to the transmitter over a limited-rate feedback link, the CSI conveyed to the transmitter is not perfect anymore [31–33]. This is notably the case in MIMO systems where the feedback requirements grow with the product of the transmit antennas and the receive antennas, and thus as the number of antennas increase, the CSI obtained at the transmitter through limited-rate feedback gets more inaccurate [34, 35]. Furthermore, using feedback for conveying channel estimates involves some delay as well, which makes the estimates outdated especially in relatively fast time-varying channels. In such channels, more frequent feedback is required which will introduce overhead that can be prohibitive [36].

Note that explicit feedback links can be expensive and there are open-loop wireless systems that do not provide any protocol means for the receiver to convey CSI to the transmitter via a feedback link [37, 38]. In such systems, an ARQ (Automatic Repeat reQuest) feedback scheme is used for transmission rate control and by using some pilot sequences from the ARQ feedback scheme, the CSI can be estimated at the transmitter as well. Since these schemes are not designed for highly accurate CSI estimation, the CSI estimates are imperfect in such cases.

In this thesis, assuming that the channel gains are perfectly known at the receiver, first, the performance of spatial diversity, beamforming, and spatial multiplexing techniques is studied in multi-antenna systems with perfect CSI at the transmitter (perfect CSIT). Then, we address the case that the channel gains are imperfectly known at the transmitter (imperfect CSIT) by modelling the channel estimation error to be Gaussian distributed (which is a reasonable assumption in estimation methods such as maximum likelihood

(ML) estimation [39]). This model allows us to obtain statistical closed-form results and find how the performance is affected due to the lack of perfect CSIT.

1.4   Applications of Multi-antenna Technology

In wireless communications, co-channel interference may be imposed from one wireless system to another, such as in cognitive radio systems [40, 41], or from a wireless user to another user in the same wireless system, such as in heterogeneous networks [42, 43]. Such co-existence in the same frequency band is sometimes intentionally allowed and is aimed to increase spectral efficiency. Therefore, the use of interference management techniques (e.g. beamforming) is essential to handle the interference among such wireless users and systems.

1.4.1   Beamforming in Cognitive Radio Systems

As stated earlier, in order to manage the co-channel interference caused by simultaneous operation of wireless systems in the same geographical region, traditionally, distinct frequency bands have been allocated to different wireless systems. But, according to radio spectrum occupancy measurements, such a fixed spectrum allocation policy has failed to accommodate wireless systems in an efficient manner, and has led to an under-utilization of frequency bands both temporally and geographically [44, 45]. In other words, monitoring of the frequency spectrum has revealed that some frequency bands are not occupied at all times and thus are not utilized efficiently.

Motivated by these measurements and observations, there has been recent interest in finding a new communication paradigm that allows different wireless systems to operate in the same frequency band (as opposed to the current fixed spectrum allocation policy). This spectrum sharing approach has the potential to increase spectral efficiency and enables optimal accommodation of wireless services in the limited frequency spectrum. In this regard, the idea of cognitive radio systems was introduced [46], that has evoked much

interest in frequency regulatory bodies to provide new spectrum allocation policies in order to support cognitive radio [44, 47].

In the literature, cognitive radio frameworks mainly consist of a secondary system that can use the licensed frequency bands originally allocated to a primary system in its vicinity (see Fig. 1.3) [40, 48–50]. This helps the secondary system to achieve better performance and increase spectral efficiency especially when the primary system is inactive. The co-channel interference resulting from the secondary system's activity must be controlled such that the primary system's performance does not degrade severely.

A widely recognized approach to enhance secondary system's performance in cognitive radio systems is to exploit spatial diversity by using multiple antennas at the secondary transmitter [51–57]. If the secondary receiver has a single antenna, this leads to a MISO system. In such systems, with a secondary transmit power constraint, the transmit co-variance matrix at the secondary transmitter can be chosen appropriately to satisfy the primary interference-power constraints while maximizing the secondary received power. As shown in [51–53], the optimal covariance matrix in a MISO system is rank-one which implies that beamforming is optimal. Therefore, taking advantage of channel fading, the secondary transmitter should set the beamforming weights such that the received signals from different transmit antennas combine destructively at the primary receiver and constructively at the secondary receiver.

Beamforming has been broadly investigated in the literature as a technique for inter-ference management in MISO cognitive radio frameworks. Under perfect channel state information at the transmitter (CSIT), [53, 54] derive optimal solutions to their associated optimization problems and subsequently evaluate the performance of beamforming, whereas [55] takes a numerical approach to the same problem. In the realistic case of imperfect CSIT, [56, 59] consider probabilistic interference constraints for the interfered system and subsequently study robust beamforming numerically. In [60, 61], even though the actual channel gains are unknown at the beamformer, the knowledge of some uncertainty regions containing the actual gains is assumed to be available. Under such assumptions, [60] solves the associated problem numerically while [61] solves it analytically. In this thesis,
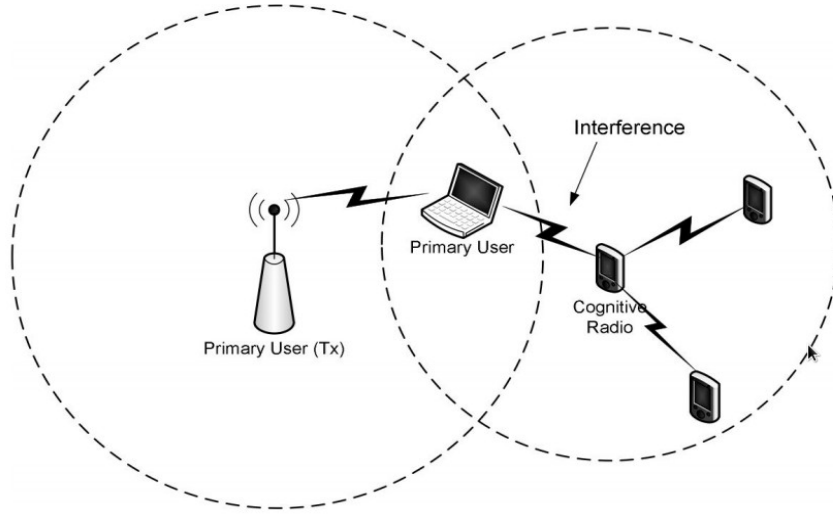
Figure 1.3: Cognitive radio framework; courtesy of [58]

beamforming and its performance analysis with imperfect CSIT is studied but a different approach is taken compared to the previous works. More specifically, in this thesis, an optimal beamforming vector is found merely based on the estimates of the channel gains and unlike [60, 61], it is not assumed that uncertainty regions containing the actual gains are known at the beamformer.

1.4.2   Beamforming in Heterogeneous Networks

Current cellular systems are not only being used for transmission of voice but also have become a significant means for accessing the Internet anytime and anywhere. This, in turn, has led to the generation of high volumes of data traffic. Clearly, seamless operation of such applications requires an underlying infrastructure that supports reliable communications with high data rates.

Large tower-mounted macro base stations (BS) are the current cellular system's main component that serve as the gateways connecting mobile users to the core network. These

BSs are roughly located based on carefully planned layouts and are almost identical in terms of transmit power, antenna configurations, etc, creating a homogeneous network.

To decrease the total number of macro BSs that cover a geographical region, which, in turn, decreases the total cost of BS installation, maintenance, etc, the coverage area (cell-site) of each BS must be expanded. This leads to a greater distance between mobile users and the BSs and thus increases large-scale propagation losses. In addition, with a bigger cell-site, the data rate offered for each mobile user is less because the same bandwidth is shared among a larger number of users. Therefore, macro BSs with a large cell-size do not effectively support high demands of traffic in dense areas with more users.

The idea of heterogeneous networks has emerged to overcome the aforementioned issue [42, 43, 62–64]. It simply suggests the deployment of additional BSs with smaller cell-sizes (based on their range and power level are referred to as micro, pico, and femto stations) in dense areas that overlap in coverage with a traditional macro BS (MBS) as shown in Fig. 1.4. This brings users closer to their serving BSs and thus, intuitively, improves coverage and overall system performance. But since the smaller BSs operate in the same frequency band as the macro BSs, such performance improvement can be achieved only if the excessive co-channel interference is successfully managed using techniques such as beamforming.

1.4.3   Spatial Diversity in Massive MIMO systems

Massive MIMO, which has received considerable attention recently in academia and industry, is an emerging technology that takes advantage of having a large number of antennas at the transmitter and/or the receiver [66–68]. Massive MIMO can be enabled for example by employing small-sized antennas when using mmWave frequency bands for radio access in wireless systems [69]. Having such a large number of antennas potentially allows for orders of magnitude improvement in data rates especially if suitable spatial diversity techniques are performed.

Water-filling power allocation can provide throughputs that scale with the number of antennas assuming perfect CSI is available at the transmitter and the receiver side. How-
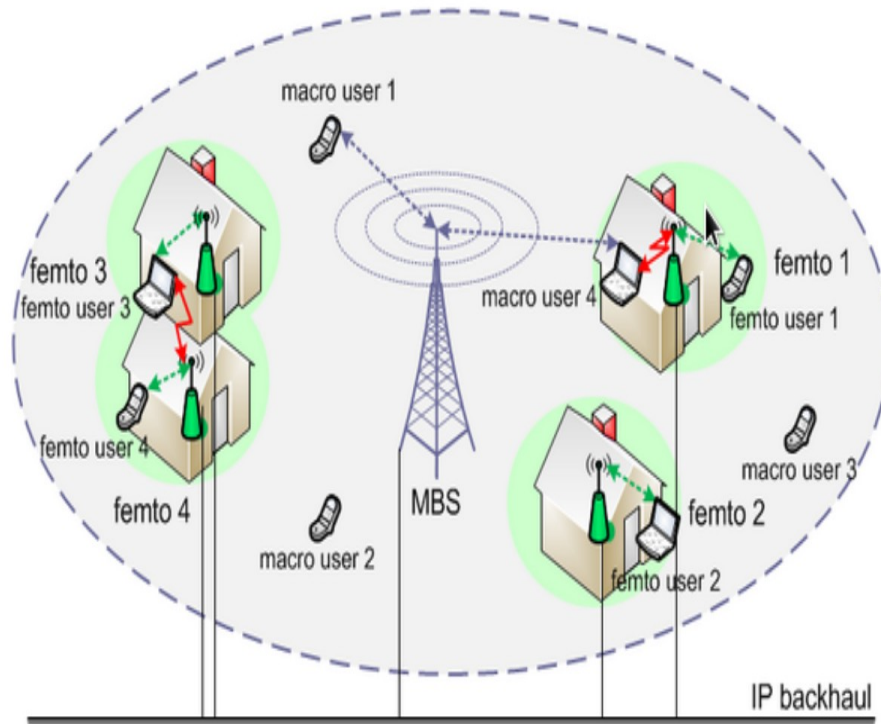
Figure 1.4: Heterogeneous network consisting of macro BSs and femto BSs; courtesy of [65]

ever, in practice, when the number of antennas is large, it becomes extremely challenging to obtain accurate CSI especially at the transmitter. Thus, it is worthwhile to evaluate how water-filling performs in such systems when channel estimation error is present, and obtain the asymptotic water-filling throughput as a function of the power of the estimation error.

1.5 System Model and Summary of the Results

1.5.1 Performance of Beamforming with Channel Estimation Error

In Chapter 2, spatial diversity and transmit beamforming are studied in a MISO wireless communication system. In the MISO model considered in Chapter 2, it is assumed that a transmitter with n antennas aims to communicate with a single-antenna receiver (the intended receiver). Since the transmitter has multiple antennas that are separated far enough and the environment is rich scattering, it can exploit spatial diversity to increase reliability and send a single stream of data over multiple antennas with different beamforming weights. The beamforming weights can be set appropriately based on the channel gains so that, with limited transmit power P, the received power at the intended receiver is maximized (transmit maximal-ratio combining). In such a scenario, the maximum received power is shown to grow linearly with the number of transmit antennas in the case of perfect CSIT.

The main contribution of Chapter 2 is to study transmit beamforming when m single-antenna unintended receivers exist in the vicinity of the transmitter and the focus is to limit the co-channel interference on such receivers as well as maximizing the received power at the intended receiver. In this regard, two different scenarios of null-steering beamforming and ϵ-threshold beamforming are considered in sections 2.3 and 2.4, respectively. Null-steering beamforming simply refers to the scenario in which the unintended receivers tolerate no interference from the transmitter, i.e., the interference at the unintended receivers is constrained to be 0. In ϵ-threshold beamforming, a small amount of interference (ϵ) is allowed at the unintended receivers.

In each beamforming scenario, two separate cases of perfect CSIT and imperfect CSIT are investigated. In each case, an optimization problem is obtained that maximizes the received power at the intended receiver while satisfying the interference constraints and the transmit power constraint. Unlike most related prior works which solve the associated optimization problems numerically using suitable optimization algorithms and only pro-

vide numerical analysis, in this thesis, a closed-form expression for the maximum received power is derived which can then be statistically analyzed.

Null-steering Beamforming: In the null-steering beamforming scenario, Rician fading with a general K-factor (ratio of the line-of sight component's power to the scattering component's power) is assumed to model the channel gains between the transmitter and its intended receiver and Rayleigh fading is assumed to model the channel gains between the transmitter and the unintended receivers. First, the case of perfect CSIT is examined, and with limited total transmit power, a suitable beamforming vector is found that attains maximum received power at the intended receiver while nullifying the interference power at $m < n$ unintended receivers. Then the corresponding maximum received power at the intended receiver, G^{null} is derived.

The results in the case of perfect CSIT imply that:

- When the number of transmit antennas is greater than the number of unintended receivers $(n > m)$, the transmitter can successfully perform null-steering beamforming and achieve a nonzero value for the mean received power $\mathsf{E}[G^{\mathrm{null}}]$ which grows linearly with $(n - m)$. Intuitively, in such a case, m degrees of freedom are used for nullifying the interference at m unintended receivers while the remaining $n - m$ degrees of freedom are used towards spatial diversity in the intended link. Therefore, either a larger number of antennas or smaller number of unintended receivers leads to higher values of $\mathsf{E}[G^{\mathrm{null}}]$.

- The mean received power $\mathsf{E}[G^{\mathrm{null}}]$ is directly proportional to the power of the line-of-sight component represented by the K-factor. Thus, a better line-of-sight component in the intended link provides a better performance.

- After finding the variance of G^{null}, it is shown that G^{null} converges in probability to $\mathsf{E}[G^{\mathrm{null}}]$ as n increases. This implies that the stochastic average $\mathsf{E}[G^{\mathrm{null}}]$ becomes a more precise indicator of the realization of G^{null}, as the number of antennas increases.

To obtain a more realistic framework, in the next approach, imperfect CSIT is consid-

ered with Gaussian-modeled estimation errors. Consequently, the transmitter is assumed to ignore or be ignorant of the existence of estimation errors. Therefore, beamforming weights are computed using the channel estimates only. Subsequently, the actual received power at the intended receiver, $G^{\text{null, Im}}$, and the actual interference power at the unintended receivers $I^{\text{null, Im}}$, are derived.

In the case of imperfect CSIT, it is found that:

- The mean received power at the intended receiver, $\mathsf{E}[G^{\text{null, Im}}]$, is upper bounded by $\mathsf{E}[G^{\text{null}}]$ obtained in the perfect CSIT case because of the detrimental effect of channel estimation error. However, it still grows with $(n - m)$ and a moderate K-factor can significantly reduce the effect of estimation error on the performance.

- Since null-steering beamforming is performed based on channel estimates which are imperfect, the interference at the unintended receivers is not zero anymore. However, it is found that the mean interference power at the ith unintended receiver, $\mathsf{E}[I_i^{\text{null, Im}}]$, $i = 1, \ldots, m$, is independent of the number of antennas n and, in turn, it does not grow with n. As a result, more antennas can be used to enhance the performance of the intended link (due to spatial diversity) without increasing the average interference power at the unintended receivers.

ϵ-threshold Beamforming: In the ϵ-threshold beamforming scenario, the same channel gain statistics are assumed except that for tractability, we take $K = 0$. Also, for simplicity in finding a closed form expression for the received power, we consider the special case that $m = 1$. Note that for $m > 1$, the corresponding optimization problem can be formulated as a quadratically constrained quadratic program (QCQP) which in general does not have a closed-form solution. However, the problem can be numerically solved as shown in [59], using techniques such as second order cone programming (SOCP), or semi definite programming (SDP).

In this scenario, similarly to the null-steering scenario, we first examine the case of perfect CSIT and find a suitable beamforming vector that, with a limited transmit power (P), attains maximum received power at the intended receiver while satisfying the interference

power constraint at the unintended receiver (i.e. the interference power must be less than the threshold ϵ (where $\epsilon = \alpha P$ for $\alpha \geq 0$). After showing the existence of such a beamforming vector, its corresponding received power at the intended receiver, G^{th}, is found. In the case of imperfect CSIT, beamforming weights are computed using the channel estimates only. Subsequently, the actual received power at the intended receiver, $G^{\text{th, Im}}$, and the actual interference power at the unintended receiver $I^{\text{th, Im}}$, are derived.

The results indicate that:

- The mean received power with perfect CSIT, $\mathsf{E}[G^{\text{th}}]$, consists of a term independent of α (or the interference threshold ϵ). This term is identical to $\mathsf{E}[G^{\text{null}}]$ which is the corresponding null-steering result with $K = 0$ and $m = 1$, and it grows linearly with $(n-1)$. Therefore, employing more transmit antennas results in higher $\mathsf{E}[G^{\text{th}}]$.

- The mean of actual received power with imperfect CSIT, $\mathsf{E}[G^{\text{th, Im}}]$, consists of a term independent of α as well. This term is identical to $\mathsf{E}[G^{\text{null, Im}}]$ which is the corresponding null-steering result with $K = 0$ and $m = 1$, and it grows linearly with $(n-1)$.

- The expressions for $\mathsf{E}[G^{\text{th}}]$ and $\mathsf{E}[G^{\text{th, Im}}]$ have additional terms dependent on α with a dominant term that grows as $\sqrt{\alpha}$. This allows to obtain the following tradeoffs:

  - Under perfect CSIT, a small increase in α can result in a moderate improvement in $\mathsf{E}[G^{\text{th}}]$ when the number of antennas, n, is small. Particularly, increasing α from zero to 0.1 leads to an increase of 31% for $\mathsf{E}[G^{\text{th}}]$ when $n = 2$. However, the amount of increase reduces as n gets larger.

  - Under imperfect CSIT, the enhancement in $\mathsf{E}[G^{\text{th, Im}}]$ is less compared to the perfect CSIT case.

- Under imperfect CSIT, the dominant parts of the mean actual interference power, $\mathsf{E}[I^{\text{th, Im}}]$, are independent of n. This implies that while employing additional transmit antennas can benefit the intended link by increasing $\mathsf{E}[G^{\text{th, Im}}]$, it results in little extra interference on the unintended receiver.

18

- Since the interfered system's allowance of a small nonzero interference threshold results in no significant enhancement to $\mathsf{E}[G^{\text{th, Im}}]$ compared to $\mathsf{E}[G^{\text{null, Im}}]$, there is no significant loss in performance to the intended link if the transmitter performs null-steering beamforming instead of ϵ-threshold beamforming.

The results of this chapter have been published in [70] and [71].

1.5.2 Performance of Water-filling with Channel Estimation Error

In Chapter 3, we study spatial multiplexing in a MIMO wireless communication model that has n antennas at the transmitter and n antennas at the receiver. The antennas are sufficiently far apart at the transmitter and the receiver and the environment is assumed to be rich scattering. Therefore, the corresponding $n \times n$ channel gain matrix between the transmitter and the receiver has i.i.d. entries. By applying singular value decomposition to the channel gain matrix, the MIMO system can be converted to multiple independent and parallel SISO channels, over which multiple independent streams of data can be sent. The transmit power can thus be split between these streams based on the knowledge of the channel gain matrix at the transmitter.

We consider three different cases of i) no CSIT, ii) perfect CSIT, and iii) imperfect CSIT and find the throughput of the MIMO system in each of these three cases. We then perform asymptotic analysis to derive the corresponding normalized (by $1/n$) throughput as n grows large.

In the case of no CSIT, the transmitter has no knowledge of the channel gain matrix. Therefore, the transmitter distributes the limited transmit power equally among the n transmit antennas as this is the most reasonable way to allocate power in such a case [19]. In other words, the same amount of power is allocated to all eigen-directions regardless of how strong their eigenvalues are (equal power allocation). Doing such a power allocation, it is shown in [19] that, even though the channel gain matrix is random, the normalized (by $1/n$) throughput converges to a non-random value denoted by I^{N}, almost surely as n

increases. This result implies that the throughput scales linearly with n as n grows large which is, intuitively, due to spatial multiplexing.

In the case of perfect CSIT, knowing the eigenvalues of the channel, the transmitter can optimally distribute power among antennas according to water-filling power allocation and thus achieve the maximum throughput (the capacity of the MIMO system). As found in [19], in this case too, the normalized (by $1/n$) throughput converges to a non-random value denoted by, C^{P}, almost surely as n increases.

At high SNR, as shown in [19], C^{P} and I^{N} are asymptotically equal. This implies that there is no loss in the throughput if power is distributed equally among antennas (which requires no CSIT) instead of performing water-filling power allocation. As the SNR decreases, the gap between C^{P} and I^{N} increases and water-filling power allocation results in better throughputs compared to equal power allocation. At low SNR, as shown in [19], $C^{\mathrm{P}} \approx 4I^{\mathrm{N}}$. Thus, at low SNR, it is crucially important how the transmitter allocates power among antennas.

The contribution of Chapter 3 is to study the imperfect CSIT case at low SNR, and find the throughput when the transmitter performs water-filling power allocation based on channel estimates that are imperfect.

Note that, in the high SNR regime, as stated earlier, equal power allocation (which requires no CSIT) provides the same asymptotic throughput as water-filling with perfect CSIT. Hence, in that regime, it is preferable to employ equal power allocation instead of water-filling and thus studying the effect of channel estimation error on the performance of water-filling is of no interest.

We measure the quality of channel estimation error by a parameter called the signal to estimation error ratio (SER), i.e., the ratio of the scattering component's power to the power of the estimation error. High SER corresponds to good channel estimates at the transmitter, whereas low SER corresponds to bad estimates.

The results indicate that:

- Even though the channel gain matrix is random, at low SNR, the normalized water-

filling throughput based on imperfect CSIT converges to a non-random value (asymptotic growth rate) denoted by R, almost surely as n increases.

- Since the gap between C^{P} and I^{N} is high at low SNR, different SER values lead to different values of R that fall between the best-case result C^{P}, and the worst-case result I^{N}. Using the asymptotic growth rate for the capacity with perfect CSIT, C^{P}, as a baseline for comparison, we thus can compare R with C^{P} as a function of the SER.

- At low SNR, for moderate values of the SER, water-filling based on erroneous channel estimates can still asymptotically achieve significant throughputs. In particular, for SER values such as 5 dB, 0 dB, and -5 dB, R is found to be 86%, 70%, and 52% of C^{P}, respectively.

- At low SNR and low SER, water-filling based on erroneous channel estimates asymptotically achieves the same throughput as equal power allocation.

The results of this chapter have been published in [72].

1.6 Notations

Throughout this thesis, boldface uppercase and lowercase letters denote matrices and vectors, respectively.

Notations $(.)^*$, $(.)^T$, and $(.)^\dagger$ respectively refer to complex conjugate, transpose, and conjugate transpose of a vector or a matrix.

The phase of a complex number is denoted by $\arg(.)$ and $\det[.]$ refers to the determinant of a matrix.

For a vector $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, we define $\mathbf{y_{-m}} = (y_m, y_{m+1}, \ldots, y_n)$ when $m \leq n$ and $\|\mathbf{y}\|$ is the Euclidean norm of $\mathbf{y}$.

Furthermore, $\mathbf{I}_n$ denotes the $n \times n$ identity matrix, $\mathsf{Tr}(\mathbf{Q})$ is the trace of a square matrix $\mathbf{Q}$, and $\mathbf{Q} \succeq 0$ means that $\mathbf{Q}$ is positive semi-definite.

The distribution of a circularly symmetric complex Gaussian (CSCG) vector with the mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ is written as $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

For functions $p(x)$ and $q(x)$ defined on some subset of real numbers, we have $p(x) = O(q(x))$ as $x \to 0$, if and only if there exist positive numbers δ and M such that $|p(x)| \leq M|q(x)|$ for all $|x| < \delta$. Furthermore, we define $p(x) = \Theta(q(x))$, if and only if there exist positive numbers ξ, M_1, and M_2 such that $M_1|q(x)| \leq |p(x)| \leq M_2|q(x)|$ for all $|x| < \xi$.

Finally, $\mathsf{E}[.]$ denotes the expectation of a random variable.

Chapter 2

Performance of Beamforming with Channel Estimation Error

2.1 Introduction

In this chapter, we explore spatial diversity as well as interference management via beamforming in MISO systems. As stated earlier, if a wireless transmitter has multiple antennas that are separated far enough and the environment is rich scattering, independent channels can be established between the transmit antennas and a single-antenna receiver with high probability. Thus, a single stream of data can be sent over these independent channels simultaneously using the same frequency band to increase the received signal's power due to spatial diversity.

In particular, with a transmit power constraint, the transmit covariance matrix can be chosen such that the received power is maximized at some intended receivers and at the same time, co-channel interference is limited at some unintended receivers operating in the same frequency band. As shown in [51–53], the optimal covariance matrix that achieves such a goal in a MISO system is rank-one which implies that beamforming is optimal. Therefore, taking advantage of the variations in the channel gains over space, the transmitter can set beamforming weights such that the received signals from different

transmit antennas combine destructively at the unintended receivers and constructively at the intended receiver [51–57].

In MISO systems, obtaining closed-form expressions for performance metrics such as maximum received power at the intended receiver and the interference power at the interfered receivers are undoubtedly crucial, especially in the realistic case of imperfect CSIT. Such expressions provide remarkable insights into the impact of different parameters (such as the estimation error's variance) on the performance of wireless systems operating in the same frequency band. This chapter aims to obtain such closed-form results in both cases of perfect and imperfect CSIT. A key contribution of this chapter is to introduce an alternative analytic approach to optimal beamforming in the perfect CSIT case which is also applicable to the case of imperfect CSIT.

2.2   System Model

Throughout this chapter, we consider the MISO system model shown in Fig. 2.1. The system consists of a transmitter equipped with n antennas (referred to as node Tx or Tx for short). This multi-antenna transmitter aims to communicate with its intended single-antenna receiver (referred to as node Rx or Rx for short). We assume that such a communication is one-way and is from the Tx to the Rx. In addition, the total transmit power over n antennas is denoted by P. We further consider that there are m unintended single-antenna receivers which we refer to as $\mathsf{Rx}_1, \ldots, \mathsf{Rx}_m$. These unintended receivers are subject to co-channel interference from the Tx upon its operation.

We assume that the transmit antennas, which are all located at Tx, are separated far enough so that independent channels are established between the transmit antennas and the single-antenna receivers. In order to perform an effective transmit beamforming, we exploit the variations in the channel gains over space due to small-scale fading. Throughout this chapter, we do not consider large-scale propagation effects such as path loss because path loss does not affect the final results of this chapter except for introducing a scaling factor.
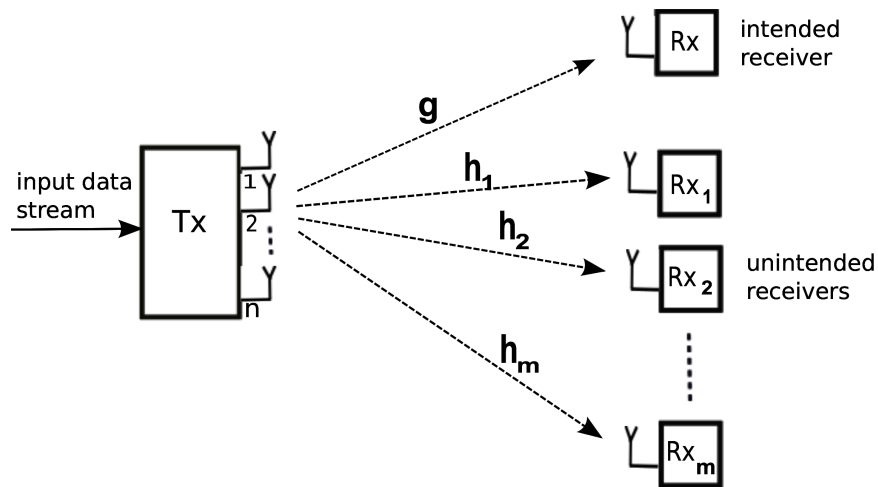
Figure 2.1:   MISO System model under consideration; a multi-antenna transmitter, Tx, is in communication with its intended single-antenna receiver, Rx, and imposing co-channel interference on m unintended single-antenna receivers, $\mathsf{Rx}_1, \ldots, \mathsf{Rx}_m$.

Assuming a flat-fading channel, the channel gain between the Tx and the Rx (the intended link) over the entire frequency band is denoted by $\mathbf{g} \in \mathbb{C}^{n \times 1}$. By splitting $\mathbf{g}$ into its zero-mean scattering component $\mathbf{u}$ and its mean vector $\boldsymbol{\mu}$, we obtain

$$\mathbf{g} = \mathbf{u} + \boldsymbol{\mu}, \tag{2.1}$$

where $\mathbf{u} \sim \mathcal{CN}(\mathbf{0}, 2\sigma^2 \mathbf{I}_n)$. In other words, the entries of $\mathbf{g}$ have a magnitude that is Rician distributed. Note that because of the separation of the n antennas in the Tx, n independent channels are assumed to be established between each transmit antenna in the Tx and the antenna in the Rx.

The mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^T$ in (2.1) is defined as the line-of-sight vector between the Tx and the Rx, and μ_l is the line-of-sight component between the Rx and the lth antenna of the Tx. Thus, we refer to $\|\boldsymbol{\mu}\|^2 = \sum_{l=1}^{n} |\mu_l|^2$ as the line-of-sight component between the Tx and the Rx.

Even though the line-of-sight component, $\|\boldsymbol{\mu}\|^2$, is a non-random value, we assume that all the real and imaginary parts of $\mu_1, \ldots, \mu_n$ are random variables that are identically (but

25

not independently) distributed and the random point $\boldsymbol{\mu}$ is uniformly located on the surface of a $(2n)$-dimensional sphere with the non-random and fixed radius $\|\boldsymbol{\mu}\|$.

We define the K-factor as the ratio of $\|\boldsymbol{\mu}\|^2$ to the total power of the zero-mean scattering component $\mathbf{u}$ as

$$K = \frac{\|\boldsymbol{\mu}\|^2}{2\sigma^2 n}. \tag{2.2}$$

As shown in Fig. 2.1, the channel gain between the Tx and the ith unintended receiver (Rx_i), $i = 1, \ldots, m$, is denoted by $\mathbf{h_i} \in \mathbb{C}^{n \times 1}$ where $\mathbf{h_i} \sim \mathcal{CN}(\mathbf{0}, 2\sigma_I^2 \mathbf{I}_n)$. Since in reality, a typical transmitter tends to be located closer to its intended receiver and farther away from the unintended receivers, it is reasonable to assume that there is a line-of-sight link between the Tx and the Rx, while no line-of-sight link is established between the Tx and the unintended receivers. Without loss of generality, we assume that $\sigma_I^2 = \sigma^2$. Furthermore, having a rich scattering envirnment, the entries of $\mathbf{g}$ are assumed to be independent of the entries of $\mathbf{h_i}$ for $i = 1, \ldots, m$.

Considering the system model described above, the first goal is to study *null-steering beamforming* in Section 2.3, which is the case that the unintended receivers $\mathsf{Rx}_1, \mathsf{Rx}_2, \ldots, \mathsf{Rx}_m$ tolerate no co-channel interference from the Tx. Then, we take one step further and in Section 2.4, we examine ϵ-*threshold beamforming* which is the case that the interference constraint is relaxed to some extent, i.e., co-channel interference less than a certain threshold ϵ is allowed at the unintended receivers. For tractability, in the analysis in Section 2.4, we take $K = 0$ and $m = 1$.

It is worth noting that we do not consider bidirectional communication between the transmitter and the intended receiver. To facilitate the reverse communication in the intended link and at the same time satisfy the interference constraints, the intended receiver also needs to employ multiple antennas and thus perform beamforming effectively.

2.3   Null-steering Beamforming

In this section[1], we focus on characterizing a suitable beamforming vector at the Tx so that with a limited transmit power P, the received power at the intended receiver Rx is maximized while the interference power at each of the unintended receivers $Rx_1, \ldots, Rx_m$, is nullified. We investigate two separate cases of perfect CSIT and imperfect CSIT in the sequel.

2.3.1   Perfect CSIT

We first study the perfect CSIT case, i.e., we assume that perfect knowledge of the channel gain vectors $\mathbf{g}$, and $\mathbf{h_i}$, $i = 1, 2, \ldots, m$, is available at the Tx. Accordingly, we formulate an optimization problem based on the system model. We then solve the optimization problem to find the optimal beamforming vector, and subsequently find the maximum received power at the Rx.

Problem Formulation

We denote the signal transmitted by the Tx at the discrete time instant k by the $n \times 1$ vector $\mathbf{s}[k]$. Therefore, the received signal at the Rx at time instant k can be expressed as $r[k] = \mathbf{g}^T\mathbf{s}[k] + o[k]$, where $o[k]$ is the additive white Gaussian noise at the Rx. Similarly, we have $r_i[k] = \mathbf{h_i}^T\mathbf{s}[k] + o_i[k]$ as the received signal at Rx_i, for $i = 1, 2, \ldots, m$, at time instant k, where $o_i[k]$ is the additive white Gaussian noise at Rx_i.

Let $\mathbf{Q} = \mathsf{E}[\mathbf{s}[k]\mathbf{s}^\dagger[k]]$ be the transmit covariance matrix of the Tx. Since we aim to perform null-steering to divert all the interference from each unintended receiver, the optimization problem can then be formulated as (**P0**)

$$G^{\mathrm{null}} = \max \quad \mathbf{g}^T\mathbf{Q}\mathbf{g}^* \tag{2.3}$$

$$\text{subject to:} \quad \mathbf{h_i}^T\mathbf{Q}\mathbf{h_i}^* = 0, \ \text{for all } i = 1, \ldots, m,$$

$$\mathsf{Tr}(\mathbf{Q}) \leq P, \quad \mathbf{Q} \succeq 0, \tag{2.4}$$

---

[1]The results presented in this section have been published in [70].

where the optimization (2.3) is over the transmit covariance matrix $\mathbf{Q}$. In $\mathbf{P0}$, $\mathbf{g}^T\mathbf{Q}\mathbf{g}^*$ is the received power at the Rx that is aimed to be maximized and $\mathbf{h_i}^T\mathbf{Q}\mathbf{h_i}^*$ is the received power at the ith unintended receiver Rx$_i$ that is aimed to be nullified. Also, in (2.4), the transmit power of the Tx is restricted to P.

Since we aim to study transmit beamforming, the covariance matrix $\mathbf{Q}$ in problem $\mathbf{P0}$ should be rank-one [51–53]. Thus $\mathbf{Q}$ can be expressed as $\mathbf{Q} = \mathbf{x}\mathbf{x}^\dagger$ where $\mathbf{x} \in \mathbb{C}^{\mathbf{n}\times\mathbf{1}}$ is called the *beamforming* vector. Consequently, by replacing $\mathbf{Q} = \mathbf{x}\mathbf{x}^\dagger$ in $\mathbf{P0}$, we obtain the equivalent optimization problem over the beamforming vector $\mathbf{x}$ as ($\mathbf{P1}$)

$$G^{\mathrm{null}} = \max \quad |\mathbf{x}^T\mathbf{g}|^2 \tag{2.5}$$
$$\text{subject to:} \quad |\mathbf{x}^T\mathbf{h_i}|^2 = 0, \text{ for all } i = 1,\ldots,m, \tag{2.6}$$
$$\|\mathbf{x}\|^2 \leq P,$$

where the optimization (2.7) is over the beamforming vector $\mathbf{x}$. We further refer to the m constraints in (2.6) as the *null-steering constraints*. Therefore, the focus now is to find a suitable beamforming vector $\mathbf{x}$ that satisfies the m null-steering constraints and, at the same time, maximizes the received power at the Rx.

It is worth noting that, beamforming is performed here at each discrete time instant k based on the corresponding channel knowledge at that instant. Consequently, the null-steering constraints (2.6) are satisfied at each time instant k and thus are characterized as the peak interference power constraint [73].

In the case that no unintended receiver is present ($m = 0$), the transmit beamforming problem above simply reduces to the transmit maximal-ratio combining that maximizes the received power at the Rx over the vector $\mathbf{x}$. The average received power in this case grows linearly with the number of transmit antennas n.

Solving the Optimization Problem

The problem **P1** can be rewritten as (**P2**)

$$G^{\text{null}} = \max \quad |\mathbf{x}^T\mathbf{g}|^2 \tag{2.7}$$

$$\text{subject to:} \quad \mathbf{x}^T\mathbf{H} = \mathbf{0}, \tag{2.8}$$

$$\|\mathbf{x}\|^2 \leq P,$$

where $\mathbf{H} \in \mathbb{C}^{n \times m}$ and its ith column is the vector $\mathbf{h_i}$ for all $i = 1, 2, \ldots, m$.

The singular value decomposition (SVD) of $\mathbf{H}$ yields $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}^{1/2}\mathbf{V}$, where $\mathbf{U} \in \mathbb{C}^{n \times n}$ and $\mathbf{V} \in \mathbb{C}^{m \times m}$ are unitary matrices and $\boldsymbol{\Sigma} \in \mathbb{C}^{n \times m}$ is diagonal. Therefore, we can express the null-steering constraint $\mathbf{x}^T\mathbf{H} = \mathbf{0}$ as $\mathbf{x}^T\mathbf{U}\boldsymbol{\Sigma}^{1/2}\mathbf{V} = \mathbf{0}$, which is further reduced to $\mathbf{x}^T\mathbf{U}\boldsymbol{\Sigma}^{1/2} = \mathbf{0}$. Defining $\mathbf{x}^T\mathbf{U} = \mathbf{z}^T$, where $\mathbf{z} = (z_1, \ldots, z_n)^T$ is a new coordinate system, we obtain $\mathbf{z}^T\boldsymbol{\Sigma}^{1/2} = \mathbf{0}$ as the equivalent null-steering constraint.

Note that due to the statistical independence of vectors $\mathbf{h_i}$ for all i, the realizations of $\mathbf{h_i}$ for all i are linearly independent with high probability and thus $\mathbf{H}$ is full-rank which leads to $\boldsymbol{\Sigma}$ being full-rank as well. Therefore, satisfying the constraint $\mathbf{z}^T\boldsymbol{\Sigma}^{1/2} = \mathbf{0}$ yields the elimination of m dimensions (assuming $m < n$).

Furthermore, we find the new transmit power constraint as $\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2 \leq P$ and the new objective function $|\mathbf{x}^T\mathbf{g}|^2 = |\mathbf{x}^T\mathbf{U}\mathbf{U}^{-1}\mathbf{g}|^2 = |\mathbf{z}^T\tilde{\mathbf{g}}|^2$, where $\tilde{\mathbf{g}} = \mathbf{U}^{-1}\mathbf{g}$ is the rotated version of the vector $\mathbf{g}$. Thus, a new formulation for **P2** in terms of the new coordinate system $\mathbf{z}$, can be expressed as (**P3**)

$$G^{\text{null}} = \max \quad |\mathbf{z}^T\tilde{\mathbf{g}}|^2 \tag{2.9}$$

$$\text{subject to:} \quad \mathbf{z}^T\boldsymbol{\Sigma}^{1/2} = \mathbf{0} \tag{2.10}$$

$$\|\mathbf{z}\|^2 \leq P. \tag{2.11}$$

Having $\mathbf{g} = \mathbf{u} + \boldsymbol{\mu}$ from (2.1), we obtain

$$\tilde{\mathbf{g}} = \mathbf{U}^{-1}\mathbf{g} = \mathbf{U}^{-1}\mathbf{u} + \mathbf{U}^{-1}\boldsymbol{\mu} = \tilde{\mathbf{u}} + \tilde{\boldsymbol{\mu}}. \tag{2.12}$$

It is crucial to point out the following important facts regarding the rotations of the channel gain vectors:

1. Note that the random point $\boldsymbol{\mu}$ is uniformly located on the surface of a $(2n)$-dimensional sphere with the non-random and fixed radius $\|\boldsymbol{\mu}\|$. This point is then rotated by the rotation matrix $\mathbf{U}^{-1}$. Therefore, the new random point $\tilde{\boldsymbol{\mu}}$ is also uniformly located on the same sphere, which implies $\|\tilde{\boldsymbol{\mu}}\|^2 = \|\boldsymbol{\mu}\|^2$, and all the real and imaginary parts of $\tilde{\mu}_1, \ldots, \tilde{\mu}_n$ are also identically (but not independently) distributed.

2. Since the rotation matrix $\mathbf{U}^{-1}$ is a unitary matrix, the zero-mean CSCG-distributed vector $\mathbf{u}$ is statistically invariant under rotation and thus we have $\tilde{\mathbf{u}} \sim \mathcal{CN}(\mathbf{0}, 2\sigma^2 \mathbf{I}_n)$.

Denoting the optimal solution to **P3** as $\mathbf{z}^{\mathsf{opt}}$, maximization of the received power $|\mathbf{z}^T \tilde{\mathbf{g}}|^2$ in (2.9) yields $\arg(z_i^{\mathsf{opt}}) = -\arg(\tilde{g}_i)$, for $i = 1, \ldots, n$. Therefore, making (2.11) tight in the constraint, we obtain

$$\mathbf{z}^{\mathsf{opt}} = \frac{\sqrt{P} \left(0, \ldots, 0, \tilde{g}_{m+1}, \ldots, \tilde{g}_n\right)^{\dagger}}{\sqrt{\sum_{l=m+1}^{n} |\tilde{g}_l|^2}}. \tag{2.13}$$

Note that if $n \leq m$, we will have $\mathbf{z}^{\mathsf{opt}} = \mathbf{0}$ which is not interesting. Therefore, in the sequel, we use the assumption that $m < n$.

Maximum Received Power at the Intended Receiver Rx

Having $\mathbf{z}^{\mathsf{opt}}$ as in (2.13), the maximum received power at the Rx is

$$G^{\mathrm{null}} = |\tilde{\mathbf{g}}^T \mathbf{z}^{\mathsf{opt}}|^2 = P \sum_{l=m+1}^{n} |\tilde{g}_l|^2 = P \sum_{l=m+1}^{n} |\tilde{u}_l + \tilde{\mu}_l|^2, \tag{2.14}$$

where $\tilde{u}_l$ is the lth entry of the rotated vector $\tilde{\mathbf{u}}$. Conditioning on $\tilde{\boldsymbol{\mu}}$, the random variable $G^{\mathrm{null}}/P\sigma^2$ has non-central chi-square distribution with $2(n-m)$ degrees of freedom (see

[74]). Thus, the expected value of G^{null} in (2.14) can be derived as

$$\mathsf{E}\left[G^{\text{null}}\right] = \mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\mathsf{E}\left[G^{\text{null}}|\tilde{\boldsymbol{\mu}}\right]\right]$$

$$= P\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\mathsf{E}\left[\sum_{l=m+1}^{n}\left(\mathrm{Re}\left(\tilde{u}_l\right) + \mathrm{Re}\left(\tilde{\mu}_l\right)\right)^2 |\tilde{\boldsymbol{\mu}}\right]\right]$$

$$+ P\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\mathsf{E}\left[\sum_{l=m+1}^{n}\left(\mathrm{Im}\left(\tilde{u}_l\right) + \mathrm{Im}\left(\tilde{\mu}_l\right)\right)^2 |\tilde{\boldsymbol{\mu}}\right]\right]$$

$$= P\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\sum_{l=m+1}^{n}\mathsf{E}\left[\left(\mathrm{Re}\left(\tilde{u}_l\right) + \mathrm{Re}\left(\tilde{\mu}_l\right)\right)^2 |\tilde{\boldsymbol{\mu}}\right]\right]$$

$$+ P\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\sum_{l=m+1}^{n}\mathsf{E}\left[\left(\mathrm{Im}\left(\tilde{u}_l\right) + \mathrm{Im}\left(\tilde{\mu}_l\right)\right)^2 |\tilde{\boldsymbol{\mu}}\right]\right]$$

$$= P\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\sum_{l=m+1}^{n}\left(\sigma^2 + \mathrm{Re}\left(\tilde{\mu}_l\right)^2\right)\right] + P\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\sum_{l=m+1}^{n}\left(\sigma^2 + \mathrm{Im}\left(\tilde{\mu}_l\right)^2\right)\right]$$

$$= 2P\sigma^2\left(n - m\right) + P\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\sum_{l=m+1}^{n}\mathrm{Re}\left(\tilde{\mu}_l\right)^2 + \sum_{l=m+1}^{n}\mathrm{Im}\left(\tilde{\mu}_l\right)^2\right]. \tag{2.15}$$

Suppose we are given a random point $\boldsymbol{\mu}_R = (\mathrm{Re}(\tilde{\mu}_1), \mathrm{Im}(\tilde{\mu}_1), \ldots, \mathrm{Re}(\tilde{\mu}_n), \mathrm{Im}(\tilde{\mu}_n))$ uniformly distributed on a $(2n)$-dimensional sphere with radius $\|\boldsymbol{\mu}\|$. Define $X = \mathrm{Re}(\tilde{\mu}_1)$ as the projection of $\boldsymbol{\mu}_R$ to its first coordinate. According to [75], X has the following probability distribution function

$$f_X(x) = \frac{\left(\sqrt{\|\boldsymbol{\mu}\|^2 - x^2}\right)^{2n-3}}{c}, \tag{2.16}$$

for $-\|\boldsymbol{\mu}\| \leq x \leq \|\boldsymbol{\mu}\|$, where $c = \sqrt{\pi}\|\boldsymbol{\mu}\|^{2n-2}\Gamma(n - \frac{1}{2})/\Gamma(n)$ is an appropriate scaling factor in which $\Gamma(n)$ is the Gamma function. Consequently, we can find

$$\mathsf{E}[X^2] = \frac{1}{c}\int_{-\|\boldsymbol{\mu}\|}^{\|\boldsymbol{\mu}\|}x^2\left(\sqrt{\|\boldsymbol{\mu}\|^2 - x^2}\right)^{2n-3}dx = \frac{2\|\boldsymbol{\mu}\|^{2n}\Gamma(3/2)\Gamma(n)}{\|\boldsymbol{\mu}\|^{2n-2}\Gamma(1/2)\Gamma(n+1)} = \frac{\|\boldsymbol{\mu}\|^2}{2n},$$

using the fact that (see [76])

$$\int_{0}^{\|\boldsymbol{\mu}\|}x^2\left(\sqrt{\|\boldsymbol{\mu}\|^2 - x^2}\right)^{2n-3}dx = \frac{\|\boldsymbol{\mu}\|^{2n}\,\Gamma(3/2)\Gamma(n-1/2)}{2\,\Gamma(n+1)}.$$

Therefore following from (2.15) and knowing that the entries of $\boldsymbol{\mu}_R$ are identically distributed, we find

$$\mathsf{E}[G^{\mathrm{null}}] = 2P\sigma^2(n-m) + \frac{P\|\boldsymbol{\mu}\|^2(n-m)}{n}$$

$$= 2P\sigma^2(n-m)(K+1), \tag{2.17}$$

in terms of the K-factor defined in (2.2). As shown in (2.17), the expected value of the maximum received power at the Rx grows linearly with $(n-m)$ and is directly proportional to the K-factor.

The variance of G^{null} in (2.14) can be expressed as (see [77])

$$\mathsf{Var}[G^{\mathrm{null}}] = \mathsf{Var}_{\tilde{\boldsymbol{\mu}}}[\mathsf{E}[G^{\mathrm{null}}|\tilde{\boldsymbol{\mu}}]] + \mathsf{E}_{\tilde{\boldsymbol{\mu}}}[\mathsf{Var}[G^{\mathrm{null}}|\tilde{\boldsymbol{\mu}}]]. \tag{2.18}$$

From (2.14), we have

$$\mathsf{E}\left[G^{\mathrm{null}}|\tilde{\boldsymbol{\mu}}\right] = P\mathsf{E}\left[\sum_{l=m+1}^{n}\left(\mathrm{Re}\left(\tilde{u}_l\right) + \mathrm{Re}\left(\tilde{\mu}_l\right)\right)^2|\tilde{\boldsymbol{\mu}}\right] + P\mathsf{E}\left[\sum_{l=m+1}^{n}\left(\mathrm{Im}\left(\tilde{u}_l\right) + \mathrm{Im}\left(\tilde{\mu}_l\right)\right)^2|\tilde{\boldsymbol{\mu}}\right]$$

$$= 2P\sigma^2\left(n-m\right) + P\left(\sum_{l=m+1}^{n}\mathrm{Re}\left(\tilde{\mu}_l\right)^2 + \sum_{l=m+1}^{n}\mathrm{Im}\left(\tilde{\mu}_l\right)^2\right).$$

Thus, we can write

$$\mathsf{Var}_{\tilde{\boldsymbol{\mu}}}[\mathsf{E}[G^{\mathrm{null}}|\tilde{\boldsymbol{\mu}}]] = P^2\mathsf{Var}_{\tilde{\boldsymbol{\mu}}}\left[\sum_{l=m+1}^{n}\mathrm{Re}\left(\tilde{\mu}_l\right)^2 + \sum_{l=m+1}^{n}\mathrm{Im}\left(\tilde{\mu}_l\right)^2\right]$$

$$= 2P^2(n-m)\mathsf{Var}[X^2] + 2P^2(n-m)(2n-2m-1)\mathsf{Cov}[X^2,Y^2], \quad (2.19)$$

where as stated earlier, $(X,Y) = (\mathrm{Re}(\tilde{\mu}_1), \mathrm{Im}(\tilde{\mu}_1))$. Knowing $\mathsf{Var}[X^2] = \mathsf{E}[X^4] - \mathsf{E}^2[X^2]$ and finding from (2.16) that

$$\mathsf{E}[X^4] = \frac{(2n-1)\|\boldsymbol{\mu}\|^4}{4n^2(n+1)},$$

we derive

$$\mathsf{Var}[X^2] = \frac{(n-2)\|\boldsymbol{\mu}\|^4}{4n^2(n+1)}.$$

32

Furthermore, since

$$\mathsf{Var}\left[\|\boldsymbol{\mu}\|^2\right] = 2n\mathsf{Var}[X^2] + 2n(2n-1)\mathsf{Cov}[X^2, Y^2] = 0,$$

we obtain

$$\mathsf{Cov}[X^2, Y^2] = \frac{\mathsf{Var}[X^2]}{1-2n}.$$

Consequently, we find the first term of (2.18) as

$$\begin{aligned}
\mathsf{Var}_{\tilde{\boldsymbol{\mu}}}[\mathsf{E}[G^{\mathrm{null}}|\tilde{\boldsymbol{\mu}}]] &= P^2 m \|\boldsymbol{\mu}\|^4 \frac{(n-m)(n-2)}{n^2(n+1)(2n-1)} \\
&= 4P^2 K^2 \sigma^4 m \frac{(n-m)(n-2)}{(n+1)(2n-1)}.
\end{aligned} \tag{2.20}$$

in terms of the K-factor.

From (2.14), we have

$$\begin{aligned}
\mathsf{Var}[G^{\mathrm{null}}|\tilde{\boldsymbol{\mu}}] &= P^2 \mathsf{Var}\left[\sum_{l=m+1}^{n}(\mathrm{Re}(\tilde{u}_l)+\mathrm{Re}(\tilde{\mu}_l))^2 + \sum_{l=m+1}^{n}(\mathrm{Im}(\tilde{u}_l)+\mathrm{Im}(\tilde{\mu}_l))^2|\tilde{\boldsymbol{\mu}}\right] \\
&= P^2 \mathsf{Var}\left[\sum_{l=m+1}^{n}(\mathrm{Re}(\tilde{u}_l)+\mathrm{Re}(\tilde{\mu}_l))^2|\tilde{\boldsymbol{\mu}}\right] + P^2 \mathsf{Var}\left[\sum_{l=m+1}^{n}(\mathrm{Im}(\tilde{u}_l)+\mathrm{Im}(\tilde{\mu}_l))^2|\tilde{\boldsymbol{\mu}}\right] \\
&= P^2 \sigma^4 \mathsf{Var}\left[\frac{1}{\sigma^2}\sum_{l=m+1}^{n}(\mathrm{Re}(\tilde{u}_l)+\mathrm{Re}(\tilde{\mu}_l))^2|\tilde{\boldsymbol{\mu}}\right] \\
&\quad + P^2 \sigma^4 \mathsf{Var}\left[\frac{1}{\sigma^2}\sum_{l=m+1}^{n}(\mathrm{Im}(\tilde{u}_l)+\mathrm{Im}(\tilde{\mu}_l))^2|\tilde{\boldsymbol{\mu}}\right] \\
&= 2P^2 \sigma^4 \left(n-m+\frac{2}{\sigma^2}\sum_{l=m+1}^{n}\mathrm{Re}(\tilde{\mu}_l)^2\right) + 2P^2 \sigma^4 \left(n-m+\frac{2}{\sigma^2}\sum_{l=m+1}^{n}\mathrm{Im}(\tilde{\mu}_l)^2\right),
\end{aligned} \tag{2.21}$$

which is found knowing that given $\tilde{\boldsymbol{\mu}}$, the random variable $\frac{1}{\sigma^2}\sum_{l=m+1}^{n}(\mathrm{Re}(\tilde{u}_l)+\mathrm{Re}(\tilde{\mu}_l))^2$ in (2.21) is chi-square distributed with $n-m$ degrees of freedom (see [74]).

Therefore, recalling $\mathsf{E}[\mathrm{Re}(\tilde{\mu}_l)^2] = \mathsf{E}[\mathrm{Im}(\tilde{\mu}_l)^2] = \|\boldsymbol{\mu}\|^2/2n$, for $l = m+1, \ldots, n$, we find the second term of (2.18) as

$$\begin{aligned}
\mathsf{E}_{\tilde{\boldsymbol{\mu}}}\left[\mathsf{Var}\left[G^{\mathrm{null}}|\tilde{\boldsymbol{\mu}}\right]\right] &= 4P^2 \sigma^4 (n-m) + \frac{4P^2 \sigma^2 \|\tilde{\boldsymbol{\mu}}\|^2(n-m)}{n} \\
&= 4P^2 \sigma^4 (n-m)(1+2K),
\end{aligned} \tag{2.22}$$

33

in terms of the K-factor.

By replacing (2.20) and (2.22) in (2.18), we obtain

$$\mathsf{Var}[G^{\mathrm{null}}] = 4P^2\sigma^4(n-m)\left(\frac{m(n-2)}{(n+1)(2n-1)}K^2 + 2K + 1\right).$$

Using Chebychev's inequality [77], we have for every $\zeta > 0$

$$\mathsf{Pr}\left[\left|G^{\mathrm{null}} - \mathsf{E}[G^{\mathrm{null}}]\right| \geq \zeta\mathsf{E}[G^{\mathrm{null}}]\right] \leq \frac{\mathsf{Var}[G^{\mathrm{null}}]}{\zeta^2\mathsf{E}^2[G^{\mathrm{null}}]}. \tag{2.23}$$

Since

$$\frac{\mathsf{Var}[G^{\mathrm{null}}]}{\mathsf{E}^2[G^{\mathrm{null}}]} = \frac{\frac{m(n-2)}{(n+1)(2n-1)}K^2 + 2K + 1}{(n-m)(K+1)^2},$$

converges to zero as $n \to \infty$, therefore G^{null} converges to $\mathsf{E}[G^{\mathrm{null}}]$ in probability as n increases. Thus, $\mathsf{E}[G^{\mathrm{null}}]$ becomes a more precise indicator of the realization of G^{null} when the number of antennas increases.

Performance Evaluation

Summarizing the results obtained in this section in the case of perfect CSIT, we can point out the following facts:

1. When the number of transmit antennas is greater than the number of unintended receivers $(n > m)$, the transmitter can successfully perform null-steering beamforming and achieve a nonzero value for the mean received power $\mathsf{E}[G^{\mathrm{null}}]$ which grows linearly with $(n-m)$. Intuitively, in such a case, m degrees of freedom are used for nullifying the interference at m unintended receivers while the remaining $n-m$ degrees of freedom are used towards spatial diversity in the intended link. Therefore, either a larger number of antennas or smaller number of unintended receivers leads to higher values of $\mathsf{E}[G^{\mathrm{null}}]$.

2. The mean received power $\mathsf{E}[G^{\mathrm{null}}]$ is directly proportional to the power of the line-of-sight component represented by the K-factor. Thus, a better line-of-sight component in the intended link provides a better performance.

3. Since the ratio $\mathsf{Var}[G^{\mathrm{null}}]/\zeta^2\mathsf{E}^2[G^{\mathrm{null}}]$ approaches zero as n grows large for any $\zeta > 0$, the stochastic average $\mathsf{E}[G^{\mathrm{null}}]$ becomes a more precise indicator of the realization of G^{null} as the number of antennas increases.

2.3.2   Imperfect CSIT

As previously discussed in Section 1.3, obtaining perfect CSIT is very challenging in practice. Thus, a more general and realistic scenario is to consider the case when the channel gain vectors are not perfectly known at the Tx which happens due to nonzero channel estimation errors. In this section, we address the case of imperfect CSIT by modelling the channel estimation error to be Gaussian distributed (which is a reasonable assumption in estimation methods such as maximum likelihood (ML) estimation [39]). This model allows us to obtain statistical closed-form results and find how the performance is affected due to the lack of perfect CSIT.

With ML estimation [39], the estimated channel gain vectors are expressed as

$$\hat{\mathbf{g}} = \mathbf{g} + \mathbf{w}, \quad \hat{\mathbf{h}_i} = \mathbf{h_i} + \mathbf{w_i}, \tag{2.24}$$

where $\mathbf{g}$ and $\mathbf{h_i}$ for $i = 1, \ldots, m$ are the actual channel gain vectors. $\mathbf{w}$ and $\mathbf{w_i}$, are the estimation error vectors which are distributed as $\mathcal{CN}(\mathbf{0}, 2\sigma_e^2\mathbf{I}_n)$, for $i = 1, \ldots, m$ and are independent for all i. Furthermore, we assume that the actual channel gains and their respective estimation errors are independent.

Throughout this chapter, we denote the ratio of the scattering component's power to the power of the estimation error as $\gamma = \sigma^2/\sigma_e^2$ and interchangeably refer to it as the signal to estimation error ratio (SER).

We assume that the Tx does not have a priori knowledge of the error vectors or σ_e^2, or any uncertainty regions containing the actual channels. Therefore, uncertainty analysis can not be performed. Instead, we find an optimal beamforming vector using the estimates of channel gains instead of the actual channel gains. Consequently, we derive the received power at the Rx (respectively interference power at $\mathsf{Rx}_1, \ldots, \mathsf{Rx}_m$) corresponding to this

beamforming vector and refer to it as the "actual" received power (respectively interference power). Then, we analyze the effect of channel estimation error on these results. Intuitively, estimation error leads to extra interference at the unintended receivers $\mathsf{Rx}_1, \ldots, \mathsf{Rx}_m$ and less received power at the Rx compared to the corresponding values in the case of perfect CSIT. Therefore, the results in the case of perfect CSIT are upper bounds for the results obtained in this section in the case of imperfect CSIT.

Problem Formulation

Having the estimated channel gain vectors in (2.24), we obtain the following optimization problem (**P4**) under the system model described in Section 2.2

$$
\begin{aligned}
\max \quad & |\mathbf{x}^T \hat{\mathbf{g}}|^2 \\
\text{subject to:} \quad & |\mathbf{x}^T \hat{\mathbf{h}_i}|^2 = 0, \ \text{for all } i = 1, \ldots, m, \\
& \|\mathbf{x}\|^2 \leq P.
\end{aligned}
\tag{2.25}
$$

P4 has the same form as **P1** in Section 2.3.1 with estimated channel gains instead. Therefore, to solve **P4** we follow the same approach as in Section 2.3.1 and obtain the optimal solution as

$$
\mathbf{z}^{\mathsf{opt}} = \frac{\sqrt{P} \left(0, \ldots, 0, \tilde{\hat{g}}_{m+1}, \ldots, \tilde{\hat{g}}_n\right)^\dagger}{\sqrt{\sum_{l=m+1}^n |\tilde{\hat{g}}_l|^2}},
\tag{2.26}
$$

where $\tilde{\hat{g}}_{m+1}$ denotes the $(m+1)$th entry of $\tilde{\hat{\mathbf{g}}}$.

Actual Interference Power at $\mathsf{Rx}_1, \ldots, \mathsf{Rx}_m$

Because of nonzero estimation errors, the interference power at each unintended receiver is not zero anymore. We denote the actual interference power at the ith unintended receiver as $I_i^{\text{null, Im}}$ and derive it as follows

$$
I_i^{\text{null, Im}} = |\mathbf{h_i}^T \mathbf{x}^{\mathsf{opt}}|^2 = |(\hat{\mathbf{h}_i} - \mathbf{w_i})^T \mathbf{x}^{\mathsf{opt}}|^2 = |\mathbf{w_i}^T \mathbf{x}^{\mathsf{opt}}|^2,
$$

for $i = 1, 2, \ldots, m$. Note that in the derivation of $I_i^{\text{null, Im}}$ above, we use (2.24) and the fact that according to (2.25), for the optimal beamforming vector $\mathbf{x}^{\text{opt}}$, $\hat{\mathbf{h}}_{\mathbf{i}}^T \mathbf{x}^{\text{opt}} = 0$ for $i = 1, 2, \ldots, m$. Therefore,

$$I_i^{\text{null, Im}} = |\mathbf{w}_{\mathbf{i}}^T \mathbf{x}^{\text{opt}}|^2 = |\tilde{\mathbf{w}}_{\mathbf{i}}^T \mathbf{z}^{\text{opt}}|^2,$$

where $\tilde{\mathbf{w}}_{\mathbf{i}}$ is the rotated version of $\mathbf{w}_{\mathbf{i}}$ which is also distributed as $\mathcal{CN}(\mathbf{0}, 2\sigma_e^2 \mathbf{I}_n)$. Hence, using (2.26), we obtain

$$
\begin{aligned}
I_i^{\text{null, Im}} &= \frac{P \left| \sum_{l=m+1}^n \tilde{\hat{g}}_l^* \tilde{w}_{il} \right|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2} = \frac{P \left(\sum_{l=m+1}^n \tilde{\hat{g}}_l^* \tilde{w}_{il} \right) \left(\sum_{r=m+1}^n \tilde{\hat{g}}_r \tilde{w}_{ir}^* \right)}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2} \\
&= \frac{P \sum_{l=m+1}^n |\tilde{\hat{g}}_l|^2 |\tilde{w}_{il}|^2 + P \sum_{\substack{l,r \\ l \neq r}}^n \tilde{\hat{g}}_l \tilde{\hat{g}}_r^* \tilde{w}_{il}^* \tilde{w}_{ir}}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2},
\end{aligned}
$$

where $\sum_{\substack{l,r \\ l \neq r}}^n$ is a shorter notation for $\sum_{l=m+1, l \neq r}^n \sum_{r=m+1}^n$ and $\tilde{w}_{il}$ is the lth entry of $\tilde{\mathbf{w}}_{\mathbf{i}}$.

Taking into account that the entries of $\tilde{\mathbf{w}}_{\mathbf{i}}$ (the error vector corresponding to the ith unintended link) and $\tilde{\hat{\mathbf{g}}}$ are independent for all i, the expected value of $I_i^{\text{null, Im}}$ can then be derived as

$$\mathsf{E}[I_i^{\text{null, Im}}] = P \sum_{l=m+1}^n \mathsf{E}\left[\frac{|\tilde{\hat{g}}_l|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right] \mathsf{E}[|\tilde{w}_{il}|^2] + P \sum_{\substack{l,r \\ l \neq r}}^n \mathsf{E}\left[\frac{\tilde{\hat{g}}_l \tilde{\hat{g}}_r^*}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right] \mathsf{E}[\tilde{w}_{il}^*] \mathsf{E}[\tilde{w}_{ir}].$$

Since $\mathsf{E}[\tilde{w}_{il}] = 0$ and $\mathsf{E}[|\tilde{w}_{il}|^2] = 2\sigma_e^2$ for every i and l, we obtain

$$\mathsf{E}[I_i^{\text{null, Im}}] = 2P\sigma_e^2 \sum_{l=m+1}^n \mathsf{E}\left[\frac{|\tilde{\hat{g}}_l|^2}{\sum_{l=m+1}^n |\tilde{\hat{g}}_l|^2} \right] = 2P\sigma_e^2 \mathsf{E}\left[\frac{\sum_{l=m+1}^n |\tilde{\hat{g}}_l|^2}{\sum_{l=m+1}^n |\tilde{\hat{g}}_l|^2} \right] = 2P\sigma_e^2, \quad (2.27)$$

for $i = 1, 2, \ldots, n$.

Maximum Received Power at the Intended Receiver Rx

Using the actual channel gain vector $\mathbf{g}$, the actual received power at the Rx can be expressed as

$$G^{\text{null, Im}} = |\mathbf{g}^T \mathbf{x}^{\text{opt}}|^2 = |\hat{\mathbf{g}}^T \mathbf{x}^{\text{opt}} - \mathbf{w}^T \mathbf{x}^{\text{opt}}|^2 = |\tilde{\hat{\mathbf{g}}}^T \mathbf{z}^{\text{opt}} - \tilde{\mathbf{w}}^T \mathbf{z}^{\text{opt}}|^2, \quad (2.28)$$

which is derived using (2.24).

We can find (see Appendix A) a lower bound on $\mathsf{E}[G^{\text{null, Im}}]$ as

$$
\begin{aligned}
\mathsf{E}[G^{\text{null, Im}}] &\geq 2P\sigma^2(n-m)\left(K+1-\frac{\sigma_e^2}{\sigma^2}+\frac{\sigma_e^4\sigma^2(1-K)+\sigma_e^6}{\sigma^2\left(\sigma^2+\sigma_e^2\right)^2}+\frac{\sigma_e^2}{(\sigma^2+\sigma_e^2)(n-m)}\right) \\
&= 2P\sigma^2(n-m)\left(K+1-\frac{1}{\gamma}+\frac{1-K+\frac{1}{\gamma}}{(1+\gamma)^2}+\frac{1}{(n-m)(1+\gamma)}\right),
\end{aligned}
\tag{2.29}
$$

with equality when $K = 0$ (Rayleigh fading).

As $\gamma \to \infty$, the lower bound on $\mathsf{E}[G^{\text{null, Im}}]$ in (2.29) approaches the result $\mathsf{E}[G^{\text{null}}]$ in the case of perfect CSIT found in Section 2.3.1. Therefore, we can write

$$
\mathsf{E}[G^{\text{null, Im}}] \leq \mathsf{E}[G^{\text{null}}] = 2P\sigma^2(n-m)(K+1),
\tag{2.30}
$$

with equality when $\gamma = \infty$.

Performance Evaluation

Summarizing the results obtained in this section in the case of imperfect CSIT, we can point out the following facts:

1. The closed form expression in (2.27), shows how the two parameters P and σ_e^2 contribute to $\mathsf{E}[I_i^{\text{null, Im}}]$, which is the average interference power at the ith unintended receiver. Interestingly, $\mathsf{E}[I_i^{\text{null, Im}}]$ in (2.27), does not depend on the number of antennas n for $i = 1, 2, \ldots, m$. Thus, for a fixed P and σ_e^2, n can be increased to improve the intended link's performance (due to spatial diversity) without increasing the average interference power at the unintended receivers.

2. In the case that the estimation is very erroneous, σ_e^2 grows larger and thus $\mathsf{E}[I_i^{\text{null, Im}}]$ in (2.27) increases. To prevent such an increase in the average interference, the transmit power P can be decreased (note that $\mathsf{E}[I_i^{\text{null, Im}}]$ is directly proportional to P and σ_e^2). Since $\mathsf{E}[I_i^{\text{null, Im}}]$ is insensitive to n, the Tx gets the freedom to employ additional antennas to compensate for its performance loss due to the reduction of P without resulting in an increase in $\mathsf{E}[I_i^{\text{null, Im}}]$.

3. In the case of Rayleigh fading between the Tx and the Rx ($K = 0$), the lower bound on $\mathsf{E}[G^{\text{null, Im}}]$ in (2.29) becomes tight and we obtain

$$\mathsf{E}[G^{\text{null, Im}}] = 2P\sigma^2(n-m)\left(1 - \frac{1}{\gamma} + \frac{1}{\gamma(1+\gamma)} + \frac{1}{(n-m)(1+\gamma)}\right). \qquad (2.31)$$

4. When $K = 0$, according to (2.31), if $\gamma \gg 1$, the mean received power $\mathsf{E}[G^{\text{null, Im}}]$ approaches the best-case result ($\mathsf{E}[G^{\text{null}}]$) in (2.30). Particularly, for $\gamma = 0$ dB, $\gamma = 4$ dB, and $\gamma = 6$ dB, the mean received power $\mathsf{E}[G^{\text{null, Im}}]$ is 2.5 dB, 1.8 dB, and 1.1 dB away from $\mathsf{E}[G^{\text{null}}]$ respectively when $n = 10$. This result can be observed in Fig. 2.2 that plots $\mathsf{E}[G^{\text{null, Im}}]$ in (2.31) versus n for $K = 0$, $P = 1$, $m = 2$, $\sigma = 1$, and different values of γ. In Fig. 2.2, the solid line is $\mathsf{E}[G^{\text{null}}]$ which is the upper bound of $\mathsf{E}[G^{\text{null, Im}}]$.

5. when $K \neq 0$, according to the lower bound in (2.29), for $K \gg \frac{1}{\gamma}$ the estimation error can be dominated by K and thus the mean received power $\mathsf{E}[G^{\text{null, Im}}]$ approaches the upper bound $\mathsf{E}[G^{\text{null}}]$. Fig. 2.3 plots the lower bound of $\mathsf{E}[G^{\text{null, Im}}]$ versus n for $P = 1$, $m = 2$, $\sigma = 1$, $K = 6$ dB, and different values of γ. This choice of the K-factor is reasonable according to measurements conducted on wireless channels [78]. As it is seen in this figure, for $\gamma = 0$ dB, $K = 6$ dB leads to 0.97 dB difference between the lower bound and the upper bound when $n = 3$, and 1.47 dB difference when $n = 10$. These differences are regardless of σ^2. Furthermore, as γ increases in Fig. 2.3, the lower bound of $\mathsf{E}[G^{\text{null, Im}}]$ approaches the upper bound.

2.4 ϵ-threshold Beamforming

In Section 2.3, we studied null-steering beamforming which is the case when the receivers $\mathsf{Rx}_1, \ldots, \mathsf{Rx}_m$ tolerate no interference from the Tx. Subsequently, under such a condition and with perfect CSIT, we derived the maximum received power at the Rx and showed that its expected value grows linearly with $(n-m)$. Intuitively, if a small nonzero interference (say ϵ) is allowed at the unintended receivers, a higher maximum received power is anticipated at the Rx. In this section[2], we find such a tradeoff between the interference threshold
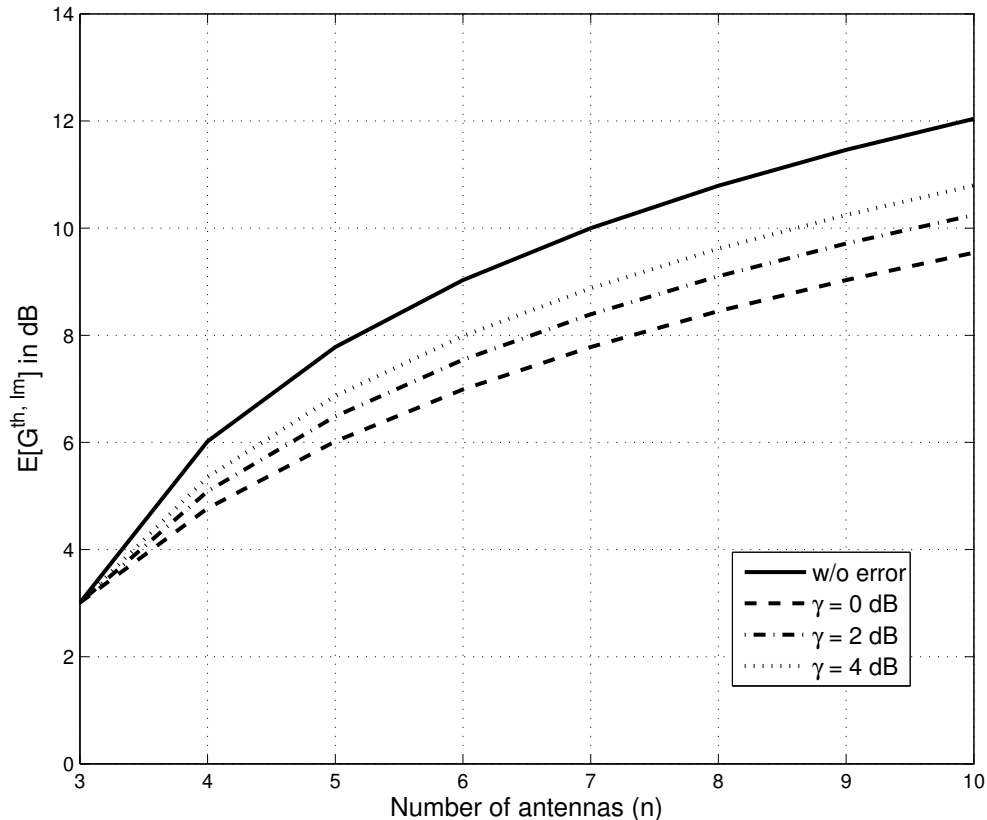
Figure 2.2: $\mathsf{E}[G^{\mathrm{null,\ Im}}]$ in dB versus n, for $K = 0$ and different values of γ when $P = 1$ and $\sigma = 1$. Since $m = 2$, at least $n = 3$ antennas are needed for successful null-steering.

ϵ and the maximum received power at the Rx in the perfect CSIT case. Furthermore, we examine the case of imperfect CSIT as well and similarly to Section 2.3, we formulate the degradation in the performance due to channel estimation errors.

The system model considered in this section is similar to the model in Section 2.2 except that in this section, for simplicity and tractability, we study the special case that $K = 0$ and $m = 1$. In other words, we assume that there is only one unintended receiver (Rx_1) and $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}, 2\sigma^2 \mathbf{I}_n)$. Note that for $m > 1$, the corresponding optimization problem can

---

[2]The results presented in this section have been published in [71].
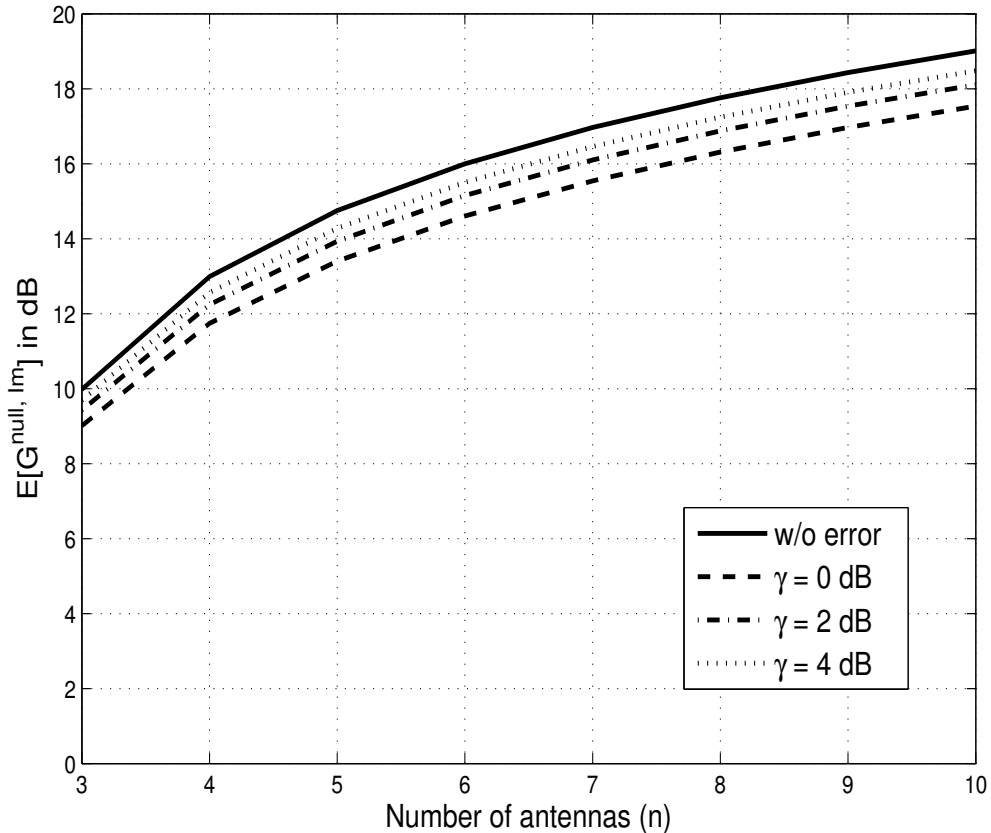
Figure 2.3: Upper bound and lower bound of $\mathsf{E}[G^{\text{null, Im}}]$ in dB versus n, for $K = 6$ dB and different values of γ when $P = 1$ and $\sigma = 1$.

be formulated as a quadratically constrained quadratic program (QCQP) which in general does not have a closed-form solution. However, the problem can be numerically solved as shown in [59], using techniques such as second order cone programming (SOCP), or semi definite programming (SDP).

2.4.1 Perfect CSIT

Here, we examine the case of perfect CSIT and find the maximum received power at the Rx while the interference power at Rx_1 is limited to the threshold $\epsilon = \alpha P$ $(\alpha \geq 0)$.

Problem Formulation

When the Tx has perfect knowledge of the channel gain vectors $\mathbf{g}$ and $\mathbf{h}_1$, the optimization problem is formulated as (**P5**)

$$G^{\text{th}} = \max \quad |\mathbf{x}^T\mathbf{g}|^2 \tag{2.32}$$

$$\text{subject to:} \quad |\mathbf{x}^T\mathbf{h}|^2 \leq \epsilon \tag{2.33}$$

$$\|\mathbf{x}\|^2 \leq P. \tag{2.34}$$

where the optimization (2.32) is over the beamforming vector $\mathbf{x}$. Therefore, the focus is to find a suitable beamforming vector $\mathbf{x}$ that with the power limited to P (according to (2.34)) satisfies the interference constraint (2.33) and at the same time, maximizes the received power at the Rx.

Solving the Optimization Problem

To solve **P5**, we apply a rotation matrix $\mathbf{U} \in \mathbb{C}^{n \times n}$ that rotates the channel gain vector $\mathbf{h}$ to $\|\mathbf{h}\|\mathbf{e_1}$, where $\mathbf{e_1}$ is the unit vector in the direction of the first coordinate of $\mathbf{h}$. In other words, we have $\mathbf{Uh} = (\|\mathbf{h}\|, 0, \ldots, 0)$. Thus, we can express the constraint $|\mathbf{x}^T\mathbf{h}|^2 \leq \epsilon$ in (2.33) as $|\mathbf{x}^T\mathbf{h}|^2 = |(\mathbf{x}^T\mathbf{U}^{-1})(\mathbf{Uh})|^2 \leq \epsilon$. Defining $\mathbf{x}^T\mathbf{U}^{-1} = \mathbf{y}^T$, where $\mathbf{y} = (y_1, \ldots, y_n)^T$ is a new coordinate system, we obtain $|\mathbf{x}^T\mathbf{h}|^2 = \|\mathbf{h}\|^2|y_1|^2 \leq \epsilon$.

Consequently, applying a change of coordinates according to the rotation $\mathbf{U}$, we get a new formulation for **P5** in terms of $\mathbf{y}$ expressed as (**P6**)

$$G^{\text{th}} = \max \quad |\mathbf{y}^T\tilde{\mathbf{g}}|^2$$

$$\text{subject to:} \quad \|\mathbf{h}\|^2|y_1|^2 \leq \epsilon,$$

$$\|\mathbf{y}\|^2 \leq P.$$

In **P6**, the vector $\tilde{\mathbf{g}} = \mathbf{Ug}$ is a rotated vector and we have $\tilde{\mathbf{g}} \sim \mathcal{CN}(\mathbf{0}, 2\sigma^2\mathbf{I}_n)$. Furthermore, the entries of $\tilde{\mathbf{g}}$ are independent of the entries of $\mathbf{h}$. **P6** implies that only y_1 contributes to the interference power at Rx_1. Using $|y_1|$ as a slack variable, we can convert

P6 to the equivalent problem (**P7**)

$$G^{\text{th}} = \max \quad |\mathbf{y}^T \tilde{\mathbf{g}}|^2 \tag{2.35}$$

$$\text{subject to:} \quad |y_1| \leq \frac{\sqrt{\epsilon}}{\|\mathbf{h}\|},$$

$$\|\mathbf{y_{-1}}\|^2 \leq P - |y_1|^2. \tag{2.36}$$

Using the triangle inequality and Cauchy-Schwarz inequality, we can write $|\mathbf{y}^T \tilde{\mathbf{g}}|$ in (2.35) as

$$|\mathbf{y}^T \tilde{\mathbf{g}}| \leq |y_1 \tilde{g}_1| + |\mathbf{y_{-1}}^T \tilde{\mathbf{g}}_{-1}| \leq |y_1||\tilde{g}_1| + \|\mathbf{y_{-1}}\| \|\tilde{\mathbf{g}}_{-1}\|. \tag{2.37}$$

The upper bound on the right hand side of (2.37) can be achieved when $\arg(y_1) = -\arg(\tilde{g}_1)$ and $\mathbf{y_{-1}} = \kappa \tilde{\mathbf{g}}^*_{-1}$, where κ is a constant. Furthermore, the upper bound can be maximized when the constraint (2.36) is tight. Therefore, the objective function in (2.35) is written as

$$f(\mathbf{y}) = |\mathbf{y}^T \tilde{\mathbf{g}}|^2 = \left(|y_1||\tilde{g}_1| + \sqrt{P - |y_1|^2} \, \|\tilde{\mathbf{g}}_{-1}\| \right)^2.$$

Denoting the optimal solution to **P7** as $\mathbf{y}^{\text{opt}}$, we thus obtain $\mathbf{y_{-1}}^{\text{opt}} = \frac{\tilde{\mathbf{g}}^*_{-1}}{\|\tilde{\mathbf{g}}_{-1}\|} \sqrt{P - |y_1|^2}$ (by making (2.36) tight) and in the sequel, we find $|y_1^{\text{opt}}|$ by maximizing $f(\mathbf{y})$ over $|y_1| \leq \frac{\sqrt{\epsilon}}{\|\mathbf{h}\|}$.

Maximization of $f(\mathbf{y})$ over $|y_1|$ yields the two following cases:

1. **Case 1; when** $\|\mathbf{h}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}$: In this case, we obtain $|y_1^{\text{opt}}| = \frac{|\tilde{g}_1|}{\|\tilde{\mathbf{g}}\|} \sqrt{P}$. Thus, the optimal solution to **P7** becomes $\mathbf{y}^{\text{opt}} = \frac{\tilde{\mathbf{g}}^*}{\|\tilde{\mathbf{g}}\|} \sqrt{P}$ and consequently, the maximum received power at the Rx is found to be $G^{\text{th}} = P \|\tilde{\mathbf{g}}\|^2$.

2. **Case 2; when** $\|\mathbf{h}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}$: In this case, we obtain $|y_1^{\text{opt}}| = \frac{\sqrt{\epsilon}}{\|\mathbf{h}\|}$, and thus $\mathbf{y_{-1}}^{\text{opt}} = \frac{\tilde{\mathbf{g}}^*_{-1}}{\|\tilde{\mathbf{g}}_{-1}\|} \sqrt{P - \frac{\epsilon}{\|\mathbf{h}\|^2}}$. Therefore, the maximum received power at the Rx is found to be $G^{\text{th}} = \left(\frac{\sqrt{\epsilon}|\tilde{g}_1|}{\|\mathbf{h}\|} + \sqrt{P - \frac{\epsilon}{\|\mathbf{h}\|^2}} \, \|\tilde{\mathbf{g}}_{-1}\| \right)^2$.

To obtain the solution to the original optimization problem **P5**, we can use the fact that $\mathbf{x} = \mathbf{U}^T \mathbf{y}$, and find $\mathbf{x}^{\text{opt}} = \mathbf{U}^T \mathbf{y}^{\text{opt}}$.

Maximum Received Power at the Intended Receiver Rx

Considering the two different cases for the optimal solution, the maximum received power at the Rx can be expressed compactly as

$$G^{\text{th}} = P\|\tilde{\mathbf{g}}\|^2 \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 \le \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}} + \left(\frac{|\tilde{g}_1|\sqrt{\epsilon}}{\|\mathbf{h}\|} + \sqrt{P - \frac{\epsilon}{\|\mathbf{h}\|^2}}\|\tilde{\mathbf{g}}_{-1}\|\right)^2 \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}}. \quad (2.38)$$

Since $\mathbf{1}_{\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}} \ne 0$ implies that $\epsilon < P\|\mathbf{h}\|^2|\tilde{g}_1|^2/\|\tilde{\mathbf{g}}\|^2 < P\|\mathbf{h}\|^2$, we have the expansion

$$\left(\sqrt{P - \frac{\epsilon}{\|\mathbf{h}\|^2}}\right) \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}} = \sqrt{P}\left(1 - \frac{\epsilon}{2\|\mathbf{h}\|^2 P} - O\left(\left(\frac{\epsilon}{P}\right)^2\right)\right) \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}}. \quad (2.39)$$

Thus, having P as a fixed constant and $\epsilon \ge 0$ as a variable, we can write the second term in (2.38) as

$$\left(\frac{|\tilde{g}_1|\sqrt{\epsilon}}{\|\mathbf{h}\|} + \sqrt{P - \frac{\epsilon}{\|\mathbf{h}\|^2}}\|\tilde{\mathbf{g}}_{-1}\|\right)^2 \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}} \quad (2.40)$$

$$= \left(P\|\tilde{\mathbf{g}}_{-1}\|^2 + \sqrt{\epsilon}\left(\frac{2|\tilde{g}_1|\|\tilde{\mathbf{g}}_{-1}\|}{\|\mathbf{h}\|}\sqrt{P}\right) - \epsilon\left(\frac{\|\tilde{\mathbf{g}}_{-1}\|^2 - |\tilde{g}_1|^2}{\|\mathbf{h}\|^2}\right) - O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right)\right) \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}}. \quad (2.41)$$

We can find the expected value of Gth in (2.38) using $\mathsf{E}[G^{\text{th}}] = \mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\mathbf{h}}\left[G^{\text{th}}|\tilde{\mathbf{g}}\right]\right]$. Therefore,

$$\mathsf{E}[G^{\text{th}}] = P\mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\mathbf{h}}\left[\|\tilde{\mathbf{g}}\|^2 \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 \le \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}}|\tilde{\mathbf{g}}\right]\right]$$

$$+ \mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\mathbf{h}}\left[\left(\frac{|\tilde{g}_1|\sqrt{\epsilon}}{\|\mathbf{h}\|} + \sqrt{P - \frac{\epsilon}{\|\mathbf{h}\|^2}}\|\tilde{\mathbf{g}}_{-1}\|\right)^2 \cdot \mathbf{1}_{\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right\}}|\tilde{\mathbf{g}}\right]\right]. \quad (2.42)$$

Using the expansion in (2.41) and knowing that the entries of $\tilde{\mathbf{g}}$ and $\mathbf{h}$ are independent,

we obtain

$$\begin{aligned}
\mathsf{E}[G^{\text{th}}] = {}& P\mathsf{E}_{\tilde{\mathbf{g}}}\left[\|\tilde{\mathbf{g}}\|^2 \mathsf{Pr}\left\{\|\mathbf{h}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}|\tilde{\mathbf{g}}\right\}\right] \\
& + P\mathsf{E}_{\tilde{\mathbf{g}}}\left[\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2 \mathsf{Pr}\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}|\tilde{\mathbf{g}}\right\}\right] \\
& + 2\sqrt{P\epsilon}\,\mathsf{E}_{\tilde{\mathbf{g}}}\left[|\tilde{g}_1|\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|\mathsf{E}_{\mathbf{h}}\left[\frac{\mathbf{1}_{\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\}}}{\|\mathbf{h}\|}|\tilde{\mathbf{g}}\right]\right] \\
& - \epsilon\mathsf{E}_{\tilde{\mathbf{g}}}\left[\left(\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2 - |\tilde{g}_1|^2\right)\mathsf{E}_{\mathbf{h}}\left[\frac{\mathbf{1}_{\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\}}}{\|\mathbf{h}\|^2}|\tilde{\mathbf{g}}\right]\right] \\
& - O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right).
\end{aligned}$$
(2.43)

The random variable $\|\mathbf{h}\|^2/\sigma^2$ is chi-square distributed with $2n$ degrees of freedom (see [74]). Therefore,

$$\mathsf{Pr}\left\{\|\mathbf{h}\|^2 > u\right\} = \left(1 + \frac{u}{2\sigma^2} + \frac{u^2}{(2\sigma^2)^2 2!} + \ldots + \frac{u^{n-1}}{(2\sigma^2)^{n-1}(n-1)!}\right)e^{-\frac{u}{2\sigma^2}},$$

where u is a positive real number (see [74]), and thus using the series expansion of $e^{-\frac{u}{2\sigma^2}}$ with $u = \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}$ for a fixed $\tilde{\mathbf{g}}$, we find

$$\mathsf{Pr}\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}|\tilde{\mathbf{g}}\right\} = 1 - \frac{1}{n!}\left(\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\right)^n\left(\frac{\epsilon}{2P\sigma^2}\right)^n + O\left(\left(\frac{\epsilon}{P}\right)^{n+1}\right).$$
(2.44)

Therefore, by (2.44), we can derive

$$\mathsf{E}_{\tilde{\mathbf{g}}}\left[\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2 \cdot \mathsf{Pr}\left\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}|\tilde{\mathbf{g}}\right\}\right] = \mathsf{E}\left[\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2\right] - \mathsf{E}\left[f_1\left(\tilde{\mathbf{g}}\right)\right]O\left(\left(\frac{\epsilon}{P}\right)^n\right),$$
(2.45)

and

$$\mathsf{E}_{\tilde{\mathbf{g}}}\left[\|\tilde{\mathbf{g}}\|^2 \cdot \mathsf{Pr}\left\{\|\mathbf{h}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}|\tilde{\mathbf{g}}\right\}\right] = \mathsf{E}\left[f_2\left(\tilde{\mathbf{g}}\right)\right]O\left(\left(\frac{\epsilon}{P}\right)^n\right),$$
(2.46)

where f_1, f_2 are suitable functions of the entries of $\tilde{\mathbf{g}}$.

Furthermore, knowing that $\|\mathbf{h}\|^2/\sigma^2$ is chi-square distributed and thus $\mathsf{E}\left[\frac{1}{\|\mathbf{h}\|}\right] = \frac{\sqrt{2}\Gamma(n+\frac{1}{2})}{\sigma\Gamma(n)(2n-1)}$

(see [74]), for a fixed $\tilde{\mathbf{g}}$, we can write

$$
\mathsf{E_h}\left[\frac{\mathbf{1}_{\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\}}}{\|\mathbf{h}\|}\Big|\tilde{\mathbf{g}}\right] = \mathsf{E}\left[\frac{1}{\|\mathbf{h}\|}\right] - \frac{2^{-n}}{\sigma\Gamma(n)}\int_0^{\frac{\epsilon}{P\sigma^2}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}} x^{n-\frac{3}{2}}e^{-\frac{x}{2}}\,dx
$$

$$
= \frac{\sqrt{2}\,\Gamma(n+\frac{1}{2})}{\sigma\Gamma(n)(2n-1)} - \frac{2^{-n}}{\sigma\Gamma(n)}\int_0^{\frac{\epsilon}{P\sigma^2}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}} x^{n-\frac{3}{2}}\left(1 + \sum_{n=1}^{\infty}\frac{(-1)^n(x^n)}{2^n}\right)\,dx,
$$

(2.47)

which after caclulating the integral on the right hand side of (2.47), results in

$$
\mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E_h}\left[\frac{\mathbf{1}_{\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\}}}{\|\mathbf{h}\|}\Big|\tilde{\mathbf{g}}\right]\right] = \frac{\sqrt{2}\,\Gamma(n+\frac{1}{2})}{\sigma\Gamma(n)(2n-1)} - O\left(\left(\frac{\epsilon}{P}\right)^{n-\frac{1}{2}}\right).
$$

(2.48)

Using the same approach as above, since $\mathsf{E}\left[\frac{1}{\|\mathbf{h}\|^2}\right] = \frac{1}{2(n-1)\sigma^2}$ (see [74]), we obtain

$$
\mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E_h}\left[\frac{\mathbf{1}_{\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\}}}{\|\mathbf{h}\|^2}\Big|\tilde{\mathbf{g}}\right]\right] = \frac{1}{2(n-1)\sigma^2} - O\left(\left(\frac{\epsilon}{P}\right)^{n-1}\right).
$$

(2.49)

Therefore, by (2.48) and (2.49), we can derive

$$
\mathsf{E}_{\tilde{\mathbf{g}}}\left[|\tilde{g}_1|\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|\mathsf{E}\left[\frac{\mathbf{1}_{\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\}}}{\|\mathbf{h}\|}\Big|\tilde{\mathbf{g}}\right]\right] = \frac{\sqrt{2}\,\Gamma(n+\frac{1}{2})}{\sigma\Gamma(n)(2n-1)}\mathsf{E}\left[|\tilde{g}_1|\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|\right]
$$

$$
- \mathsf{E}\left[f_3\left(\tilde{\mathbf{g}}\right)\right]O\left(\left(\frac{\epsilon}{P}\right)^{n-\frac{1}{2}}\right),
$$

(2.50)

and

$$
\mathsf{E}_{\tilde{\mathbf{g}}}\left[\left(\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2 - |\tilde{g}_1|^2\right)\mathsf{E}\left[\frac{\mathbf{1}_{\{\|\mathbf{h}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\tilde{g}_1|^2}\}}}{\|\mathbf{h}\|^2}\Big|\tilde{\mathbf{g}}\right]\right] = \frac{\mathsf{E}\left[\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2 - |\tilde{g}_1|^2\right]}{2(n-1)\sigma^2}
$$

$$
- \mathsf{E}\left[f_4\left(\tilde{\mathbf{g}}\right)\right]O\left(\left(\frac{\epsilon}{P}\right)^{n-1}\right),
$$

(2.51)

respectively, where f_3, f_4 are suitable functions of the entries of $\tilde{\mathbf{g}}$.

Thus, following from (2.43) and using (2.46)–(2.51), the expected value of Gth can be written as

$$
\mathsf{E}[G^{\text{th}}] = P\mathsf{E}\left[\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2\right] + \sqrt{2P}\epsilon\frac{\Gamma(n+\frac{1}{2})}{\sigma\Gamma(n)(n-\frac{1}{2})}\mathsf{E}\left[|\tilde{g}_1|\right]\mathsf{E}\left[\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|\right]
$$

$$
+ \epsilon\frac{\mathsf{E}\left[|\tilde{g}_1|^2\right] - \mathsf{E}\left[\|\tilde{\mathbf{g}}_{-\mathbf{1}}\|^2\right]}{2(n-1)\sigma^2} - O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right).
$$

(2.52)

The random variable, $\|\tilde{\mathbf{g}}_{-1}\|^2/\sigma^2$ is chi-square distributed and the random variable, $\|\tilde{\mathbf{g}}_{-1}\|/\sigma$ is chi distributed both with $2(n-1)$ degrees of freedom. Therefore, we have (see [74])

$$\mathsf{E}\left[\|\tilde{\mathbf{g}}_{-1}\|^2\right] = 2(n-1)\sigma^2$$

$$\mathsf{E}\left[\|\tilde{\mathbf{g}}_{-1}\|\right] = \sqrt{2}\sigma\frac{\Gamma(n-\frac{1}{2})}{\Gamma(n-1)}$$

$$\mathsf{E}\left[|\tilde{g}_1|^2\right] = 2\sigma^2$$

$$\mathsf{E}\left[|\tilde{g}_1|\right] = \sqrt{2}\,\Gamma\left(\frac{3}{2}\right)\sigma.$$

Thus, replacing the above quantities in (2.52), we obtain

$$\mathsf{E}[G^{\text{th}}] = 2P\sigma^2(n-1) + 2\sqrt{2P\epsilon}\,\sigma\left(\frac{\Gamma\left(\frac{3}{2}\right)\Gamma(n-\frac{1}{2})\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(n)\Gamma(n-1)\left(n-\frac{1}{2}\right)}\right) - \epsilon\left(\frac{n-2}{n-1}\right) - O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right),$$

$$= 2P\sigma^2(n-1) + 2\sqrt{2\alpha}P\sigma\left(\frac{\Gamma\left(\frac{3}{2}\right)\Gamma(n-\frac{1}{2})\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(n)\Gamma(n-1)\left(n-\frac{1}{2}\right)}\right) - \alpha P\left(\frac{n-2}{n-1}\right) - O\left(\alpha^{\frac{3}{2}}\right),$$

$$(2.53)$$

where $\epsilon = \alpha P$. Note that each term in $\mathsf{E}[G^{\text{th}}]$ above is deliberately chosen to be the product of P and a unitless quantity. These unitless quantities in turn are expressed in terms of the unitless ratio $\alpha = \epsilon/P$, to show explicitly how increasing the tolerable interference ratio α affects $\mathsf{E}[G^{\text{th}}]$.

Performance Evaluation

Summarizing the results obtained in this section in the case of perfect CSIT, we can point out the following facts:

1. When the interference threshold is zero ($\alpha = 0$) as in Section 2.3.1, the result in (2.53) reduces to the result in (2.17) (note that $K = 0$ and $m = 1$ in this section). Thus, the first term in (2.53), which corresponds to such null-steering beamforming approach, is referred to as the *null-steering result*.

2. If we exclude the higher order terms in (2.53) (the terms of order $\alpha^{\frac{3}{2}}$ and higher), we obtain a second order approximation of $\mathsf{E}[G^{\text{th}}]$ in $\sqrt{\alpha}$ (throughout this chapter, second order approximation refers to second order approximation in $\sqrt{\alpha}$).

3. Fig. 2.4 plots the mean received power $\mathsf{E}[G^{\mathrm{th}}]$ in (2.42) (the exact value) and its second order approximation versus n for $P = 1$ and $\sigma = 1$. Fig. 2.5 plots $\mathsf{E}[G^{\mathrm{th}}]$ in (2.42) and its second order approximation versus α for $P = 1$ and $\sigma = 1$. As shown in the figures, the second order approximations track well the exact values. Therefore, the second order approximation of $\mathsf{E}[G^{\mathrm{th}}]$ is accurate for α at least as large as 0.5.

4. As n increases, the terms $\frac{\Gamma\left(n-\frac{1}{2}\right)\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(n)\Gamma(n-1)\left(n-\frac{1}{2}\right)}$ and $\frac{n-2}{n-1}$ in (2.53), both converge to 1. Particularly, for $n = 10$, we have $\frac{\Gamma\left(n-\frac{1}{2}\right)\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(n)\Gamma(n-1)\left(n-\frac{1}{2}\right)} = 0.97$ and $\frac{n-2}{n-1} = 0.89$. Therefore, for large n, the second order approximation of $\mathsf{E}[G^{\mathrm{th}}]$ can be written as

$$\mathsf{E}[G^{\mathrm{th}}] \simeq 2P\sigma^2(n-1) + 2\sqrt{2\alpha}\, P\sigma\, \Gamma\left(\frac{3}{2}\right) - \alpha P. \tag{2.54}$$

We are interested in finding tradeoffs between the interference threshold, $\epsilon = \alpha P$, and the mean received power at the Rx, i.e., $\mathsf{E}[G^{\mathrm{th}}]$. In other words, we aim to study the improvement in $\mathsf{E}[G^{\mathrm{th}}]$ as α grows slightly larger than zero. In (2.54), since n is sufficiently large, for small α the first term is clearly dominant. Thus, a small increase in α from zero results in a slight relative increase in $\mathsf{E}[G^{\mathrm{th}}]$. Therefore, when the number of antennas is large, a nonzero interference threshold does not lead to a significant gain in $\mathsf{E}[G^{\mathrm{th}}]$ compared to the null-steering result. This fact can also be observed in Fig. 2.4 as the gap between the the performance for $\alpha = 0$ (null-steering result) and the performance corresponding to different values of α decreases as n increases.

5. On the other hand, for smaller values of n, since the first term in the second order approximation of (2.53) is not dominant, an increase in α from zero does lead to a moderate increase in $\mathsf{E}[G^{\mathrm{th}}]$. As shown in Fig. 2.5, for $n = 2$, $n = 4$, and $n = 6$, the change of α from 0 to 0.1 leads to an increase of 1.2 dB (31%), 0.5 dB (11%), and 0.3 dB (6%) in $\mathsf{E}[G^{\mathrm{th}}]$ respectively.

2.4.2 Imperfect CSIT

In the realistic case of imperfect CSIT, we assume that the Tx ignores or is ignorant of the existence of estimation errors. Therefore, beamforming weights are computed using

the channel estimates. Subsequently, the actual received power at the Rx, $G^{\text{th, Im}}$, and the actual interference power at Rx_1, $I^{\text{th, Im}}$, are derived and then statistically analyzed.

Problem Formulation

Since only the estimated channel gains are available at the Tx, the optimization problem considered in this section has the same formulation as **P5** in Section 2.4.1 but with the estimated channel gains instead. With ML estimation [39], the estimated channel gain vectors $\hat{\mathbf{g}}$ and $\hat{\mathbf{h}}$ can be expressed as

$$\hat{\mathbf{g}} = \mathbf{g} + \mathbf{w}, \quad \hat{\mathbf{h}} = \mathbf{h} + \mathbf{v}, \tag{2.55}$$

where $\mathbf{g}$ and $\mathbf{h}$ are the actual channel gain vectors, and $\mathbf{w}$ and $\mathbf{v}$, are the estimation error vectors which are distributed as $\mathcal{CN}(\mathbf{0}, 2\sigma_e^2 \mathbf{I}_n)$ and are independent. Furthermore, we assume that the actual channel gains and their respective estimation errors are independent.

Therefore, we obtain **P8** as

$$
\begin{aligned}
\max \quad & |\mathbf{x}^T \hat{\mathbf{g}}|^2 \\
\text{subject to:} \quad & |\mathbf{x}^T \hat{\mathbf{h}}|^2 \leq \epsilon, \\
& \|\mathbf{x}\|^2 \leq P.
\end{aligned}
\tag{2.56}
$$

Solving the Optimization Problem

To solve **P8**, we can follow the same approach as in Section 2.4.1 and similarly, exploit $\mathbf{U}'$ as a rotation matrix to obtain an equivalent optimization problem as follows (**P9**)

$$
\begin{aligned}
\max \quad & |\mathbf{y}^T \tilde{\hat{\mathbf{g}}}|^2 \\
\text{subject to:} \quad & \|\hat{\mathbf{h}}\|^2 |y_1|^2 \leq \epsilon, \\
& \|\mathbf{y}\|^2 \leq P.
\end{aligned}
\tag{2.57}
$$

Similarly to Section 2.4.1, solving **P9** yields the following two cases:

1. **Case 1; when** $\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}$: In this case, we obtain the optimal solution to **P9** as $\mathbf{y}^{\mathsf{opt}} = \frac{\tilde{\hat{\mathbf{g}}}^*}{\|\tilde{\hat{\mathbf{g}}}\|} \sqrt{P}$.

2. **Case 2; when** $\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}$: In this case, we obtain $|y_1^{\mathsf{opt}}| = \frac{\sqrt{\epsilon}}{\|\hat{\mathbf{h}}\|}$, and $\mathbf{y}_{-1}^{\mathsf{opt}} = \frac{\tilde{\hat{\mathbf{g}}}_{-1}^*}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|} \sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}$.

Actual Interference Power at the Unintended Receiver Rx_1

Denote the optimal solution to **P8** as $\mathbf{x}^{\mathsf{opt}}$. Then, using (2.55), the actual interference power at Rx_1 can be written as

$$I^{\text{th, Im}} = |\mathbf{h}^T\mathbf{x}^{\mathsf{opt}}|^2 = |\hat{\mathbf{h}}^T\mathbf{x}^{\mathsf{opt}} - \mathbf{v}^T\mathbf{x}^{\mathsf{opt}}|^2 = |\|\hat{\mathbf{h}}\|y_1^{\mathsf{opt}} - \tilde{\mathbf{v}}^T\mathbf{y}^{\mathsf{opt}}|^2$$

$$= \|\hat{\mathbf{h}}\|^2|y_1^{\mathsf{opt}}|^2 + |\tilde{\mathbf{v}}^T\mathbf{y}^{\mathsf{opt}}|^2 - \|\hat{\mathbf{h}}\|y_1^{\mathsf{opt}}\tilde{\mathbf{v}}^\dagger\mathbf{y}^{\mathsf{opt}*} - \|\hat{\mathbf{h}}\|y_1^{\mathsf{opt}*}\tilde{\mathbf{v}}^T\mathbf{y}^{\mathsf{opt}}, \qquad (2.58)$$

where $\tilde{\mathbf{v}}$ is the rotated version of $\mathbf{v}$ and thus it is also distributed as $\mathcal{CN}(\mathbf{0}, 2\sigma_e^2\mathbf{I}_n)$.

The first term in (2.58) is equal to the interference threshold $\epsilon = \alpha P$ because (2.57) is tight in the constraint. The remaining terms in (2.58) are introduced by the estimation error $\mathbf{v}$. Calculating the expected value of the actual interference power $I^{\text{th, Im}}$ (see Appendix B), for $\epsilon = \alpha P$, we find

$$\mathsf{E}[I^{\text{th, Im}}] = 2P\sigma_e^2 + \epsilon - O\left(\left(\frac{\epsilon}{P}\right)^n\right)$$

$$= 2P\sigma_e^2 + \alpha P - O\left(\alpha^n\right). \qquad (2.59)$$

Actual Received Power at the Intended Receiver Rx

Using (2.55), the actual received power at the intended receiver Rx is

$$G^{\text{th, Im}} = |\mathbf{g}^T\mathbf{x}^{\mathsf{opt}}|^2 = |\hat{\mathbf{g}}^T\mathbf{x}^{\mathsf{opt}} - \mathbf{w}^T\mathbf{x}^{\mathsf{opt}}|^2 = |\tilde{\hat{\mathbf{g}}}^T\mathbf{y}^{\mathsf{opt}} - \tilde{\mathbf{w}}^T\mathbf{y}^{\mathsf{opt}}|^2$$

$$= |\tilde{\hat{\mathbf{g}}}^T\mathbf{y}^{\mathsf{opt}}|^2 + |\tilde{\mathbf{w}}^T\mathbf{y}^{\mathsf{opt}}|^2 - \tilde{\hat{\mathbf{g}}}^T\mathbf{y}^{\mathsf{opt}}\tilde{\mathbf{w}}^\dagger\mathbf{y}^{\mathsf{opt}*} - \tilde{\hat{\mathbf{g}}}^\dagger\mathbf{y}^{\mathsf{opt}*}\tilde{\mathbf{w}}^T\mathbf{y}^{\mathsf{opt}}, \qquad (2.60)$$

where $\tilde{\mathbf{w}}$ is the rotated version of $\mathbf{w}$ and thus it is also distributed as $\mathcal{CN}(\mathbf{0}, 2\sigma_e^2\mathbf{I}_n)$.

Calculating the expected value of the actual received power $G^{\text{th, Im}}$ (see Appendix C), for $\epsilon = \alpha P$, we find

$$\mathsf{E}[G^{\text{th, Im}}] = Q + \sqrt{\alpha}R + \alpha T + O\left(\alpha^{\frac{3}{2}}\right), \tag{2.61}$$

where

$$Q = 2P(n-1)\sigma^2 \left(1 - \frac{1}{\gamma} + \frac{1}{\gamma(1+\gamma)} + \frac{1}{(n-1)(1+\gamma)}\right), \tag{2.62}$$

$$R = 2\sqrt{2}P\sigma \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(n-\frac{1}{2}\right)\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(n)\Gamma(n-1)\left(n-\frac{1}{2}\right)}\left(\frac{\gamma}{1+\gamma}\right)^{\frac{3}{2}}, \tag{2.63}$$

$$T = P\left(-1 + \frac{1}{n-1} - \frac{1}{(1+\gamma)^2} - \frac{\gamma}{(n-1)(1+\gamma)^2} + \frac{2n-3}{(n-1)(1+\gamma)}\right), \tag{2.64}$$

and $\gamma = \sigma^2/\sigma_e^2$ is the signal to estimation error ratio (SER). Note that in the case of no channel estimation error $(\gamma = \infty)$, (2.61) is equal to (2.53).

Performance Evaluation

Summarizing the results obtained in this section in the case of imperfect CSIT, we can point out the following facts:

1. In (2.59), the first term is equal to the actual mean interference power at Rx_1 for null-steering beamforming $(\alpha = 0)$ as derived in Section 2.3. In the case of a relatively small α, for sufficiently large n, the first two terms in (2.59) which are independent of n, become dominant. Therefore, similarly to the null-steering scenario, the Tx can employ more antennas to improve the intended link's performance (due to spatial diversity) without increasing the average interference power at the unintended receiver.

2. The empirical average of the actual interference at the unintended receiver (the exact $\mathsf{E}[I^{\text{th, Im}}]$) and the second order approximation of $\mathsf{E}[I^{\text{th, Im}}]$ in (2.59) are plotted versus n in Fig. 2.6. As shown, the approximation tracks well the exact value and is effectively insensitive to n. Even though the acceptable interference threshold is

$\epsilon = \alpha P$, because of channel estimation inaccuracy, $\mathsf{E}[I^{\text{th, Im}}]$ has an extra dominant term, according to (2.59), which is proportional to the power of the estimation error. Therefore, more accurate estimation of the channel gains result in interference power closer to the intended threshold ϵ.

3. The mean received power at the Rx found with perfect CSIT in (2.53) is an upper bound for the actual received power at the Rx derived in (2.61). In other words, as $\gamma \to \infty$, (2.61) converges to (2.53). Fig. 2.7 plots $\mathsf{E}[G^{\text{th, Im}}]$ and the second order approximation of the analytic result in (2.61) for $\alpha = 0$ versus n for different values of γ, when $P = 1$ and $\sigma = 1$. As shown in the figure, the analytic result tracks well the exact value. Furthermore, for a fixed n, as γ increases, $\mathsf{E}[G^{\text{th, Im}}]$ increases to approach the upper bound (the w/o error performance).

4. As n increases, the second order approximation of (2.61) converges to

$$\mathsf{E}[G^{\text{th, Im}}] \simeq \frac{2P\sigma^2}{1+\gamma} + 2P\sigma^2 \left(1 - \frac{1}{\gamma} + \frac{1}{\gamma(1+\gamma)} \right) (n-1)$$
$$+ 2\sqrt{2\alpha} P \sigma \, \Gamma \left(\frac{3}{2} \right) \left(\frac{\gamma}{1+\gamma} \right)^{\frac{3}{2}} - \alpha P \left(1 + \frac{1}{(1+\gamma)^2} \right). \qquad (2.65)$$

Fig. 2.8 and Fig. 2.9 plot the exact $\mathsf{E}[G^{\text{th, Im}}]$ in (2.61) and its second order approximation versus n for different values of γ and α when $P = 1$ and $\sigma = 1$. As shown in the figures, the second order approximation tracks well the exact value in each case. Furthermore, according to these figures, for small n, $\mathsf{E}[G^{\text{th, Im}}]$ increases moderately as α increases but for larger values of n, the increase of α leads to a small relative increase in $\mathsf{E}[G^{\text{th, Im}}]$. In other words, with the increase of n, the gap between the performance for $\alpha = 0$ (the null-steering result) and the performance corresponding to different values of α decreases.

Fig. 2.10 and Fig. 2.11 plot the exact $\mathsf{E}[G^{\text{th, Im}}]$ in (2.61) and its second order approximation versus α for $n = 5$ and $n = 9$ respectively, when $P = 1$ and $\sigma = 1$. According to the figures, the second order approximation tracks well the exact $\mathsf{E}[G^{\text{th, Im}}]$ for different values of γ. Furthermore, $\mathsf{E}[G^{\text{th, Im}}]$ does not significantly improve with the increase in α, the improvements in the case of imperfect CSIT

are less compared to the perfect CSIT case, and the improvements decrease as the number of antennas increases.

2.5   Summary and Conclusions

In this chapter, we have studied beamforming in order to manage and limit the interference from a typical transmitter to some unintended receivers while improving the performance of the intended link between the transmitter and its own receiver. An optimal beamforming vector has been characterized for the both scenarios of null-steering beamforming and ϵ-threshold beamforming under the condition that the respective desired interference constraint is satisfied. Subsequently, closed form expressions for the average received power at the intended receiver are derived in both scenarios in the case of perfect CSIT and imperfect CSIT.

Intuitively, the allowance of a small nonzero interference at the unintended receivers as in the ϵ-threshold beamforming scenario, should improve the received power at the intended receiver. But the analysis in this thesis has shown that this enhancement is marginal and not worthwhile, notably in the case of imperfect CSIT. More specifically, when the interference threshold of the unintended receiver is a small value relative to the transmit power P, ϵ-threshold beamforming does not lead to a significant increases in the average received power at the intended receiver compared to the null-steering result. Therefore, there is no significant loss in the performance of the intended link if the transmitter performs null-steering beamforming instead of ϵ-threshold beamforming.

For the null-steering beamforming scenario, we have shown that the mean received power at the intended receiver grows with the number of antennas minus the number of unintended receivers. Therefore, it can be enhanced either by employing additional antennas or by reduction in the number of unintended receivers.

Thus, the transmitter can perform null-steering beamforming and use a large number of antennas to increase the received power at the intended receiver and boost its performance (due to spatial diversity) instead of relying on the interference threshold to improve the

performance. In this way, since the average interference power is only $2P\sigma_e^2$ and is independent of n, using more antennas does not significantly impact the unintended receivers. In fact, if the estimation error variance $2\sigma_e^2$ is small enough, the imposed average interference power may even be less than the interference threshold ϵ regulated by the unintended receiver (regardless of the number of antennas).

Furthermore, the mean received power at the intended receiver is directly proportional to the power of the line-of-sight component of the intended link for the null-steering scenario and as shown in Fig. 2.2, in the case of imperfect CSIT, a moderate line-of-sight component can significantly reduce the effect of estimation error on the performance of the intended link.

An application of null-steering beamforming can be for example in cognitive radio systems. In such systems, the secondary transmitter can employ multiple antennas and thus perform null-steering beamforming to nullify the co-channel interference at the primary receivers. Therefore, the concurrent operation of the secondary system in the same licensed frequency band does not degrade the performance of the primary system. Furthermore, by using additional secondary antennas, not only the performance of the primary system remains unchanged but also the performance of the secondary system is improved.

Figure 2.4: The exact value of $\mathsf{E}[G^{\text{th}}]$ in the case of perfect CSIT and its second order approximation versus n for different values of α when $P = 1$ and $\sigma = 1$. For different values of α from zero to 0.5, the approximation tracks well the exact value.

For a fixed n, as α increases, $\mathsf{E}[G^{\text{th}}]$ increases, i.e., intentionally allowing additional interference to Rx_1 provides additional received power at the Rx.

With the increase of n, the gap between the null-steering result (performance for $\alpha = 0$) and the performance corresponding to other values of α decreases, which implies that for large n, a nonzero interference threshold does not lead to a significant gain in $\mathsf{E}[G^{\text{th}}]$ compared to the null-steering result.

Figure 2.5: The exact value of $\mathsf{E}[G^{\mathrm{th}}]$ in the case of perfect CSIT and its second order approximation versus α for different values of n when $P = 1$ and $\sigma = 1$. The approximation tracks well the exact value.

$\mathsf{E}[G^{\mathrm{th}}]$ increases as α increases. Smaller n leads to a higher increase in $\mathsf{E}[G^{\mathrm{th}}]$. For $n = 2$, $n = 3$, $n = 4$, $n = 6$, and $n = 9$, the increase in $\mathsf{E}[G^{\mathrm{th}}]$ is 1.2 dB (31%), 0.7 dB (16%), 0.5 dB (11%), 0.3 dB (6%), and 0.2 dB (4%) respectively, when α increases from 0 to 0.1.
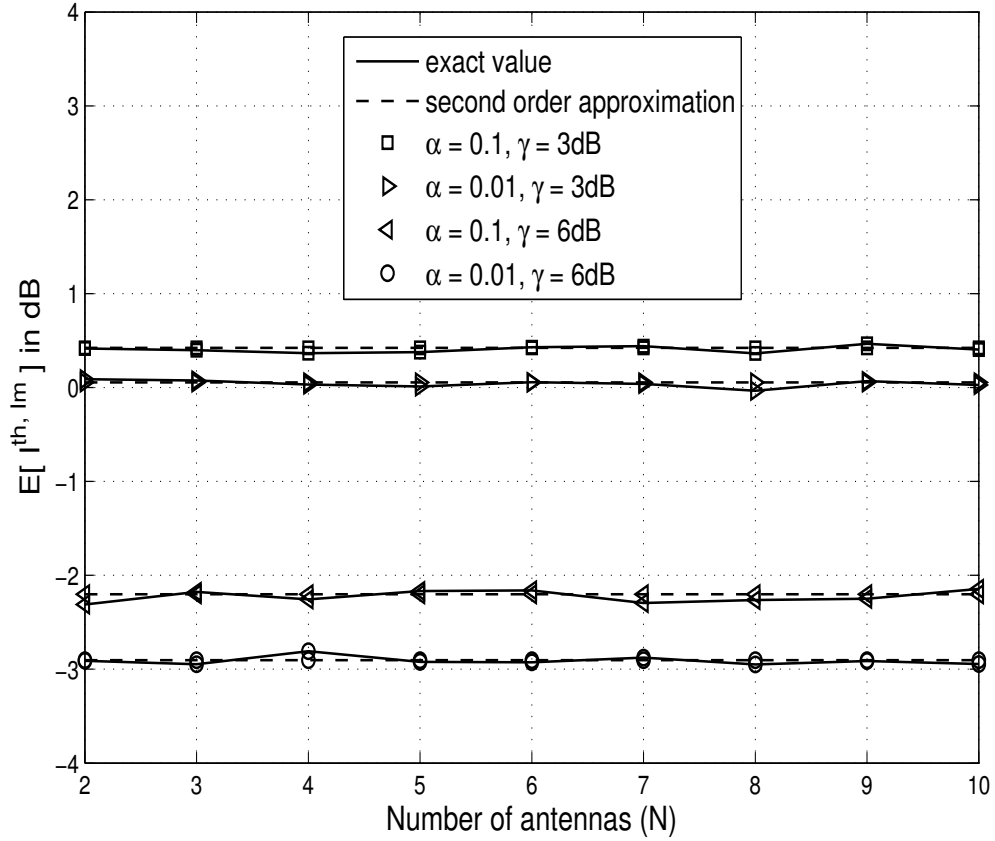
Figure 2.6: The exact value of $\mathsf{E}[I^{\text{th, Im}}]$ in dB and its second order approximation versus n for different values of α and γ when $P = 1$ and $\sigma = 1$. The second order approximation tracks well the exact value and it is insensitive to n.
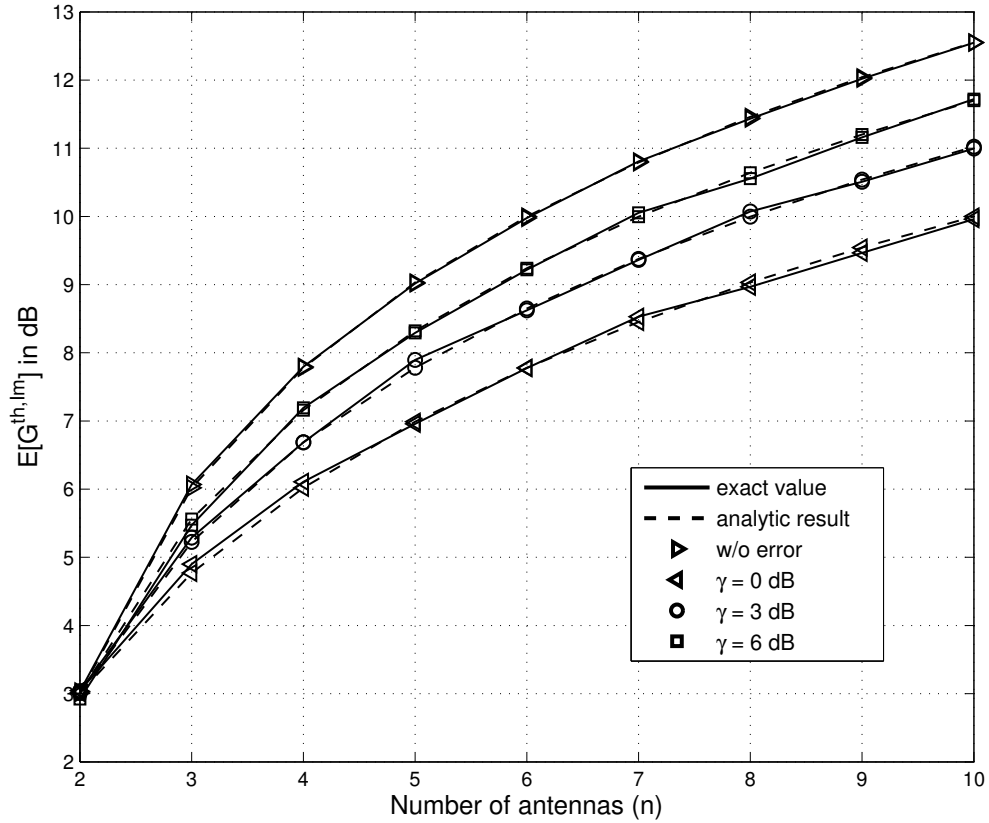
Figure 2.7: The exact value of $\mathsf{E}[G^{\text{th, Im}}]$ in the case of imperfect CSIT and its second order approximation versus n for different values of γ when $\alpha = 0$, $P = 1$, and $\sigma = 1$. The second order approximation tracks well the exact value.

For a fixed n, as γ increases, $\mathsf{E}[G^{\text{th, Im}}]$ increases to approach the upper bound (the w/o error performance).
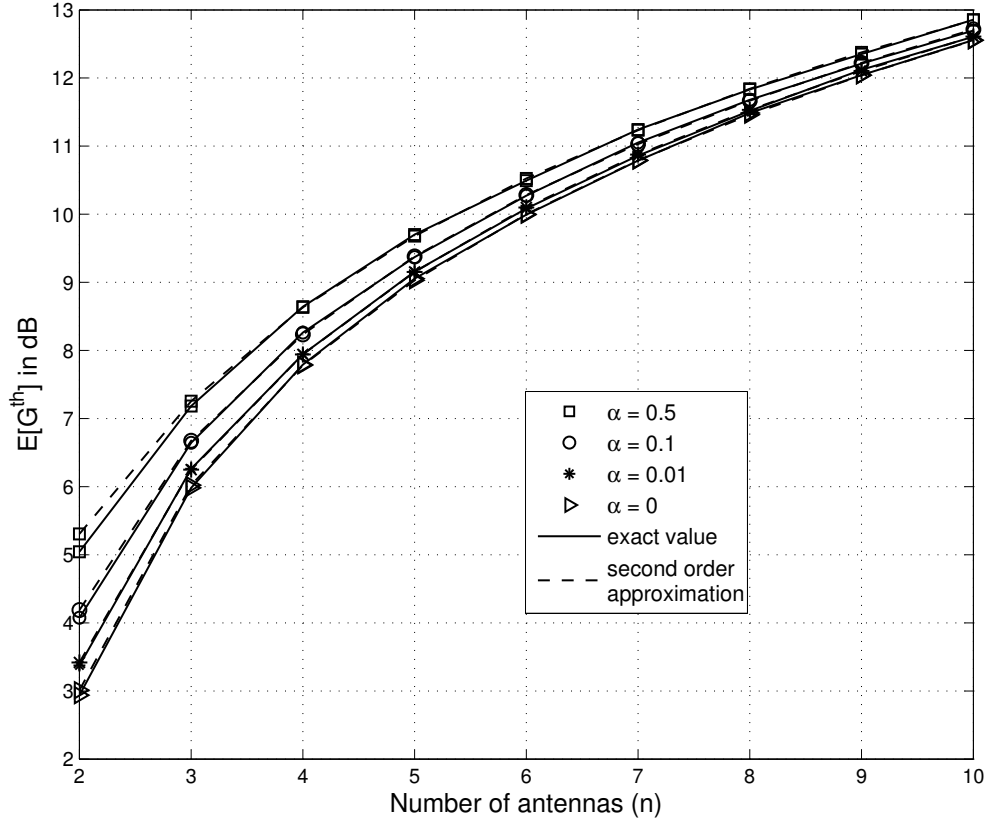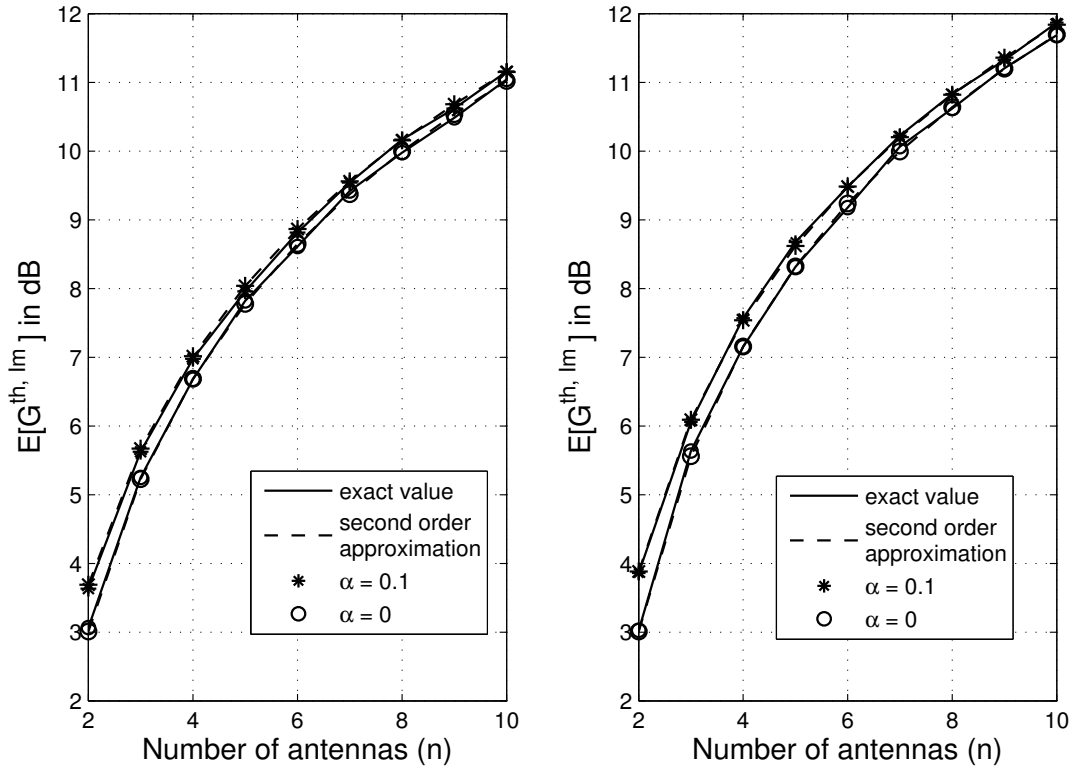
Figure 2.8: The exact value of $\mathsf{E}[G^{\mathrm{th,\ Im}}]$ in the case of imperfect CSIT and its second order approximation versus n for different values of α and γ when $P = 1$ and $\sigma = 1$. The second order approximation tracks well the exact value. Left: $\gamma = 3$ dB, Right: $\gamma = 6$ dB.
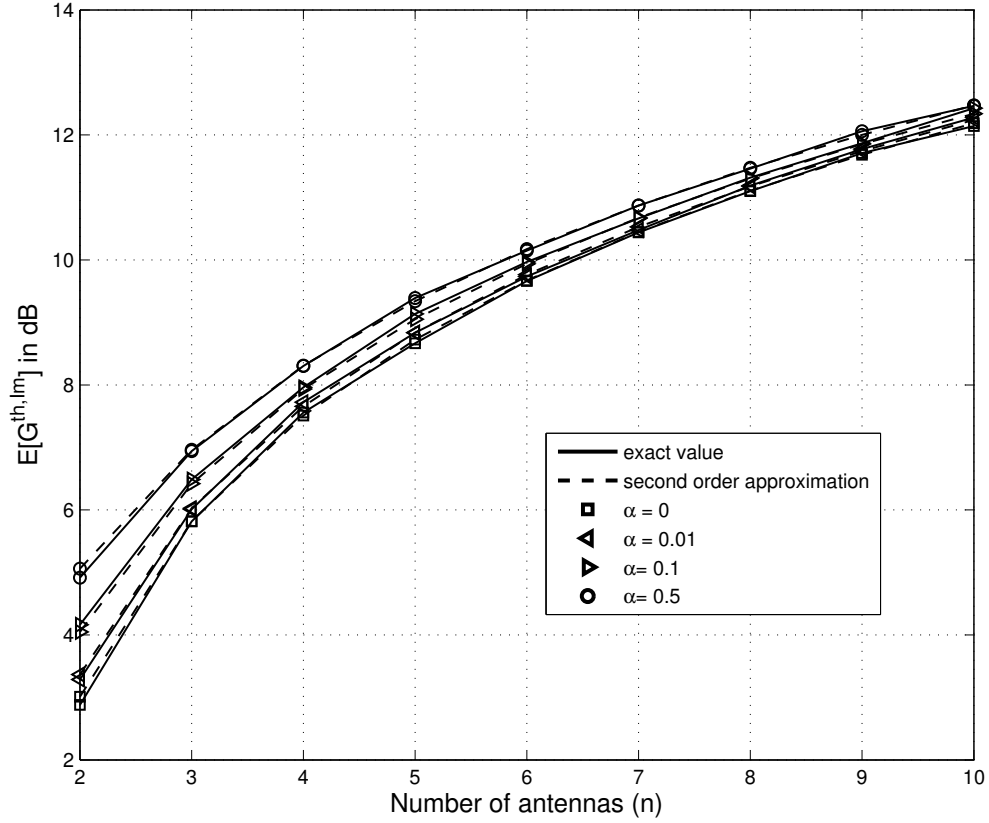
Figure 2.9: The exact value of $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ in the case of imperfect CSIT and its second order approximation versus n for different values of α when $\gamma = 10$ dB, $P = 1$, and $\sigma = 1$. The second order approximation tracks well the exact value.

For a fixed n, as α increases, $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ increases.

With the increase of n, the gap between the performance for $\alpha = 0$ (the null-steering result) and the performance corresponding to different values of α decreases.
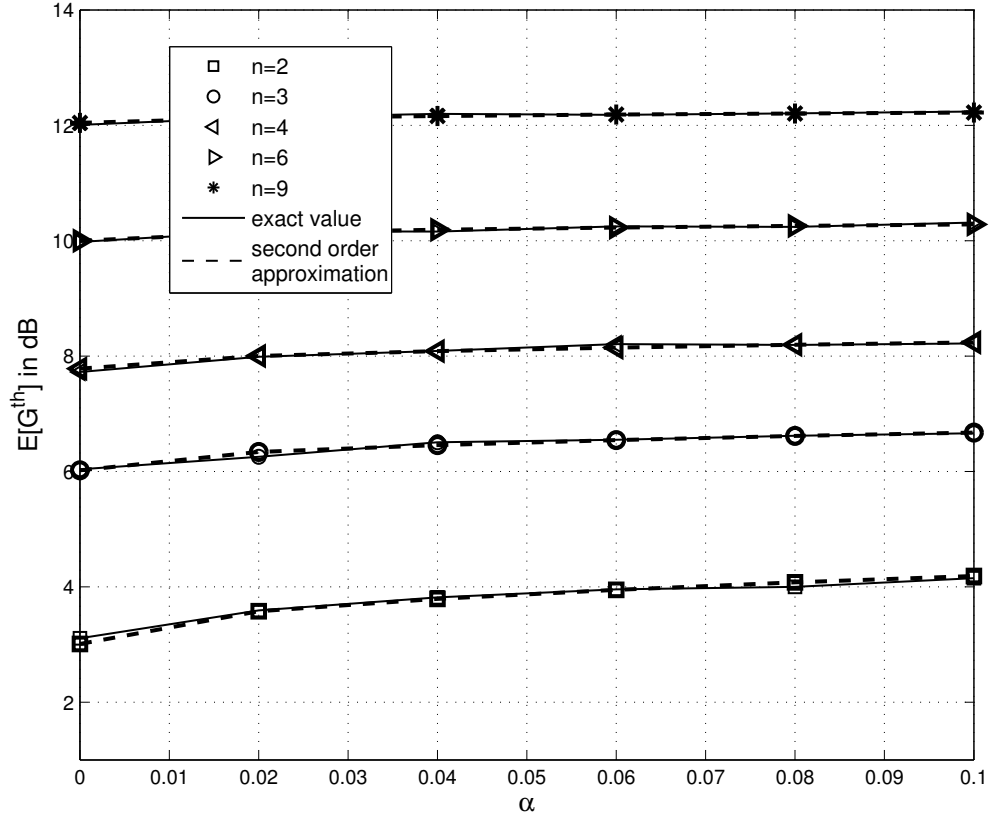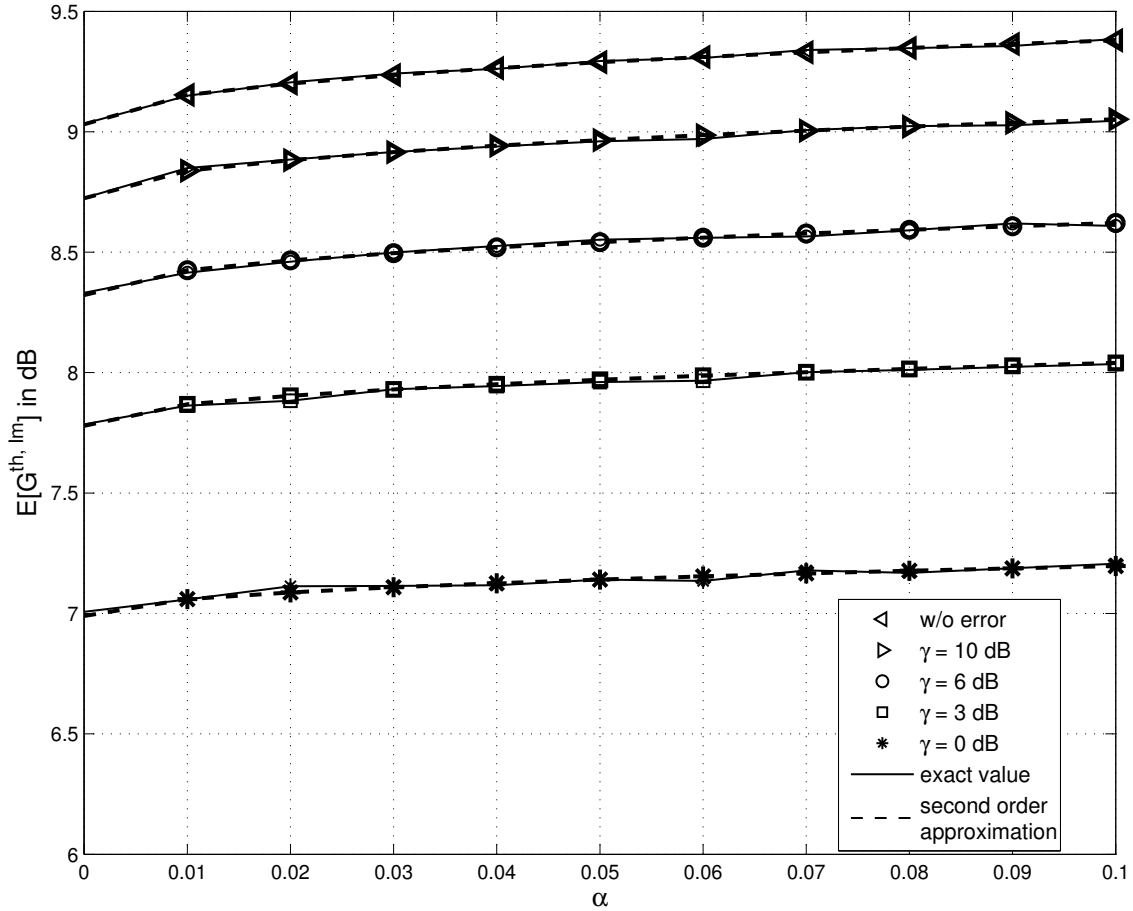
Figure 2.10: The exact value of $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ in the case of imperfect CSIT and its second order approximation versus α for different values of γ and $n = 5$, when $P = 1$ and $\sigma = 1$. The second order approximation tracks well the exact value of $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ for different values of γ. The exact $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ does not improve significantly with the increase in α. The improvements in the case of imperfect CSIT are less compared to the perfect CSIT case (w/o error performance).

61
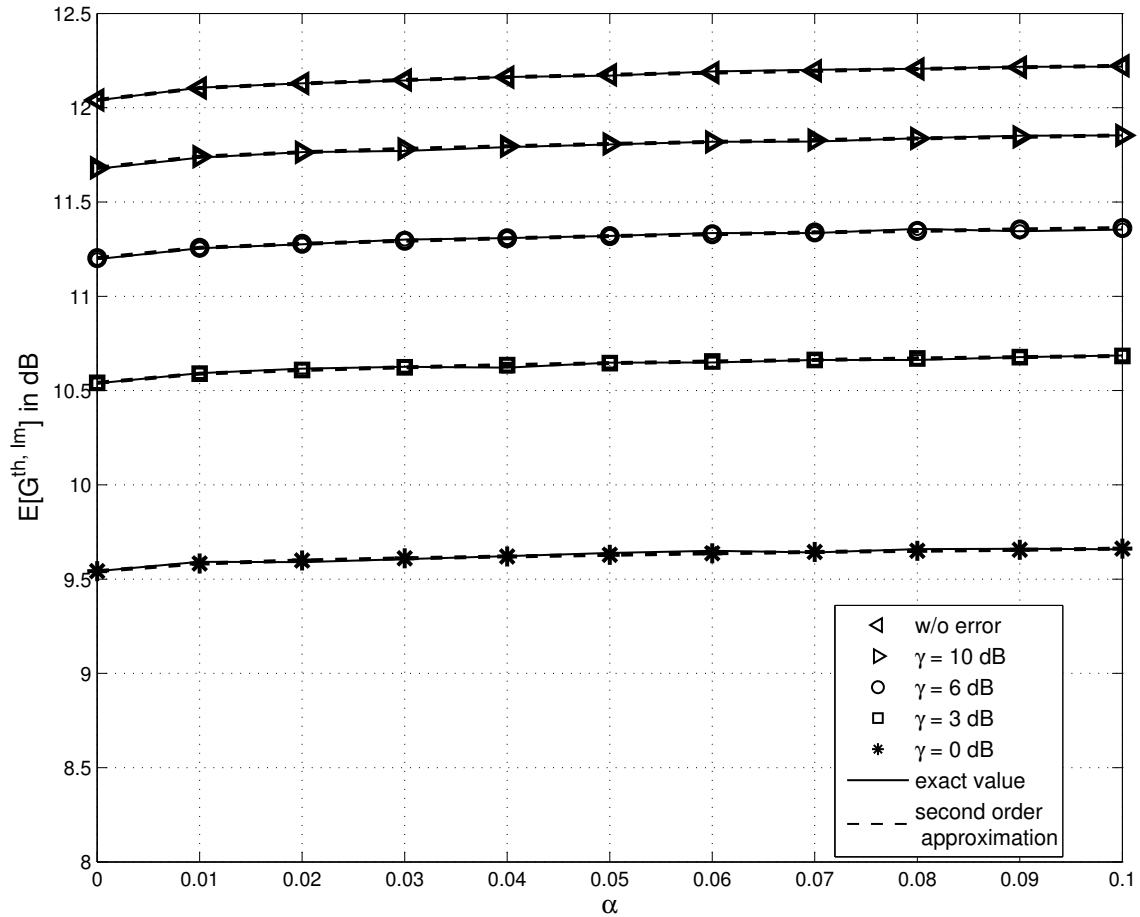
Figure 2.11: The exact value of $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ in the case of imperfect CSIT and its second order approximation versus α for different values of γ and $n = 9$, when $P = 1$ and $\sigma = 1$. The second order approximation tracks well the exact value of $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ for different values of γ. The exact $\mathsf{E}[G^{\mathrm{th,\,Im}}]$ does not improve significantly with the increase in α.

Chapter 3

Performance of Water-filling with Channel Estimation Error at Low SNR

3.1 Introduction

In this chapter[1], we address spatial multiplexing in MIMO systems and explore its capability to increase data rate (throughput) of wireless systems without increasing bandwidth or transmit power. As stated earlier, in a rich scattering environment, an $n \times n$ MIMO system can be decomposed into n parallel and independent channels over which multiple streams of data can be simultaneously sent within the bandwidth of operation. Thus, by multiplexing over these channels an n-fold increase in the data rate can be obtained [15, 16, 19].

With perfect CSIT, water-filling power allocation can achieve the maximum throughput of a MIMO system with very small probability of error (which is referred to as the capacity of the system) [16, 19]. When no CSIT is available, equal power allocation can be performed, i.e., the transmitter can distribute the transmit power equally among the

---

[1]The results presented in this chapter have been published in [72].

transmit antennas as this is the most reasonable way to allocate power in such a case [19].

In [19], having n transmit and n receive antennas in a rich scattering environment, the capacity of a MIMO system with perfect CSIT and the throughput with no CSIT at all were found using water-filling and equal power allocation, respectively and were shown to scale linearly with n as n grows large (intuitively due to spatial multiplexing). Note that MIMO systems with a large number of antennas (also known as Massive MIMO or Very Large MIMO) and their capacity scaling results have been under extensive investigation recently [66–68]. The constant multiplier associated with such linear scaling (the asymptotic growth rate) in [19], were derived by using the limiting distribution of the eigenvalues of the channel gain matrix as n grows large. It was shown in [19] that at low SNR, water-filling with perfect CSIT provides significant throughput improvement asymptotically over equal power allocation.

In reality, it requires significant effort to obtain CSIT within a reasonable accuracy, especially at low SNR and in addition, the transmitter may have no a priori reliable knowledge of the accuracy of channel estimates either. Thus, it is interesting and worthwhile to find the throughput achieved in MIMO systems if the transmitter performs water-filling based on erroneous channel estimates only and to compare such a throughput with the throughput achieved by water-filling based on perfect CSIT and thus to evaluate the performance loss as a function of the channel estimation error. In this regard, we derive the throughput achieved by water-filling with imperfect CSIT and show that at low SNR, it scales linearly with n as $n \to \infty$ with a proportionality constant that is not random, but is given by R. Then we compare R with C^{P} which is the asymptotic growth rate for the throughput of water-filling based on perfect channel gains (best-case scenario) as a function of the signal to estimation error ratio (SER).

The results obtained in this thesis indicate that at low SNR, for moderate values of the SER, water-filling based on erroneous channel estimates can still achieve significant throughputs asymptotically. In particular, for SER values such as 5 dB, 0 dB, and -5 dB, R is found to be 86%, 70%, and 52% of C^{P}, respectively. In addition, we show that, at low SER, water-filling based on erroneous channel estimates achieves the same throughput as

equal power allocation asymptotically in the low SNR regime.

Extensive research such as [79–83] has been conducted prior to this work to find the capacity of MIMO systems at low SNR. The authors in [79] found the capacity in the case that the transmitter has no CSI, whereas, [80], and [81] addressed the cases of perfect CSIT, and statistical CSIT as well. On the other hand, in works such as [82, 83], the capacity of MIMO systems was derived at low SNR, assuming both the transmitter and the receiver have no CSI. Note that in this chapter, we do not aim to find the capacity of a MIMO system such as in the works outlined above and furthermore, we do not claim that water-filling is optimal. Instead, we derive the throughput of a MIMO system if the transmitter performs water-filling (which is the optimal power allocation strategy with perfect CSIT) based on erroneous channel estimates only. We evaluate the asymptotic behaviour of such a throughput and compare it to the capacity with perfect CSIT in order to understand how much throughput is lost due to channel estimation error.

In [84], the authors derived a lower bound on the ergodic capacity of a MIMO system given the erroneous channel estimates and the power of the channel estimation error $(2\sigma_e^2)$ at the transmitter and the receiver (since the channel gains change in each time slot, an estimation of $2\sigma_e^2$ can be found based on observed channel gains in previous time slots). They showed that a strategy that achieves such a lower bound to be a modified water-filling over spatial (antenna) and temporal (fading) domains. In other words, at every time slot, new water-filling is performed based on new channel gains. In this thesis, unlike in [84], a quasi-static fading model is assumed as opposed to ergodic fading, i.e., the channel gains are fixed for the entire duration of transmission. Furthermore, we assume that the transmitter has no a priori knowledge of error variance $2\sigma_e^2$ and that it only knows the channel estimates based on which it performs water-filling.

As the SNR goes to 0, which is the very low SNR regime, and with perfect CSIT, the water-filling solution reduces to allocating all the transmit power to the strongest eigen-direction. This happens intuitively because all the other eigen-directions will be above the water-filling level and thus no transmit power will be allocated to them. Now, while the SNR under consideration here is low (which makes it practically sound), we do not

take the limit of SNR$\to 0$. Thus, in this case, more than one eigen-direction is used, i.e., the transmit power is distributed among multiple eigen-directions. Additionally, with no CSIT, the strongest eigen-direction is not even known, and all eigen-directions should be used. Indeed, in our scheme, as the SER becomes small enough, all the eigen-directions are in fact used (even at low SNR) and water-filling based on imperfect CSIT gives the same throughput as equal power allocation.

Note that, in the high SNR regime, as shown in [19], equal power allocation (which requires no CSIT) provides the same asymptotic throughput as water-filling with perfect CSIT. Hence, in that regime, it is preferable to employ equal power allocation instead of water-filling and thus studying the effect of channel estimation error on the performance of water-filling is of no interest.

3.2   System Model

Throughout this chapter, we consider the point-to-point MIMO model with flat-fading channels shown in Fig. 3.1. The system consists of a transmitter, Tx for short, and a receiver, Rx for short, each equipped with n antennas. We assume that the communication is one-way and is from Tx to the Rx in a rich scattering environment.

Denoting the complex baseband signal transmitted at the Tx by the $n \times 1$ vector $\mathbf{s}$, the received vector at the Rx can be expressed as

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{z},$$

where $\mathbf{z}$ is the $n \times 1$ additive white Gaussian noise vector at the Rx. We assume that the entries of $\mathbf{z}$ are i.i.d. and zero-mean CSCG-distributed with variance N, i.e., $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, N\mathbf{I}_n)$. For simplicity, we take $N = 1$. Furthermore, $\mathbf{H}$ is the channel gain matrix whose entries are i.i.d. and zero-mean CSCG-distributed with variance $2\sigma^2$, i.e., σ^2 per real dimension.

We assume that the channel is quasi-static fading, i.e., as far as the Tx is concerned, $\mathbf{H}$ is fixed for the duration of the transmission. As stated in Section 1.3, since known
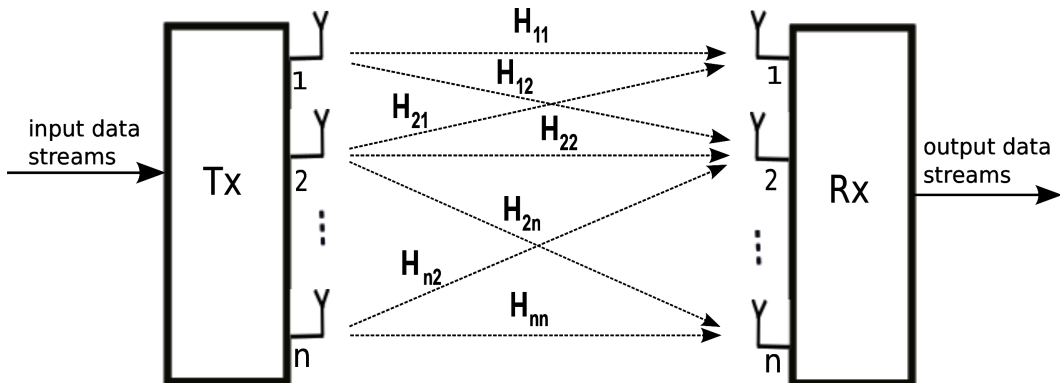
Figure 3.1:   MIMO system model under consideration; a multi-antenna transmitter, Tx, is in communication with a multi-antenna receiver, the Rx, in a rich scattering environment.

pilot sequences can be sent from the Tx to the Rx for channel estimation [5, Chap. 3.9], it is reasonable to assume that $\mathbf{H}$ is perfectly known at the Rx. If there is a feedback link available from the Rx to the Tx, the channel gain knowledge can be forwarded to the Tx. Also, in bi-directional systems that perform time division duplexing (TDD), using the reciprocity properties of the channel, the channel gain knowledge can be obtained at the Tx as well.

We denote the total transmit power over n antennas as P, i.e., $\mathsf{E}\left[\mathbf{s}^\dagger\mathbf{s}\right] \leq P$ which can be equivalently written as $\mathsf{Tr}(\mathbf{Q}) \leq P$ where $\mathbf{Q} = \mathsf{E}[\mathbf{s}\mathbf{s}^\dagger]$ is the transmit covariance matrix [16]. In addition, we measure the SNR as $\rho = P/N = P$.

In the MIMO system described above, since the channel gain matrix $\mathbf{H}$ is known at the Rx, the mutual information between the Tx and the Rx can be written as

$$I(\mathbf{s}; \mathbf{r}) = H(\mathbf{r}) - H(\mathbf{r}|\mathbf{s}) = H(\mathbf{r}) - H(\mathbf{z}),$$

where $H(\mathbf{r})$ is the entropy of the vector $\mathbf{r}$ [85]. For a fixed $\mathbf{H}$, among all input distributions of $\mathbf{s}$ with a given covariance matrix $\mathbf{Q}$, the Gaussian distribution $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$ maximizes $H(\mathbf{r})$ and thus $I(\mathbf{s}; \mathbf{r})$. Therefore, we obtain [16]

$$I(\mathbf{s}; \mathbf{r}) = \log\det\left[\mathbf{I}_n + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right], \tag{3.1}$$

which is equal to the data rate (throughput) that can be transmitted over the channel with arbitrarily small error probability. The mutual information in (3.1) depends on how the Tx chooses the input covariance matrix $\mathbf{Q}$ which is based on the knowledge of the channel gain matrix $\mathbf{H}$ at the Tx (i.e. CSIT). We consider the three different cases of no CSIT, perfect CSIT, and imperfect CSIT based on the availability of CSIT at the Tx.

3.3 Analysis of No CSIT and Perfect CSIT cases

3.3.1 Throughput with No CSIT

When the Tx has no knowledge of the channel gain matrix $\mathbf{H}$, we assume that it distributes the total transmit power P equally over n transmit antennas as this is the most reasonable way to allocate power in such a case [4]. Therefore, the input covariance matrix $\mathbf{Q}$ is equal to the normalized identity matrix, i.e., $\mathbf{Q} = P\mathbf{I}_n/n$. Thus, denoting $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ as the eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$, following from (3.1), the mutual information with equal power allocation is

$$I_n^{\mathrm{N}} = \log \det \left[\mathbf{I}_n + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right] = \log \det \left[\mathbf{I}_n + \frac{P\mathbf{H}\mathbf{H}^\dagger}{n}\right] = \sum_{i=1}^{n} \log \left(1 + \frac{P}{n}\lambda_i\right), \qquad (3.2)$$

where the logarithm is with respect to base 2. Therefore, I_n^{N} above is the throughput of the MIMO system with equal power allocation.

3.3.2 Capacity with Perfect CSIT

In this scenario, $\mathbf{H}$ is perfectly known at the Tx. Thus, based on the the knowledge of $\mathbf{H}$, the Tx can maximize the mutual information in (3.1) over the transmit covariance matrix $\mathbf{Q}$. The maximum mutual informarion obtained is referred to as the capacity of the MIMO system with perfect CSIT which can be derived as

$$C_n^{\mathrm{P}} = \max_{\mathbf{Q}:\mathsf{Tr}(\mathbf{Q})\leq P} \log \det \left[\mathbf{I}_n + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right].$$
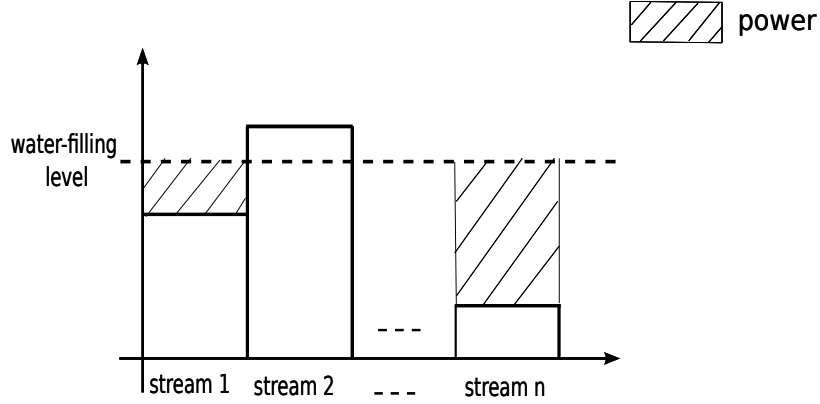
Figure 3.2: Water-filling power allocation

Let the singular value decomposition of the channel matrix $\mathbf{H}$ be $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}^{1/2}\mathbf{V}^{\dagger}$ with $\boldsymbol{\Sigma} = \text{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right)$. We can write

$$
\begin{aligned}
\log\det\left[\mathbf{I}_n + \mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}\right] &= \log\det\left[\mathbf{I}_n + \mathbf{U}\boldsymbol{\Sigma}^{1/2}\mathbf{V}^{\dagger}\mathbf{Q}\mathbf{V}\boldsymbol{\Sigma}^{1/2}\mathbf{U}^{\dagger}\right] \\
&= \log\det\left[\mathbf{I}_n + \boldsymbol{\Sigma}^{1/2}\mathbf{V}^{\dagger}\mathbf{Q}\mathbf{V}\boldsymbol{\Sigma}^{1/2}\right].
\end{aligned}
\tag{3.3}
$$

It is shown in [16] that in order to maximize (3.3) over $\mathbf{Q}$, the matrix $\mathbf{V}^{\dagger}\mathbf{Q}\mathbf{V} = \mathbf{D}$ in (3.3) must be diagonal with the ith diagonal entry

$$
D_{ii} = \left(\nu_n - \frac{1}{\lambda_i}\right)^{+},
$$

where $(x)^{+} = \max(0, x)$ and $i = 1, 2, \ldots, n$. D_{ii} above is the power allocated to the ith stream and is found according to the water-filling power allocation where the water-filling level ν_n satisfies (Fig. 3.2)

$$
\sum_{i=1}^{n}\left(\nu_n - \frac{1}{\lambda_i}\right)^{+} = P.
\tag{3.4}
$$

Therefore, the capacity with perfect CSIT is the mutual information achieved by using $\mathbf{Q} = \mathbf{V}\mathbf{D}\mathbf{V}^{\dagger}$ and water-filling power allocation and thus following from (3.3), it is equal to

$$
\begin{aligned}
C_n^{\mathrm{P}} &= \log\det\left[\mathbf{I}_n + \boldsymbol{\Sigma}^{1/2}\mathbf{D}\boldsymbol{\Sigma}^{1/2}\right] \\
&= \sum_{i=1}^{n}\left(\log\left(\lambda_i\nu_n\right)\right)^{+},
\end{aligned}
\tag{3.5}
$$

69

where ν_n satisfies (3.4).

Alternatively, recalling the model $\mathbf{r} = \mathbf{Hs} + \mathbf{z}$ in Section 3.2 and applying the singular value decomposition $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}^{1/2}\mathbf{V}^\dagger$, we can write

$$\mathbf{r} = \mathbf{Hs} + \mathbf{z} = \mathbf{U}\boldsymbol{\Sigma}^{1/2}\mathbf{V}^\dagger\mathbf{s} + \mathbf{z} = \mathbf{U}\boldsymbol{\Sigma}^{1/2}\tilde{\mathbf{s}} + \mathbf{z},$$

where $\tilde{\mathbf{s}} = \mathbf{V}^\dagger\mathbf{s}$. Denoting $\tilde{\mathbf{r}} = \mathbf{U}^\dagger\mathbf{r}$ and $\tilde{\mathbf{z}} = \mathbf{U}^\dagger\mathbf{z}$, we can write

$$\tilde{\mathbf{r}} = \boldsymbol{\Sigma}^{1/2}\tilde{\mathbf{s}} + \tilde{\mathbf{z}},$$

where $\tilde{\mathbf{z}}$, the noise vector multiplied by a unitary matrix, has the same distribution as $\mathbf{z}$. Thus, the MIMO system can be decomposed into n parallel independent SISO channels with total transmit power limited to P, where the ith channel has the input $\tilde{s}_i$, the output $\tilde{r}_i$, noise $\tilde{z}_i$, and channel gain $\sqrt{\lambda_i}$. Thus, multiplexing over these parallel channels results in an n-fold increase in the data rate.

3.4   Water-filling Throughput with Imperfect CSIT

As previously discussed in Section 1.3, obtaining perfect CSIT is very challenging in practice. Thus, a more general and realistic scenario is to consider the case when the channel gain are not perfectly known at the Tx which happens due to nonzero channel estimation errors. In this section, we address the case of imperfect CSIT by modelling the channel estimation error to be Gaussian distributed. This model allows us to obtain statistical closed-form results and find how the performance is affected due to the lack of perfect CSIT.

With ML estimation [39], the estimated channel gain matrix $\hat{\mathbf{H}}$ is expressed as

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E}, \tag{3.6}$$

where $\mathbf{E}$ is the channel estimation error matrix that has i.i.d. and zero-mean CSCG-distributed entries with variance $2\sigma_e^2$. Furthermore, $\mathbf{E}$ and $\mathbf{H}$ are independent.

Throughout this chapter, we refer to the ratio $\gamma = \sigma^2/\sigma_e^2$ as the signal to estimation error ratio (SER). Intuitively, channel estimation error leads to degradation in the performance of water-filling in the MIMO system especially at low SNR. We aim to quantify this degradation in what follows.

From (3.6), the entries of $\hat{\mathbf{H}}$ are i.i.d. and zero-mean CSCG-distributed with variance $2\sigma^2 + 2\sigma_e^2$. Thus, the entries of $\hat{\mathbf{H}}$ and $\mathbf{H}$ are joint complex Gaussian distributed with correlation coefficient equal to $\sigma/\sqrt{\sigma^2 + \sigma_e^2}$. The actual channel gain matrix $\mathbf{H}$ can then be written as [86]

$$\mathbf{H} = \eta\hat{\mathbf{H}} + \mathbf{X},\qquad(3.7)$$

where $\eta = \gamma/(1 + \gamma)$, $\hat{\mathbf{H}}$ and $\mathbf{X}$ are independent, and the entries of $\mathbf{X}$ are i.i.d. and zero-mean CSCG-distributed with variance $2\sigma^2/(1 + \gamma)$.

In this chapter, as stated earlier, we do not aim to find the capacity of the MIMO system with imperfect CSIT as the Tx does not know the quality of channel estimation or the SER. Instead, we assume that the Tx performs water-filling power allocation merely based on erroneous channel estimates. As a consequence, the throughput obtained in such case is upper bounded by the capacity of the MIMO system with perfect CSIT as obtained in Section 3.3.2

Note that the Tx only knows the channel estimate matrix $\hat{\mathbf{H}}$ which can be decomposed as $\hat{\mathbf{H}} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{V}}^\dagger$ where $\hat{\boldsymbol{\Sigma}} = \mathrm{diag}\left(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n\right)$ with $\{\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n\}$ as the eigenvalues of $\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger$. Thus, water-filling power allocation yields the transmit covariance matrix $\hat{\mathbf{Q}} = \hat{\mathbf{V}}\hat{\mathbf{D}}\hat{\mathbf{V}}^\dagger$ where $\hat{\mathbf{D}}$ is an $n \times n$ diagonal matrix with the ith diagonal entry

$$\hat{D}_{ii} = \left(\hat{\nu}_n - \frac{1}{\hat{\lambda}_i}\right)^+,\qquad(3.8)$$

as the power allocated to the ith stream and the water-filling level $\hat{\nu}_n$ satisfies

$$\sum_{i=1}^{n}\left(\hat{\nu}_n - \frac{1}{\hat{\lambda}_i}\right)^+ = P.\qquad(3.9)$$

Therefore, following from (3.1), the water-filling throughput can be derived as

$$R_n = \log\det\left[\mathbf{I}_n + \mathbf{H}\hat{\mathbf{Q}}\mathbf{H}^\dagger\right].\qquad(3.10)$$

Knowing $\mathbf{H} = \eta\hat{\mathbf{H}} + \mathbf{X}$ from (3.7), Eq. (3.10) can be written as

$$R_n = \log\det\left[\mathbf{I}_n + \left(\eta\hat{\mathbf{H}} + \mathbf{X}\right)\hat{\mathbf{Q}}\left(\eta\hat{\mathbf{H}} + \mathbf{X}\right)^{\dagger}\right]$$

$$= \log\det\left[\mathbf{I}_n + \eta^2\hat{\mathbf{H}}\hat{\mathbf{Q}}\hat{\mathbf{H}}^{\dagger} + \eta\hat{\mathbf{H}}\hat{\mathbf{Q}}\mathbf{X}^{\dagger} + \eta\mathbf{X}\hat{\mathbf{Q}}\hat{\mathbf{H}}^{\dagger} + \mathbf{X}\hat{\mathbf{Q}}\mathbf{X}^{\dagger}\right]. \qquad (3.11)$$

Replacing $\hat{\mathbf{Q}} = \hat{\mathbf{V}}\hat{\mathbf{D}}\hat{\mathbf{V}}^{\dagger}$ and $\hat{\mathbf{H}} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{V}}^{\dagger}$ in (3.11), we find

$$R_n = \log\det\left[\mathbf{I}_n + \eta^2\hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{D}}\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{U}}^{\dagger} + \eta\hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{D}}\hat{\mathbf{V}}^{\dagger}\mathbf{X}^{\dagger}\right.$$

$$\left. + \eta\mathbf{X}\hat{\mathbf{V}}\hat{\mathbf{D}}\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{U}}^{\dagger} + \mathbf{X}\hat{\mathbf{V}}\hat{\mathbf{D}}\hat{\mathbf{V}}^{\dagger}\mathbf{X}^{\dagger}\right]$$

$$= \log\det\left[\mathbf{I}_n + \eta^2\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{D}}\hat{\boldsymbol{\Sigma}}^{1/2} + \eta\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{D}}\hat{\mathbf{V}}^{\dagger}\mathbf{X}^{\dagger}\hat{\mathbf{U}}\right.$$

$$\left. + \eta\hat{\mathbf{U}}^{\dagger}\mathbf{X}\hat{\mathbf{V}}\hat{\mathbf{D}}\hat{\boldsymbol{\Sigma}}^{1/2} + \hat{\mathbf{U}}^{\dagger}\mathbf{X}\hat{\mathbf{V}}\hat{\mathbf{D}}\hat{\mathbf{V}}^{\dagger}\mathbf{X}^{\dagger}\hat{\mathbf{U}}\right]$$

$$= \log\det\left[\mathbf{I}_n + \eta^2\hat{\boldsymbol{\Sigma}}\hat{\mathbf{D}} + \eta\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{D}}\mathbf{K}^{\dagger} + \eta\mathbf{K}\hat{\mathbf{D}}\hat{\boldsymbol{\Sigma}}^{1/2} + \mathbf{K}\hat{\mathbf{D}}\mathbf{K}^{\dagger}\right], \qquad (3.12)$$

where $\mathbf{K} = \hat{\mathbf{U}}^{\dagger}\mathbf{X}\hat{\mathbf{V}}$.

Defining the matrices

$$\mathbf{Y} = \mathbf{I}_n + \eta^2\hat{\boldsymbol{\Sigma}}\hat{\mathbf{D}}, \qquad (3.13)$$

$$\mathbf{L} = \eta\mathbf{K}\hat{\mathbf{D}}\hat{\boldsymbol{\Sigma}}^{1/2} + \eta\hat{\boldsymbol{\Sigma}}^{1/2}\hat{\mathbf{D}}\mathbf{K}^{\dagger}, \qquad (3.14)$$

$$\mathbf{T} = \mathbf{K}\hat{\mathbf{D}}\mathbf{K}^{\dagger}, \qquad (3.15)$$

Eq. (3.12) can be written as

$$R_n = \log\det\left[\mathbf{Y} + \mathbf{L} + \mathbf{T}\right]$$

$$= \log\det\mathbf{Y} + \log\det\left[\mathbf{I}_n + \mathbf{Y}^{-1/2}\left(\mathbf{L} + \mathbf{T}\right)\mathbf{Y}^{-1/2}\right]. \qquad (3.16)$$

3.5   Asymptotic Growth Rates

The throughputs I_n^{N}, C_n^{P}, and R_n presented in sections 3.3 and 3.4, are random variables because the channel gain matrix $\mathbf{H}$ and its estimate $\hat{\mathbf{H}}$ are random. However, in [19], using Theorem IV.I, it is proved that the normalized throughputs I_n^{N}/n and C_n^{P}/n converge to non-random values almost-surely as $n \to \infty$, i.e., $I_n^{\mathrm{N}}/n \to I^{\mathrm{N}}$ and $C_n^{\mathrm{P}}/n \to C^{\mathrm{P}}$.

Furthermore, for the special case that the entries of $\mathbf{H}$ have unit variance $(2\sigma^2 = 1)$, the proportionality constants (the asymptotic growth rates) I^{N} and C^{P} were derived.

In this section, we first state the result of Theorem IV.I in [19] for arbitrary σ and σ_e. Then, we find I^{N} and C^{P} for an arbitrary σ. Subsequently, we show the almost-sure convergence of the normalized water-filling throughput with imperfect CSIT, R_n/n, as $n \to \infty$ at low SNR and derive R, the asymptotic growth rate for the throughput.

Let $\mathbf{G}$ be an $n \times n$ matrix that has i.i.d. and zero-mean CSCG-distributed entries with unit variance. According to Theorem IV.I in [19], the normalized (scaled by $1/n$) empirical eigenvalue distribution of $\mathbf{GG}^\dagger$, converges to a limit which has the density

$$g(\lambda) = \begin{cases} \frac{1}{\pi}\sqrt{\frac{1}{\lambda} - \frac{1}{4}} & 0 \le \lambda \le 4 \\ 0 & \text{else.} \end{cases} \tag{3.17}$$

and the largest eigenvalue of $\mathbf{GG}^\dagger$ converges to

$$\lim_{n\to\infty} \frac{\lambda_{\max}(\mathbf{GG}^\dagger)}{n} = 4,$$

almost surely. As stated earlier, the entries of $\mathbf{H}$ are i.i.d. and zero-mean CSCG-distributed with variance $2\sigma^2$. Therefore, $\mathbf{H}$ is equal in distribution with $\sqrt{2}\sigma\mathbf{G}$ [77] and consequently the eigenvalues of $\mathbf{HH}^\dagger$ are scaled versions (by $2\sigma^2$) of the eigenvalues of $\mathbf{GG}^\dagger$. Thus, the normalized empirical eigenvalue distribution of $\mathbf{HH}^\dagger$ converges to a limit as well and the limit has a density which is a scaled version of $g(\lambda)$ in (3.17) as follows

$$h(\lambda) = \begin{cases} \frac{1}{\sqrt{2}\sigma\pi}\sqrt{\frac{1}{\lambda} - \frac{1}{8\sigma^2}} & 0 \le \lambda \le 8\sigma^2 \\ 0 & \text{else.} \end{cases} \tag{3.18}$$

In addition, almost surely, we have

$$\lim_{n\to\infty} \frac{\lambda_{\max}(\mathbf{HH}^\dagger)}{n} = 8\sigma^2.$$

Following the same approach as above, since the entries of $\hat{\mathbf{H}}$ are i.i.d. and zero-mean CSCG-distributed with variance $2\sigma^2 + 2\sigma_e^2$, $\hat{\mathbf{H}}$ is equal in distribution with $\sqrt{2(\sigma^2 + \sigma_e^2)}\mathbf{G}$ and the eigenvalues of $\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger$ are scaled versions (by $2\sigma^2 + 2\sigma_e^2$) of the eigenvalues of $\mathbf{GG}^\dagger$.

Thus, the normalized empirical eigenvalue distribution of $\hat{\mathbf{H}}\hat{\mathbf{H}}^{\dagger}$ converges to a limit with density

$$
\hat{h}(\lambda) = \begin{cases} \frac{1}{\sqrt{2(\sigma^2+\sigma_e^2)}\pi}\sqrt{\frac{1}{\lambda} - \frac{1}{8(\sigma^2+\sigma_e^2)}} & 0 \leq \lambda \leq 8(\sigma^2 + \sigma_e^2) \\ 0 & \text{else}, \end{cases}
\tag{3.19}
$$

and almost surely

$$
\lim_{n\to\infty} \frac{\lambda_{\max}(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\dagger})}{n} = 8(\sigma^2 + \sigma_e^2).
$$

3.5.1 No CSIT

Here, we analyze the behaviour of the normalized throughput I_n^{N}/n as n grows large. Since the empirical distribution of λ_i/n in (3.2) converges to a limit with density $h(\lambda)$ in (3.18), as $n \to \infty$, following from (3.2) and using the Law of Large Numbers (LLN) [77], we have [19]

$$
\frac{I_n^{\mathrm{N}}}{n} = \frac{1}{n}\sum_{i=1}^{n} \log\left(1 + P\frac{\lambda_i}{n}\right) \to I^{\mathrm{N}},
$$

where

$$
I^{\mathrm{N}} = \mathsf{E}\left[\log\left(1 + P\lambda\right)\right] = \int_0^{8\sigma^2} \log(1 + P\lambda)h(\lambda)d\lambda.
\tag{3.20}
$$

As $P \to 0$, following the same approach as in [19], a first-order approximation gives

$$
I^{\mathrm{N}} \approx P\log e \int_0^{8\sigma^2} \lambda h(\lambda)d\lambda = 2P\sigma^2 \log e,
\tag{3.21}
$$

which is in bits/sec/Hz as opposed to the result in equation (12) of [19] which is obtained in nats/sec/Hz. Furthermore, a factor of $2\sigma^2$ appears in (3.21) which is due to the fact that the channel gain matrix $\mathbf{H}$ has entries with variance $2\sigma^2$.

3.5.2 Perfect CSIT

As stated earlier, in the perfect CSIT case, ν_n is the water-filling level that asymptotically scales as $\Theta(1/n)$ [19]. We define $\nu = n \times \nu_n$ as the scaled (by n) version of ν_n, and consider

74

the limit $\lim_{n \to \infty} \nu = \nu^\star$ as the asymptotic scaled water-filling level in the perfect CSIT case.

To find the asymptotic throughput, as in [19], we relabel ν_n as $\nu \times \frac{1}{n}$ and then incorporate the $1/n$ factor with λ_i. Thus, we can write the water-filling constraint in (3.9) as

$$\sum_{i=1}^{n} \left(\nu_n - \frac{1}{\lambda_i} \right)^+ = \sum_{i=1}^{n} \left(\frac{\nu}{n} - \frac{1}{\lambda_i} \right)^+ = \frac{1}{n} \sum_{i=1}^{n} \left(\nu - \left(\frac{\lambda_i}{n} \right)^{-1} \right)^+ = P.$$

As stated earlier, the empirical distribution of $\{\lambda_i/n\}$ converges to the density $h(\lambda)$ in (3.18). Thus, using the LLN, we obtain

$$\frac{1}{n} \sum_{i=1}^{n} \left(\nu - \left(\frac{\lambda_i}{n} \right)^{-1} \right)^+ \to \int_0^{8\sigma^2} \left(\nu^\star - \frac{1}{\lambda} \right)^+ h(\lambda) d\lambda = P, \tag{3.22}$$

as $n \to \infty$. From the constraint above, we can find $\nu^\star$.

Similarly, (3.5) can be expanded as

$$\frac{C_n^P}{n} = \frac{1}{n} \sum_{i=1}^{n} \left(\log \left(\nu_n \lambda_i \right) \right)^+ = \frac{1}{n} \sum_{i=1}^{n} \left(\log \left(\left(\frac{\nu}{n} \right) \lambda_i \right) \right)^+ = \frac{1}{n} \sum_{i=1}^{n} \left(\log \left(\nu \left(\frac{\lambda_i}{n} \right) \right) \right)^+ \to C^P,$$

as $n \to \infty$, where

$$C^P = \int_0^{8\sigma^2} \left(\log \left(\nu^\star \lambda \right) \right)^+ h(\lambda) d\lambda. \tag{3.23}$$

In (3.23) and (3.22), following the same approach as in [19], we can find $dC^P/dP = \log e / \nu^\star$ and $\nu^\star \to 1/8\sigma^2$ as $P \to 0$ respectively, and thus obtain the first-order approximation

$$C^P \approx 8\sigma^2 P \log e. \tag{3.24}$$

Hence, as previously shown in [19], knowing (3.21) and (3.24), we obtain

$$\frac{C^P}{I^N} \approx 4,$$

as $P \to 0$, i.e., at low SNR, availability of perfect CSIT provides significant performance improvement for the MIMO system asymptotically compared to the no CSIT scenario.

3.5.3 Imperfect CSIT

The empirical distribution of $\hat{\lambda}_i/n$ converges to a limit with density as in (3.19). Thus, by relabeling $\hat{\nu}_n$ as $\hat{\nu}/n$, and using the LLN along with (3.19), we can rewrite the water-filling constraint in (3.9) as

$$\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\nu}-\left(\frac{\hat{\lambda}_i}{n}\right)^{-1}\right)^{+}\to\int_{0}^{8(\sigma^2+\sigma_e^2)}\left(\hat{\nu}^{\star}-\frac{1}{\lambda}\right)^{+}\hat{h}(\lambda)d\lambda=P,\qquad(3.25)$$

as $n\to\infty$, where $\hat{\nu}^{\star}=\lim_{n\to\infty}\hat{\nu}$ is the asymptotic scaled water-filling level in the imperfect CSIT case.

Next, we find the asymptotic behaviour of the normalized water-filling throughput R_n/n as $n\to\infty$. Since $\hat{\boldsymbol{\Sigma}}$ and $\hat{\mathbf{D}}$ are diagonal with the ith diagonal entry respectively as $\hat{\lambda}_i$ and (3.8), the matrix $\mathbf{Y}$ in (3.13) is diagonal and the first term in (3.16) is thus

$$\log\det\mathbf{Y}=\log\det\left[\mathbf{I}_n+\eta^2\hat{\boldsymbol{\Sigma}}\hat{\mathbf{D}}\right]=\sum_{i=1}^{n}\log(1+\eta^2(\hat{\lambda}_i\hat{\nu}_n-1)^{+}).\qquad(3.26)$$

Relabeling $\hat{\nu}_n$ as $\hat{\nu}/n$ in (3.26), we get the normalized result

$$\frac{1}{n}\log\det\mathbf{Y}=\frac{1}{n}\sum_{i=1}^{n}\log\left(1+\eta^2\left(\hat{\nu}\frac{\hat{\lambda}_i}{n}-1\right)^{+}\right),\qquad(3.27)$$

which converges to

$$\frac{1}{n}\log\det\mathbf{Y}\to\int_{0}^{8(\sigma^2+\sigma_e^2)}\log(1+\eta^2(\hat{\nu}^{\star}\lambda-1)^{+})\hat{h}(\lambda)d\lambda,\qquad(3.28)$$

as $n\to\infty$, where $\hat{\nu}^{\star}$ is found from (3.25).

Denoting $L_{i,j}$ as the entry in the ith row and the jth column of the matrix $\mathbf{L}$, for $i,j=1,\ldots,n$, we find from (3.14)

$$L_{i,j}=\eta\sqrt{\hat{\lambda}_j}\hat{D}_{jj}K_{i,j}+\eta\sqrt{\hat{\lambda}_i}\hat{D}_{ii}K_{j,i}^{*}.\qquad(3.29)$$

Using (3.8), we can write

$$L_{i,j}=\eta\sqrt{\hat{\lambda}_j}\left(\hat{\nu}_n-\frac{1}{\hat{\lambda}_j}\right)^{+}K_{i,j}+\eta\sqrt{\hat{\lambda}_i}\left(\hat{\nu}_n-\frac{1}{\hat{\lambda}_i}\right)^{+}K_{j,i}^{*},$$

and relabeling $\hat{\nu}_n$ as $\hat{\nu}/n$, we find

$$L_{i,j} = \eta \sqrt{\frac{\hat{\lambda}_j}{n}} \left(\hat{\nu} - \left(\frac{\hat{\lambda}_j}{n} \right)^{-1} \right)^{+} \frac{K_{i,j}}{\sqrt{n}} + \eta \sqrt{\frac{\hat{\lambda}_i}{n}} \left(\hat{\nu} - \left(\frac{\hat{\lambda}_i}{n} \right)^{-1} \right)^{+} \frac{K_{j,i}^{*}}{\sqrt{n}}. \tag{3.30}$$

To analyze the asymptotic behaviour of $L_{i,j}$ above, we should point out the following facts:

1. The empirical distribution of $\hat{\lambda}_j/n$ for any $j = 1, 2, \ldots, n$, converges to a limit with density as in (3.19) as $n \to \infty$. Thus, $\hat{\lambda}_j/n$ in (3.30) and consequently $(\hat{\nu} - (\hat{\lambda}_j/n)^{-1})^{+}$ scale as $\Theta(1)$ for all j.

2. The entries of $\mathbf{K} = \mathbf{U}^{\dagger}\mathbf{X}\mathbf{V}$ are i.i.d. and zero-mean CSCG-distributed with variance $2\sigma^2/(1+\gamma)$ that does not depend on n. This is because $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices (obtained from the singular value decomposition of $\hat{\mathbf{H}}$) and the entries of $\mathbf{X}$ are i.i.d. and zero-mean CSCG-distributed with variance $2\sigma^2/(1 + \gamma)$.

Therefore, because of the above facts and having the $\sqrt{n}$ term in the denominator of $L_{i,j}$ in (3.30), for any $i, j = 1, 2, \ldots, n$, $L_{i,j} \to 0$ as $n \to \infty$.

For the matrix $\mathbf{T}$ in (3.15), we have

$$T_{i,j} = \sum_{l=1}^{n} \left(\hat{\nu}_n - \frac{1}{\hat{\lambda}_l} \right)^{+} K_{i,l}K_{j,l}^{*} = \frac{1}{n} \sum_{l=1}^{n} \left(\hat{\nu} - \left(\frac{\hat{\lambda}_l}{n} \right)^{-1} \right)^{+} K_{i,l}K_{j,l}^{*}, \tag{3.31}$$

which is obtained by relabeling $\hat{\nu}_n$ as $\hat{\nu}/n$. Because of the factor of $1/n$ in the expression of $T_{i,j}$ above, using the LLN, as $n \to \infty$, $T_{i,j}$ converges to its expected value as

$$T_{i,j} \to \mathsf{E}\left[\left(\hat{\nu}^{\star} - \frac{1}{\hat{\lambda}} \right)^{+} K_{i,l}K_{j,l}^{*} \right]. \tag{3.32}$$

Since $\hat{\mathbf{H}} \to (\mathbf{U}, \mathbf{V}) \to \mathbf{K}$ is a Markov chain, thus $\mathbf{K}$ and $\hat{\mathbf{H}}$ are independent given $\mathbf{U}$ and $\mathbf{V}$. Therefore, we can write

$$\mathsf{E}\left[\left(\hat{\nu}^{\star} - \frac{1}{\hat{\lambda}} \right)^{+} K_{i,l}K_{j,l}^{*} \right] = \mathsf{E}\left[\mathsf{E}\left[\left(\hat{\nu}^{\star} - \frac{1}{\hat{\lambda}} \right)^{+} K_{i,l}K_{j,l}^{*} | \mathbf{U}, \mathbf{V} \right] \right]$$

$$= \mathsf{E}\left[\mathsf{E}\left[\left(\hat{\nu}^{\star} - \frac{1}{\hat{\lambda}} \right)^{+} | \mathbf{U}, \mathbf{V} \right] \mathsf{E}\left[K_{i,l}K_{j,l}^{*} | \mathbf{U}, \mathbf{V} \right] \right]. \tag{3.33}$$

The water-filling level $\hat{\nu}^\star$ was chosen in (3.9) such that (3.9) holds for each realization of $\hat{\mathbf{H}}$. Thus,

$$\mathsf{E}\left[\left(\hat{\nu}^\star - \frac{1}{\hat{\lambda}}\right)^+ |\mathbf{U}, \mathbf{V}\right] = P. \tag{3.34}$$

Hence, from (3.32), (3.33), and (3.34) we obtain

$$T_{i,i} \to P\mathsf{E}[|K_{i,l}|^2] = \frac{2P\sigma^2}{1+\gamma}, \tag{3.35}$$

for $i = 1, 2, \ldots, n$, and

$$T_{i,j} \to P\mathsf{E}[K_{i,l}]\mathsf{E}[K_{j,l}^*] = 0,$$

for $i, j = 1, 2, \ldots, n$, and $i \neq j$.

Although, elementwise, off-diagonal entries $Z_{i,j}$ of $\mathbf{Z} = \mathbf{I}_n + \mathbf{Y}^{-1/2}\left(\mathbf{L} + \mathbf{T}\right)\mathbf{Y}^{-1/2}$ for $i, j = 1, \ldots, n$, $i \neq j$ obtained before are such that $Z_{i,j} \to 0$ as $n \to \infty$, this does not necessarily imply that $\det \mathbf{Z} \to \prod_{i=1}^n Z_{i,i}$ as $n \to \infty$. This is because the effect of off-diagonal entries of $\mathbf{Z}$ may not necessarily be neglected since the number of these entries are increasing with n. However, for a fixed n, as $P \to 0$, the off-diagonal entries of $\mathbf{Z}$ vanish and thus $\det \mathbf{Z} \to \prod_{i=1}^n Z_{i,i}$. For example, it is easy to verify by cofactor expansion for an $n \times n$ matrix $\mathbf{A}$ with constant entries and $\epsilon \to 0$ that $\det[\mathbf{I}_n + \epsilon\mathbf{A}] = \prod_{i=1}^n(1 + \epsilon A_{i,i}) + O(\epsilon^2)$. Therefore, knowing (3.35), in the low SNR regime, we obtain

$$\det \mathbf{Z} \approx \prod_{i=1}^n Z_{i,i} = \det\left[\mathbf{I}_n + \frac{2P\sigma^2 \mathbf{Y}^{-1}}{1+\gamma}\right],$$

as $n \to \infty$. Thus, using (3.13) and using the LLN, we can write

$$\frac{1}{n}\log\det\mathbf{Z} \to \int_0^{8(\sigma^2+\sigma_e^2)}\log\left(1 + \frac{2P\sigma^2/(1+\gamma)}{1+\eta^2(\lambda\hat{\nu}^\star - 1)^+}\right)\hat{h}(\lambda)d\lambda. \tag{3.36}$$

Hence, from (3.28) and (3.36), we find $R_n/n \to R$ as $n \to \infty$, i.e., at low SNR, R_n scales linearly with n as $n \to \infty$ with a proportionality constant (the asymptotic growth rate) that is not random and is given by R as

$$R \approx \int_0^{8\sigma^2(1+1/\gamma)}\log\left(1 + \left(\gamma^2/(1+\gamma)^2\right)(\lambda\hat{\nu}^\star - 1)^+\right)\hat{h}(\lambda)d\lambda$$

$$+ \int_0^{8\sigma^2(1+1/\gamma)}\log\left(1 + \frac{2P\sigma^2/(\gamma+1)}{1 + (\gamma^2/(1+\gamma)^2)(\lambda\hat{\nu}^\star - 1)^+}\right)\hat{h}(\lambda)d\lambda, \tag{3.37}$$

where $\hat{\nu}^\star$ satisfies (3.25).

78

3.6 Performance Evaluation

As shown earlier, the throughput achieved by water-filling with imperfect CSIT scales linearly with n as $n \to \infty$, at low SNR, with a proportionality constant that is not random, but is given by (3.37). Eq. (3.37) gives an insight on how the water-filling throughput depends on the parameters of the system such as γ, P, and σ.

Using (3.37), one can evaluate the effect of channel estimation error on the performance of water-filling power allocation at low SNR. For a fixed σ, when $\gamma = \infty$ dB (high SER), the second term in (3.37) is zero and

$$R = \int_0^{8\sigma^2} \left(\log\left(\lambda\hat{\nu}^\star\right)\right)^+ h(\lambda)d\lambda = C^{\mathrm{P}},$$

which is the asymptotic growth rate for water-filling with perfect CSIT as found in (3.23). Furthermore, for a fixed σ and $\gamma = -\infty$ dB (low SER), the first term in (3.37) is zero and as $P \to 0$ we have

$$R \approx \int_0^\infty \log\left(1 + 2P\sigma^2\right)\hat{h}(\lambda)d\lambda \approx 2P\sigma^2 \log e \int_0^\infty \hat{h}(\lambda)d\lambda$$
$$= 2P\sigma^2 \log e,$$

which is equal to I^{N} in (3.21). This result implies that, at low SER, water-filling based on erroneous channel estimates achieves the same throughput as equal power allocation asymptotically in the low SNR regime. This is intuitively because at low SER, the density $\hat{h}(\lambda)$ in (3.19) is defined over a wider range and therefore, all the eigenvalues are very likely to be large, which results in power being allocated equally among all eigen-directions.

Note that since we assume that the channel is quasi-static, where the channel gain matrix $\mathbf{H}$ is fixed for the entire duration of the transmission, the meaningful notion operationally is the instantaneous throughput. In addition, for large n, we have a channel hardening effect at low SNR, i.e., the asymptotic normalized throughput in (3.37) does not depend on the realization of $\mathbf{H}$. Thus, for large n, averaging the instantaneous normalized throughput over multiple realizations would give the same result.
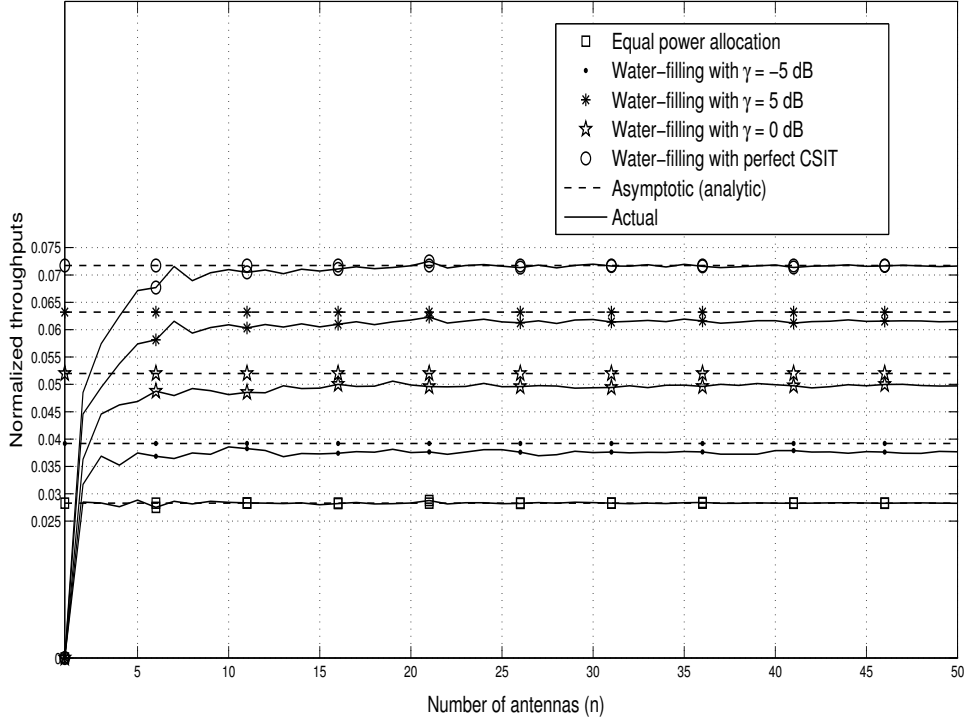
Figure 3.3: Actual and asymptotic (analytic) normalized throughputs I_n^{N}/n, C_n^{P}/n, and R_n/n versus n for different values of γ when $P = -20$ dB and $\sigma = 1$.

3.7   Numerical Results

In this section, we provide numerical results when $\sigma = 1$ and $P = -20$ dB.

Fig. 3.3 shows the convergence of the actual normalized throughputs (averaged over multiple realizations) I_n^{N}/n, C_n^{P}/n, and R_n/n to the asymptotic (analytic) results I^{N}, C^{P}, and R presented in (3.20), (3.23), and (3.37), respectively, as n increases. For $n = 50$, Fig. 3.4 illustrates the actual (averaged over multiple realizations) and the asymptotic normalized throughputs versus the SER.

Based on the figures, water-filling with imperfect CSIT monotonically interpolates between water-filling with perfect CSIT and equal power allocation. Furthermore, one can compare the result for water-filling with imperfect CSIT with the result for water-filling
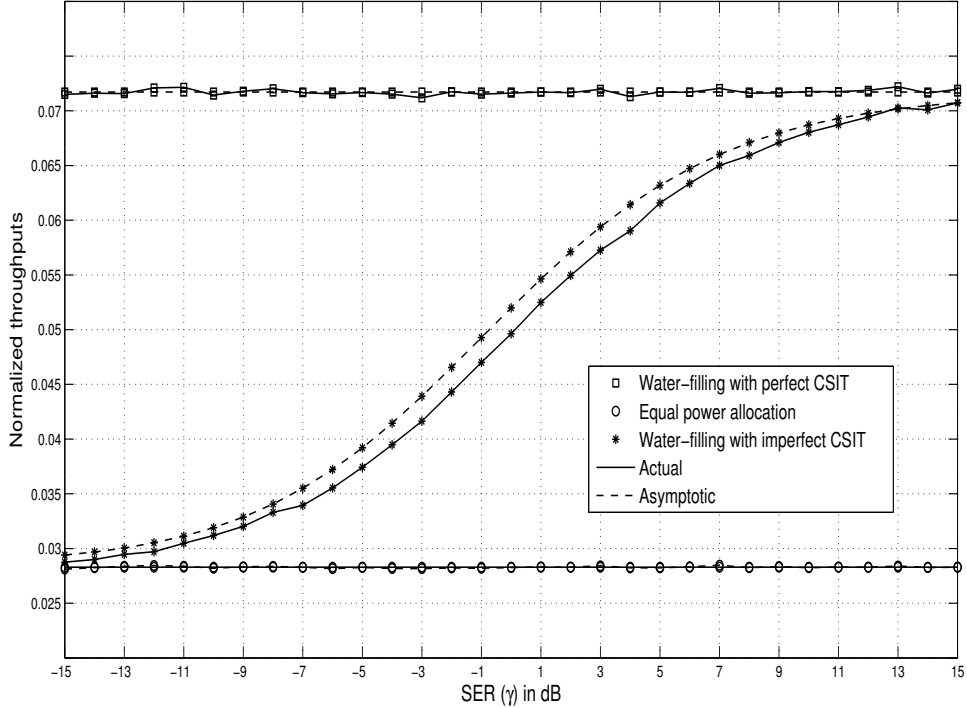
Figure 3.4: Actual and asymptotic (analytic) normalized throughputs I_n^{N}/n, C_n^{P}/n, and R_n/n versus γ, when $P = -20$ dB, $\sigma = 1$, and $n = 50$.

with perfect CSIT as a function of the SER. Particularly, with imperfect CSIT and for SER values such as 5 dB, 0 dB, and -5 dB, R is 86%, 70%, and 52% of C^{P}, respectively. In other words, at low SNR, water-filling based on erroneous channel estimates can still achieve significant throughputs asymptotically for moderate values of the SER. Further-more, at very low SER, water-filling based on erroneous channel estimates achieves the same throughput as equal power allocation.

For imperfect CSIT, the gap between the actual and the asymptotic (analytic) results in the figures is very small. When P is above -15 dB, the approximation in (3.37) starts to deviate notably. However, numerically, we find that at higher SNR values, the same conclusion holds that moderate SER of 5 dB provides performance close to C^{P}.

3.8 Summary and Conclusions

Water-filling power allocation is known to achieve the capacity of an $n \times n$ MIMO system with perfect CSIT. More specifically, knowing the eigenvalues of the channel, the transmitter can optimally distribute power among antennas according to water-filling power allocation and thus achieve the maximum throughput (the capacity of the MIMO system). Furthermore, in a rich scattering environment, such capacity has been shown to scale linearly with n as n grows large (intuitively due to spatial multiplexing). In other words, the normalized (by $1/n$) throughput converges almost surely to a non-random value denoted by, C^{P}, as n increases. This result has been derived by using the limiting distribution of the eigenvalues of the channel gain matrix as n grows large.

On the other hand, in the case of no CSIT, where the transmitter has no knowledge of the channel gain matrix, equal power allocation can be performed, i.e., the transmitter can distribute the transmit power equally among antennas. In other words, the same amount of power is allocated to all eigen-directions regardless of how strong their eigenvalues are. Doing such a power allocation, it has been shown that the normalized (by $1/n$) throughput converges to a non-random value denoted by I^{N}, almost surely as n increases.

At high SNR, C^{P} and I^{N} are asymptotically equal which implies that there is no loss in the throughput if power is distributed equally among antennas (which requires no CSIT) instead of performing water-filling power allocation. As the SNR decreases, the gap between C^{P} and I^{N} increases and water-filling power allocation results in better throughputs compared to equal power allocation. Thus, at low SNR, it is crucially important how the transmitter allocates power among antennas.

Furthermore, in reality, obtaining CSIT within a reasonable accuracy requires significant effort, especially at low SNR and in addition, the transmitter may have no a priori reliable knowledge of the accuracy of channel estimates either. Motivated by these, in this chapter, we have theoretically evaluated the performance loss of water-filling power allocation due to channel estimation error at the transmitter as a function of the signal to estimation error ratio (SER). In this regard, we have derived the throughput achieved in MIMO systems if the transmitter performs water-filling based on erroneous channel

estimates only and have shown that even though the channel gain matrix is random, at low SNR, the normalized water-filling throughput based on imperfect CSIT converges to a non-random value denoted by R, almost surely as n increases. Subsequently, using C^{P} as a baseline for comparison, we have compared R with C^{P} as a function of the SER and have found that at low SNR, for moderate values of the SER, water-filling based on erroneous channel estimates can still achieve significant throughputs asymptotically. In particular, for SER values such as 5 dB, 0 dB, and -5 dB, R is found to be 86%, 70%, and 52% of C^{P}, respectively.

In addition, we have shown that at low SNR and low SER, water-filling based on erroneous channel estimates asymptotically achieves the same throughput as equal power allocation. This result implies that at low SNR, water-filling power allocation does not perform worse than equal power allocation even when the channel estimates are very erroneous. The asymptotic analysis in this chapter, which is valid when the number of antennas is very large, can be applied to evaluate the throughput results in the massive MIMO scenario that has received considerable attention recently in the academia and industry.

Chapter 4

Conclusions and Future Work

4.1 Conclusions

Multi-antenna technology offers powerful capabilities to improve data rates of wireless systems via spatial multiplexing as well as to increase the reliability of wireless communications via spatial diversity. Furthermore, by exploiting additional spatial dimensions, transmit beamforming techniques can be used to manage co-channel interference, which is a major limiting factor in the performance of wireless systems.

In Chapter 2, transmit diversity and beamforming were studied in a MISO setting with an n-antenna transmitter, an intended single-antenna receiver, and m unintended single-antenna receivers. Two scenarios were considered, namely, null-steering beamforming and ϵ-threshold beamforming in which the allowable interference threshold at the unintended receivers is zero and $\epsilon > 0$, respectively. In each beamforming scenario, two separate cases of perfect CSIT and imperfect CSIT were investigated, and in each case, an optimization problem was obtained that maximizes the received power at the intended receiver while satisfying the interference constraints at the unintended receiver and the transmit power constraint.

With perfect CSIT, null-steering beamforming was shown to successfully nullify interference at m unintended receivers, where $m < n$, and achieve a nonzero received power

at the intended receiver with a mean value that grows linearly with $n - m$ and is directly proportional to the power of the line-of-sight component between the transmitter and the intended receiver. With imperfect CSIT, null-steering beamforming based on erroneous channel estimates was shown to result in a nonzero interference at the unintended receivers with a mean value that is interestingly independent of n. Also, it was shown that a moderate line-of-sight component can significantly reduce the effect of estimation error on the performance of the intended link.

Intuitively, the allowance of a small nonzero interference at the unintended receivers as in ϵ-threshold beamforming should improve the received power at the intended receiver. According to the analysis in Chapter 2, this enhancement is marginal and not worthwhile, notably in the case of imperfect CSIT. More specifically, it was shown that when the interference threshold of the unintended receiver is a small value relative to the transmit power, ϵ-threshold beamforming does not lead to a significant increases in the average received power at the intended receiver compared to the null-steering result. Therefore, there is no significant loss in the performance of the intended link if the transmitter performs null-steering beamforming instead. In fact, according to the analysis, the transmitter can employ additional antennas to improve the performance of the intended link without generating significant extra interference on the unintended receivers.

In general, co-channel interference may be imposed from one wireless system to another, such as in cognitive radio systems, or from a wireless user to another user in the same wireless system, such as in heterogeneous networks. Such co-existence in the same frequency band is sometimes intentionally allowed with the objective to increase spectral efficiency. Therefore, the use of interference management techniques such as beamforming is essential to handle the interference among such wireless users and systems.

In cognitive radio systems, the secondary transmitter can employ multiple antennas and perform null-steering beamforming to nullify its co-channel interference at some primary receivers. In this case, if the variance of the channel estimation error at the secondary transmitter is small enough, the imposed average interference at the primary receivers can be kept less than the interference threshold regulated by the primary system. Therefore, the

concurrent operation of the secondary system in the same frequency band as the primary system does not degrade the performance of the primary system severely. Furthermore, since the average interference power at the primary receivers is independent of the number of secondary antennas, the secondary system can use a larger number of antennas to boost its performance without significantly impacting the primary system's performance.

Beamforming can also be used in heterogenous networks such as a cellular network consisting of macrocells and some smaller cells to manage co-channel interference among them. If each base station (BS) is equipped with multiple antennas, it can perform null-steering beamforming to transmit to its associated user while null-steering towards the users with which it is interfering the most. Consequently, the amount of interference imposed on each mobile user is decreased. Therefore, the signal to interference and noise ratio (SINR) corresponding to that mobile user increases accordingly, and, intuitively, the coverage probability and the achievable rate are expected to increase. Moreover, in such cases, the performance improvement offered by increasing the number of antennas at each BS can be examined. As the number of antennas increases, a BS has more degrees of freedom for null-steering. With n antennas, a macro BS can null-steer towards $(n-1)$ mobile users. Subsequently, a tradeoff between the number of antennas and the performance of such a cellular network can be obtained analytically.

In Chapter 3, we studied spatial multiplexing in an $n \times n$ MIMO wireless model. It is known that water-filling power allocation achieves the capacity of such a MIMO system with perfect CSIT. In fact, in a rich scattering environment, such capacity was shown to scale linearly with n as n grows large. In other words, the normalized (by $1/n$) throughput converges almost surely to a non-random value, denoted by C^{P}, as n increases. This result was derived by using the limiting distribution of the eigenvalues of the channel gain matrix as n grows large. On the other hand, in the case of no CSIT, where the transmitter has no knowledge of the channel gain matrix, equal power allocation can be performed, i.e., the transmitter can distribute the transmit power equally among antennas. Doing such a power allocation, it was shown that the normalized (by $1/n$) throughput converges to a non-random value denoted by I^{N}, almost surely as n increases.

At high SNR, C^{P} and I^{N} are asymptotically equal which implies that there is no loss in throughput if power is distributed equally among antennas (which requires no CSIT) instead of performing water-filling power allocation. As the SNR decreases, the gap between C^{P} and I^{N} increases and water-filling power allocation results in better throughputs compared to equal power allocation. Thus, at low SNR, how the transmitter allocates power among antennas is crucially important.

Considering the fact that in reality, obtaining CSIT within a reasonable accuracy requires significant effort, especially at low SNR, we evaluated the performance loss of water-filling power allocation due to channel estimation error at the transmitter as a function of the SER. In this regard, we derived the throughput achieved in MIMO systems if the transmitter performs water-filling based on erroneous channel estimates only and showed that even though the channel gain matrix is random, at low SNR, the normalized water-filling throughput based on imperfect CSIT converges to a non-random value denoted by R, almost surely as n increases. Subsequently, using C^{P} as a baseline for comparison, we compared R with C^{P} as a function of the SER and found that at low SNR, for moderate values of the SER, water-filling based on erroneous channel estimates can still achieve significant throughputs asymptotically. In particular, for SER values such as 5 dB, 0 dB, and -5 dB, R is found to be 86%, 70%, and 52% of C^{P}, respectively.

In addition, we showed that at low SNR and low SER, water-filling based on erroneous channel estimates asymptotically achieves the same throughput as equal power allocation. This result implies that at low SNR, water-filling power allocation does not perform worse than equal power allocation even when the channel estimates are very noisy.

The asymptotic analysis in Chapter 3, which is valid when the number of antennas is very large, can be applied to evaluate the throughput results in the massive MIMO scenario that has received considerable attention recently in academia and industry. In practice, when the number of antennas is very large, it becomes extremely challenging to obtain accurate channel estimates at the transmitter. Thus, our results in the case of imperfect CSIT can give an insight on how water-filling performs in such systems when channel estimation error is present.

4.2 Directions for Future Work

The imperfect CSI analysis presented in this thesis can be applied to evaluate the performance of interference channels, where multiple transmit and receive user pairs communicate using the same radio resources. The interference channel is a good model for communication in cellular networks for example. Interference alignment is a proposed cooperative strategy for interference management in interference channels. By coordinating their transmissions, users can align their interfering signals in time, frequency, or space. This can provide substantial performance improvements assuming perfect CSI is available [87]. However, with channel estimation error, the same improvements are not guaranteed and thus it is important to evaluate the robustness of such a strategy to channel imperfections.

Throughout this thesis, the channel gains were assumed to be fixed for the entire duration of transmission and thus, it was not considered how rapidly the channel gains change with time. In scenarios such as mmWave communications, the wavelength is very small which leads to quick variations in channel gains with time (e.g. a slight movement of the transmitter can result in a dramatic change in channel gains) [69]. There has been recent interest in using mmWave bands for the radio access of the 5th generation of mobile systems [88] and WiGig local area networks [89]. Thus, it would be worthwhile to consider such rapid channel fluctuations in the model as well. Having such a channel model, the two-dimensional space-time water-filling scheme proposed in [90] is shown to achieve better throughputs than the one-dimensional spatial water-filing scheme that was considered in Chapter 3. Thus, a possible direction for future work is to evaluate the performance of space-time water-filling approach with imperfect CSIT and investigate how the throughput results are affected accordingly.

In Chapter 3, we assumed that the environment is rich scattering and the antennas are far apart so that the entries of the channel gain matrix $\mathbf{H}$ are independent. However, in practice, these assumptions may not be valid. This is because in reality the antennas can not be spaced very far apart and thus they will experience coupling. Furthermore, the propagation environment is not necessarily rich scattering. Therefore, the entries of $\mathbf{H}$ are not independent anymore and are thus correlated [19]. In such a case too, using appropriate

models, the performance of equal power allocation, water-filling power allocation with perfect CSIT, and water-filling power allocation with imperfect CSIT can be examined and compared similarly to the approach in Chapter 3.

APPENDICES

Appendix A

Derivation of $\mathsf{E}[G^{\text{null, Im}}]$

In this appendix, we derive the expected value of the actual received power at the Rx in the case of null-steering beamforming with imperfect CSIT.

Using (2.26), $G^{\text{null, Im}}$ in (2.28) can be written as

$$
\begin{aligned}
G^{\text{null, Im}} &= P \left| \|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\| - \frac{\sum_{l=m+1}^{n} \tilde{w}_l \tilde{\hat{g}}_l^*}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|} \right|^2 \\
&= P \sum_{l=m+1}^{n} |\tilde{\hat{g}}_l|^2 - P \sum_{l=m+1}^{n} \tilde{w}_l \tilde{\hat{g}}_l^* - P \sum_{l=m+1}^{n} \tilde{w}_l^* \tilde{\hat{g}}_l + D + E, \quad\quad (\text{A.1})
\end{aligned}
$$

where

$$
D = \frac{P \sum_{l=m+1}^{n} |\tilde{w}_l|^2 |\tilde{\hat{g}}_l|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2},
$$

$$
E = \frac{P \sum_{\substack{l,r \\ l \neq r}}^{n} \tilde{w}_l \tilde{w}_r^* \tilde{\hat{g}}_l^* \tilde{\hat{g}}_r}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}.
$$

The expected value of the first three terms in (A.1) are easy to find since we know

$$
\mathsf{E}\left[\sum_{l=m+1}^{n} |\tilde{g}_l|^2 \right] = 2\sigma^2 (n-m)(K+1),
$$

$$
\mathsf{E}\left[\sum_{l=m+1}^{n} |\tilde{\hat{g}}_l|^2 \right] = \mathsf{E}\left[\sum_{l=m+1}^{n} |\tilde{g}_l|^2 \right] + \mathsf{E}\left[\sum_{l=m+1}^{n} |\tilde{w}_l|^2 \right] = 2\sigma^2 (n-m)\left(K+1+\frac{\sigma_e^2}{\sigma^2} \right),
$$

91

and

$$\mathsf{E}\left[\sum_{l=m+1}^{n} \tilde{w}_l \tilde{\hat{g}}_l^*\right] = \mathsf{E}\left[\sum_{l=m+1}^{n} \tilde{w}_l (\tilde{g}_l + \tilde{w}_l)^*\right] = \mathsf{E}\left[\sum_{l=m+1}^{n} |\tilde{w}_l|^2\right] = 2\sigma_e^2(n-m),$$

$$\mathsf{E}\left[\sum_{l=m+1}^{n} \tilde{w}_l^* \tilde{\hat{g}}_l\right] = \mathsf{E}\left[\sum_{l=m+1}^{n} \tilde{w}_l^* (\tilde{g}_l + \tilde{w}_l)\right] = \mathsf{E}\left[\sum_{l=m+1}^{n} |\tilde{w}_l|^2\right] = 2\sigma_e^2(n-m).$$

Conditioning on the rotation matrix $\mathbf{U} = \mathbf{U_m}' \ldots \mathbf{U_1}'$ and the entries of $\tilde{\hat{\mathbf{g}}}$, we can write the expected value of the last two terms in (A.1) as

$$\mathsf{E}[D] = \mathsf{E}_{\mathbf{U}}[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}[\mathsf{E}[D|\mathbf{U}, \tilde{\hat{\mathbf{g}}}]]],$$

$$\mathsf{E}[E] = \mathsf{E}_{\mathbf{U}}[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}[\mathsf{E}[E|\mathbf{U}, \tilde{\hat{\mathbf{g}}}]]].$$

Therefore, we obtain

$$\mathsf{E}[D] = P \sum_{l=m+1}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\mathsf{E}\left[\left.\frac{|\tilde{w}_l|^2 |\tilde{\hat{g}}_l|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\right| \mathbf{U}, \tilde{\hat{\mathbf{g}}}\right]\right]\right]$$

$$= P \sum_{l=m+1}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\left.\frac{\mathsf{E}\left[|\tilde{w}_l|^2 | \mathbf{U}, \tilde{\hat{g}}_l\right] |\tilde{\hat{g}}_l|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\right| \mathbf{U}\right]\right], \tag{A.2}$$

and

$$\mathsf{E}[E] = P \sum_{\substack{l,r \\ l \neq r}}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\mathsf{E}\left[\left.\frac{\tilde{w}_l \tilde{w}_r^* \tilde{\hat{g}}_l^* \tilde{\hat{g}}_r}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\right| \mathbf{U}, \tilde{\hat{\mathbf{g}}}\right]\right]\right]$$

$$= P \sum_{\substack{l,r \\ l \neq r}}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\left.\frac{\mathsf{E}\left[\tilde{w}_l | \mathbf{U}, \tilde{\hat{g}}_l\right] \mathsf{E}\left[\tilde{w}_r^* | \mathbf{U}, \tilde{\hat{g}}_r\right] \tilde{\hat{g}}_l^* \tilde{\hat{g}}_r}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\right| \mathbf{U}\right]\right]. \tag{A.3}$$

Since we have $\tilde{g}_l = \tilde{\mu}_l + \tilde{u}_l$ and $\tilde{g}_l + \tilde{w}_l = \tilde{\hat{g}}_l$, we can write

$$\mathsf{E}\left[|\tilde{w}_l|^2 | \mathbf{U}, \tilde{\hat{g}}_l\right] = \mathsf{E}\left[|\tilde{w}_l|^2 | \mathbf{U}, \tilde{g}_l + \tilde{w}_l = \tilde{\hat{g}}_l\right] = \mathsf{E}\left[|\tilde{w}_l|^2 | \mathbf{U}, \tilde{u}_l + \tilde{w}_l = \tilde{\hat{g}}_l - \tilde{\mu}_l\right].$$

Now, define $X = \tilde{u}_l + \tilde{w}_l$ and $Y = -\frac{\sigma_e^2}{\sigma^2}\tilde{u}_l + \tilde{w}_l$. Random variables X and Y are independent since they are zero-mean CSCG-distributed and satisfy $\mathsf{E}[XY] = 0$. Writing

$\tilde{w}_l$ in terms of X and Y as $\tilde{w}_l = \frac{\sigma_e^2}{\sigma^2+\sigma_e^2}X + \frac{\sigma^2}{\sigma^2+\sigma_e^2}Y$ and knowing $\mathsf{E}\left[|Y|^2|\mathbf{U}\right] = 2\sigma_e^2(1+\frac{\sigma_e^2}{\sigma^2})$, we obtain

$$
\begin{aligned}
\mathsf{E}\left[|\tilde{w}_l|^2|\mathbf{U},\tilde{\hat{g}}_l\right] &= \mathsf{E}\left[\left|\frac{\sigma_e^2}{\sigma^2+\sigma_e^2}X + \frac{\sigma^2}{\sigma^2+\sigma_e^2}Y\right|^2\,\middle|\,\mathbf{U}, X = \tilde{\hat{g}}_l - \tilde{\mu}_l\right] \\
&= \frac{\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\mathsf{E}\left[|X|^2|\mathbf{U}, X = \tilde{\hat{g}}_l - \tilde{\mu}_l\right] + \frac{\sigma^4}{(\sigma^2+\sigma_e^2)^2}\mathsf{E}\left[|Y|^2\,\middle|\,\mathbf{U}\right] \\
&= \frac{\sigma_e^4|\tilde{\hat{g}}_l - \tilde{\mu}_l|^2}{(\sigma^2+\sigma_e^2)^2} + \frac{2\sigma^2\sigma_e^2}{\sigma^2+\sigma_e^2}.
\end{aligned}
\tag{A.4}
$$

Following the same arguments as above, we can show that

$$
\mathsf{E}\left[\tilde{w}_l|\mathbf{U},\tilde{\hat{g}}_l\right] = \frac{\sigma_e^2\left(\tilde{\hat{g}}_l - \tilde{\mu}_l\right)}{\sigma^2+\sigma_e^2},
\tag{A.5}
$$

$$
\mathsf{E}\left[\tilde{w}_r^*|\mathbf{U},\tilde{\hat{g}}_r\right] = \frac{\sigma_e^2(\tilde{\hat{g}}_r - \mu_r)^*}{\sigma^2+\sigma_e^2}.
\tag{A.6}
$$

Therefore, following from (A.2) and using (A.4), we obtain

$$
\begin{aligned}
\mathsf{E}[D] &= \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\sum_{l=m+1}^{n}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_l - \tilde{\mu}_l|^2|\tilde{\hat{g}}_l|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\,\middle|\,\mathbf{U}\right]\right] \\
&\quad + \frac{2P\sigma^2\sigma_e^2}{\sigma^2+\sigma_e^2}\sum_{l=m+1}^{n}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_l|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\,\middle|\,\mathbf{U}\right]\right] \\
&= \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\sum_{l=m+1}^{n}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_l|^4}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\,\middle|\,\mathbf{U}\right]\right] \tag{A.7} \\
&\quad + \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\sum_{l=m+1}^{n}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_l|^2|\tilde{\mu}_l|^2}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\,\middle|\,\mathbf{U}\right]\right] \tag{A.8} \\
&\quad - \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\sum_{l=m+1}^{n}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_l|^2\tilde{\hat{g}}_l\tilde{\mu}_l^*}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\,\middle|\,\mathbf{U}\right]\right] \tag{A.9} \\
&\quad - \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\sum_{l=m+1}^{n}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_l|^2\tilde{\hat{g}}_l^*\tilde{\mu}_l}{\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\|^2}\,\middle|\,\mathbf{U}\right]\right] \tag{A.10} \\
&\quad + \frac{2P\sigma^2\sigma_e^2}{\sigma^2+\sigma_e^2},
\end{aligned}
$$

93

and following from (A.3) and using (A.5) and (A.6), we find

$$\mathsf{E}[E] = \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \sum_{\substack{l,r \\ l \neq r}}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{\left(\hat{\tilde{g}}_l - \tilde{\mu}_l\right)\left(\hat{\tilde{g}}_r - \tilde{\mu}_r\right)^* \hat{\tilde{g}}_l^* \hat{\tilde{g}}_r}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right]$$

$$= \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \sum_{\substack{l,r \\ l \neq r}}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{|\hat{\tilde{g}}_l|^2 |\hat{\tilde{g}}_r|^2}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right] \qquad (A.11)$$

$$+ \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \sum_{\substack{l,r \\ l \neq r}}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{\hat{\tilde{g}}_l^* \hat{\tilde{g}}_r \tilde{\mu}_l \tilde{\mu}_r^*}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right] \qquad (A.12)$$

$$- \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \sum_{\substack{l,r \\ l \neq r}}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{|\hat{\tilde{g}}_l|^2 \hat{\tilde{g}}_r \tilde{\mu}_r^*}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right] \qquad (A.13)$$

$$- \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \sum_{\substack{l,r \\ l \neq r}}^{n} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{|\hat{\tilde{g}}_r|^2 \hat{\tilde{g}}_l^* \tilde{\mu}_l}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right]. \qquad (A.14)$$

Combination of corresponding terms in $\mathsf{E}[D]$ and $\mathsf{E}[E]$ ((A.7) with (A.11), (A.8) with (A.12), (A.9) with (A.13), and (A.10) with (A.14)), yields

$$\mathsf{E}[D + E] = \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^4}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right]$$

$$+ \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{\left(\sum_{l=m+1}^{n} \hat{\tilde{g}}_l^* \tilde{\mu}_l\right)\left(\sum_{r=m+1}^{n} \hat{\tilde{g}}_r^* \tilde{\mu}_r\right)^*}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right]$$

$$- \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{\left(\sum_{l=m+1}^{n} |\hat{\tilde{g}}_l|^2\right)\left(\sum_{r=m+1}^{n} \hat{\tilde{g}}_r \tilde{\mu}_r^*\right)}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right]$$

$$- \frac{P\sigma_e^4}{(\sigma^2 + \sigma_e^2)^2} \mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\hat{\tilde{\mathbf{g}}}}\left[\left.\frac{\left(\sum_{l=m+1}^{n} |\hat{\tilde{g}}_l|^2\right)\left(\sum_{r=m+1}^{n} \hat{\tilde{g}}_r^* \tilde{\mu}_r\right)}{\|\hat{\tilde{\mathbf{g}}}_{-\mathbf{m}}\|^2} \right| \mathbf{U} \right]\right]$$

$$+ \frac{2P\sigma^2 \sigma_e^2}{\sigma^2 + \sigma_e^2},$$

which can be simplified as

$$\mathsf{E}[D+E] = \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\mathbf{g}}}\left[\left\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\right\|^2\middle|\mathbf{U}\right]\right]$$

$$+ \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\mathbf{g}}}\left[\frac{\left(\sum_{l=m+1}^{n}\tilde{\hat{g}}_l^*\tilde{\mu}_l\right)\left(\sum_{r=m+1}^{n}\tilde{\hat{g}}_r^*\tilde{\mu}_r\right)^*}{\left\|\tilde{\hat{\mathbf{g}}}_{-\mathbf{m}}\right\|^2}\middle|\mathbf{U}\right]\right]$$

$$- \frac{2P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\mathsf{E}_{\mathbf{U}}\left[\mathsf{E}_{\tilde{\mathbf{g}}}\left[\sum_{r=m+1}^{n}\tilde{\hat{g}}_r\tilde{\mu}_r^*\middle|\mathbf{U}\right]\right]$$

$$+ \frac{2P\sigma^2\sigma_e^2}{\sigma^2+\sigma_e^2}.$$

The second term in $\mathsf{E}[D+E]$ above is positive. Thus, a lower bound on $\mathsf{E}[D+E]$ can be obtained as

$$\mathsf{E}[D+E] \geq \frac{P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\mathsf{E}_{\mathbf{U}}\left[2\sigma^2(n-m)\left(K+1+\frac{\sigma_e^2}{\sigma^2}\right)\right]$$

$$- \frac{2P\sigma_e^4}{(\sigma^2+\sigma_e^2)^2}\mathsf{E}_{\mathbf{U}}\left[\sum_{r=m+1}^{n}\mathsf{E}\left[|\tilde{\mu}_r|^2\right]\right]$$

$$+ \frac{2P\sigma^2\sigma_e^2}{\sigma^2+\sigma_e^2},$$

which is found knowing that $\mathsf{E}_{\tilde{\mathbf{g}}}[\tilde{\hat{g}}_r\tilde{\mu}_r^*|\mathbf{U}] = \mathsf{E}\left[|\tilde{\mu}_r|^2\right]$ for each r. Note that the above lower bound is tight when $K=0$ ($\|\boldsymbol{\mu}\|=0$).

Therefore, since from Section 2.3, $\sum_{r=m+1}^{n}\mathsf{E}\left[|\tilde{\mu}_r|^2\right] = \|\boldsymbol{\mu}\|^2(n-m)/n = 2\sigma^2(n-m)K$, we derive

$$\mathsf{E}[D+E] \geq \frac{2P\sigma_e^4\sigma^2(n-m)\left(K+1+\sigma_e^2/\sigma^2\right)}{(\sigma^2+\sigma_e^2)^2} - \frac{4P\sigma_e^4\sigma^2(n-m)K}{(\sigma^2+\sigma_e^2)^2} + \frac{2P\sigma^2\sigma_e^2}{\sigma^2+\sigma_e^2}.$$

Thus, a lower bound on $\mathsf{E}[G^{\text{null, Im}}]$ can be found as

$$\mathsf{E}[G^{\text{null, Im}}] \geq 2P\sigma^2(n-m)\left(K+1-\frac{\sigma_e^2}{\sigma^2}+\frac{\sigma_e^4\sigma^2(1-K)+\sigma_e^6}{\sigma^2\left(\sigma^2+\sigma_e^2\right)^2}+\frac{\sigma_e^2}{(\sigma^2+\sigma_e^2)(n-m)}\right)$$

$$= 2P\sigma^2(n-m)\left(K+1-\frac{1}{\gamma}+\frac{1-K+\frac{1}{\gamma}}{(1+\gamma)^2}+\frac{1}{(n-m)(1+\gamma)}\right),$$

with equality when $K=0$ ($\|\boldsymbol{\mu}\|=0$) where $\gamma=\frac{\sigma^2}{\sigma_e^2}$.

Appendix B

Derivation of $\mathsf{E}[I^{\text{th, Im}}]$

In this appendix, we derive the expected value of the actual interference power at Rx_1 in the case of ϵ-threshold beamforming with imperfect CSIT, with the following two different cases to consider:

1. **Case 1; when** $\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}$: In this case, the optimal solution to **P8** in Section 2.4.2 is $\mathbf{y}^{\text{opt}} = \frac{\tilde{\hat{\mathbf{g}}}^*}{\|\tilde{\hat{\mathbf{g}}}\|}\sqrt{P}$. Therefore, replacing $\mathbf{y}^{\text{opt}}$ in (2.58), the actual interference power at Rx_1 is

$$I_1^{\text{th, Im}} = \frac{P\|\hat{\mathbf{h}}\|^2|\hat{g}_1|^2}{\|\tilde{\hat{\mathbf{g}}}\|^2} + \frac{P}{\|\tilde{\hat{\mathbf{g}}}\|^2}\left|\sum_{i=1}^{n}\tilde{v}_i\tilde{\hat{g}}_i^*\right|^2 - \frac{P\tilde{\hat{g}}_1^*\|\hat{\mathbf{h}}\|}{\|\tilde{\hat{\mathbf{g}}}\|^2}\sum_{i=1}^{n}\tilde{v}_i^*\tilde{g}_i - \frac{P\tilde{g}_1\|\hat{\mathbf{h}}\|}{\|\tilde{\hat{\mathbf{g}}}\|^2}\sum_{i=1}^{n}\tilde{v}_i\tilde{\hat{g}}_i^*. \quad \text{(B.1)}$$

Consequently, we can write

$$\mathsf{E}\left[I_1^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right] = P\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_1|^2}{\|\tilde{\hat{\mathbf{g}}}\|^2}\mathsf{E}_{\hat{\mathbf{h}}}\left[\|\hat{\mathbf{h}}\|^2 \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\Big|\tilde{\hat{\mathbf{g}}}\right]\right] \quad \text{(B.2)}$$

$$+ P\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\mathsf{E}_{\hat{\mathbf{h}}}\left[|\tilde{v}_1|^2 \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\Big|\tilde{\hat{\mathbf{g}}}\right]\right] \quad \text{(B.3)}$$

$$+ Pn(n-1)\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{\tilde{\hat{g}}_1^*\tilde{\hat{g}}_2}{\|\tilde{\hat{\mathbf{g}}}\|^2}\mathsf{E}_{\hat{\mathbf{h}}}\left[\tilde{v}_1\tilde{v}_2^* \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\Big|\tilde{\hat{\mathbf{g}}}\right]\right] \quad \text{(B.4)}$$

$$- 2P\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{|\tilde{\hat{g}}_1|^2}{\|\tilde{\hat{\mathbf{g}}}\|^2}\mathsf{E}_{\hat{\mathbf{h}}}\left[\|\hat{\mathbf{h}}\|\tilde{v}_1^* \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\Big|\tilde{\hat{\mathbf{g}}}\right]\right] \quad \text{(B.5)}$$

$$- 2P\sum_{i=2}^{n}\mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[\frac{\tilde{\hat{g}}_1\tilde{\hat{g}}_i^*}{\|\tilde{\hat{\mathbf{g}}}\|^2}\mathsf{E}_{\hat{\mathbf{h}}}\left[\|\hat{\mathbf{h}}\|\tilde{v}_i \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\Big|\tilde{\hat{\mathbf{g}}}\right]\right], \quad \text{(B.6)}$$

where the term (B.2) corresponds to the first term of (B.1), the terms (B.3) and (B.4) correspond to the second term of (B.1), and the terms (B.5) and (B.6) correspond to the last two terms of (B.1).

Because of spherical symmetry, we have

$$\mathsf{E}_{\hat{\mathbf{h}}}\left[\tilde{v}_i \tilde{v}_j^* \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}} |\tilde{\tilde{\mathbf{g}}}\right] = 0,$$

and

$$\mathsf{E}_{\hat{\mathbf{h}}}\left[\|\hat{\mathbf{h}}\| \tilde{v}_i \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}} |\tilde{\tilde{\mathbf{g}}}\right] = 0,$$

for all $i, j = 1, 2, \ldots, n$ where $i \neq j$, which results in the terms (B.4), (B.5), and (B.6) being zero.

Since the random variable $\|\hat{\mathbf{h}}\|^2/\sigma^2$ is chi-square distributed with $2n$ degrees of freedom, knowing its probability distribution function , we can write

$$\begin{aligned}
\mathsf{E}_{\hat{\mathbf{h}}}\left[\|\hat{\mathbf{h}}\|^2 \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\right\}}\right] &= \frac{\sigma^2 + \sigma_e^2}{2^n \Gamma(n)} \int_0^{\frac{\epsilon}{P(\sigma^2+\sigma_e^2)}} x^{n+1} e^{-\frac{x}{2}} dx \\
&= \frac{\sigma^2 + \sigma_e^2}{2^n \Gamma(n)} \int_0^{\frac{\epsilon}{P(\sigma^2+\sigma_e^2)}} x^{n+1} \left(1 + \sum_{n=1}^{\infty} \frac{(-1)^n (x^n)}{2^n}\right) dx \\
&= \frac{4(\sigma^2 + \sigma_e^2)}{(n+2)\Gamma(n)} \left(\frac{\epsilon}{2P(\sigma^2 + \sigma_e^2)}\right)^{n+2} + O\left(\left(\frac{\epsilon}{P}\right)^{n+3}\right).
\end{aligned}$$

Consequently, the term (B.2) is found to be

$$\mathsf{E}_{\tilde{\mathbf{g}}}\left[\frac{|\tilde{\hat{g}}_1|^2}{\|\tilde{\tilde{\mathbf{g}}}\|^2} \mathsf{E}_{\hat{\mathbf{h}}}\left[\|\hat{\mathbf{h}}\|^2 \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}} |\tilde{\tilde{\mathbf{g}}}\right]\right] = \mathsf{E}\left[f_5(\tilde{\mathbf{g}})\right] O\left(\left(\frac{\epsilon}{P}\right)^{n+2}\right),$$

where f_5 is a suitable function of the entries of $\tilde{\tilde{\mathbf{g}}}$.

Therefore, we derive

$$\mathsf{E}\left[I_1^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right] = P \mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\hat{\mathbf{h}}}\left[|\tilde{v}_1|^2 \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}} |\tilde{\tilde{\mathbf{g}}}\right]\right] + O\left(\left(\frac{\epsilon}{P}\right)^{n+2}\right).$$

(B.7)

2. **Case 2; when $\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}$** : In this case, the optimal solution to **P8** in Section 2.4.2 is $y_1^{\mathsf{opt}} = \frac{\sqrt{\epsilon}e^{-j\arg(\tilde{\hat{g}}_1)}}{\|\hat{\mathbf{h}}\|}$ and $\mathbf{y_{-1}}^{\mathsf{opt}} = \frac{\tilde{\hat{\mathbf{g}}}_{\mathbf{-1}}^*}{\|\tilde{\hat{\mathbf{g}}}_{\mathbf{-1}}\|}\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}$. Therefore, replacing $\mathbf{y}^{\mathsf{opt}}$ in (2.58), the actual interference power at Rx_1 can be written as

$$
\begin{aligned}
I_2^{\text{th, Im}} = \epsilon + &\left| \frac{\sqrt{\epsilon}}{\|\hat{\mathbf{h}}\|} e^{-j\arg(\tilde{\hat{g}}_1)} \tilde{v}_1 + \frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\tilde{\hat{\mathbf{g}}}_{\mathbf{-1}}\|} \sum_{i=2}^{n} \tilde{v}_i \tilde{\hat{g}}_i^* \right|^2 \\
& - \sqrt{\epsilon} e^{-j\arg(\tilde{\hat{g}}_1)} \left(\frac{\sqrt{\epsilon}}{\|\hat{\mathbf{h}}\|} e^{j\arg(\tilde{\hat{g}}_1)} \tilde{v}_1^* + \frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\tilde{\hat{\mathbf{g}}}_{\mathbf{-1}}\|} \sum_{i=2}^{n} \tilde{v}_i^* \tilde{\hat{g}}_i \right) \\
& - \sqrt{\epsilon} e^{j\arg(\tilde{\hat{g}}_1)} \left(\frac{\sqrt{\epsilon}}{\|\hat{\mathbf{h}}\|} e^{-j\arg(\tilde{\hat{g}}_1)} \tilde{v}_1 + \frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\tilde{\hat{\mathbf{g}}}_{\mathbf{-1}}\|} \sum_{i=2}^{n} \tilde{v}_i \tilde{\hat{g}}_i^* \right).
\end{aligned}
$$

Consequently, we can write

$$
\begin{aligned}
\mathsf{E}\left[I_2^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}} \right] = {}& \epsilon \, \mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\hat{\mathbf{h}}}\left[\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}} \Big| \tilde{\hat{\mathbf{g}}} \right] \right] \\
& + \epsilon \, \mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\hat{\mathbf{h}}}\left[\frac{|\tilde{v}_1|^2}{\|\hat{\mathbf{h}}\|^2} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}} \Big| \tilde{\hat{\mathbf{g}}} \right] \right] \\
& + \mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\hat{\mathbf{h}}}\left[|\tilde{v}_2|^2 \left(P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2} \right) \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}} \Big| \tilde{\hat{\mathbf{g}}} \right] \right]. \quad \text{(B.8)}
\end{aligned}
$$

Note that in the derivation of (B.8), similarly to the previous case, we use the fact that some terms are zero due to spherical symmetry.

Using (2.44) with $\hat{\mathbf{h}}$ in place of $\mathbf{h}$, we can write

$$
\epsilon \, \mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\hat{\mathbf{h}}}\left[\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}} \Big| \tilde{\hat{\mathbf{g}}} \right] \right] = \epsilon - O\left(\left(\frac{\epsilon}{P} \right)^n \right),
$$

and knowing that $\tilde{v}_1$ is equal in distribution with $\tilde{v}_2$, we find

$$
\mathsf{E}\left[I_2^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}} \right] = \epsilon + P\mathsf{E}_{\tilde{\mathbf{g}}}\left[\mathsf{E}_{\hat{\mathbf{h}}}\left[|\tilde{v}_2|^2 \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}} \Big| \tilde{\hat{\mathbf{g}}} \right] \right] - O\left(\left(\frac{\epsilon}{P} \right)^n \right).
$$

$$\text{(B.9)}$$

Thus, by combining (B.7) and (B.9) for $\epsilon = \alpha P$, we obtain

$$
\begin{aligned}
\mathsf{E}[I^{\text{th, Im}}] &= \mathsf{E}_{\hat{\mathbf{h}}}\left[I_1^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\hat{g}_1|^2}\right\}}\right] + \mathsf{E}_{\hat{\mathbf{h}}}\left[I_2^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\mathbf{g}}\|^2}{|\hat{g}_1|^2}\right\}}\right] \\
&= P\mathsf{E}\left[|\tilde{v}_1|^2\right] + \epsilon - O\left(\left(\frac{\epsilon}{P}\right)^n\right) \\
&= 2P\sigma_e^2 + \alpha P - O\left(\alpha^n\right).
\end{aligned}
$$

Appendix C

Derivation of $\mathsf{E}[G^{\text{th, Im}}]$

In this appendix, we derive the expected value of the actual received power at the Rx in the case of ϵ-threshold beamforming with imperfect CSIT, with the following two different cases to consider:

1. **Case 1; when** $\|\hat{\mathbf{h}}\|^2 \leq \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}$: In this case, the optimal solution to **P8** in Section 2.4.2 is $\mathbf{y}^{\mathsf{opt}} = \frac{\tilde{\hat{\mathbf{g}}}^*}{\|\tilde{\hat{\mathbf{g}}}\|}\sqrt{P}$. Therefore, replacing $\mathbf{y}^{\mathsf{opt}}$ in (2.60), the actual received power at the Rx is

$$G_1^{\text{th, Im}} = P\|\tilde{\hat{\mathbf{g}}}\|^2 + \frac{P}{\|\tilde{\hat{\mathbf{g}}}\|^2}\left|\sum_{i=1}^{n}\tilde{w}_i\tilde{\hat{g}}_i^*\right|^2 - P\sum_{i=1}^{n}\tilde{w}_i^*\tilde{\hat{g}}_i - P\sum_{i=1}^{n}\tilde{w}_i\tilde{\hat{g}}_i^*.$$

   Thus, using (2.44) with $\hat{\mathbf{h}}$ in place of $\mathbf{h}$, we obtain

$$\mathsf{E}\left[G_1^{\text{th, Im}}\cdot\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2\leq\frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right] = \mathsf{E}_{\tilde{\hat{\mathbf{g}}}}\left[G_1^{\text{th, Im}}\cdot\mathsf{Pr}\left\{\|\hat{\mathbf{h}}\|^2\leq\frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\bigg|\tilde{\hat{\mathbf{g}}}\right\}\right]$$
$$= \mathsf{E}\left[f_6(\tilde{\hat{\mathbf{g}}})\right]O\left(\left(\frac{\epsilon}{P}\right)^n\right), \tag{C.1}$$

   where f_6 is a suitable function of the entries of $\tilde{\hat{\mathbf{g}}}$.

2. **Case 2; when** $\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}$: In this case, the optimal solution to **P8** in Section 2.4.2 is $y_1^{\mathsf{opt}} = \frac{\sqrt{\epsilon}e^{-j\arg(\tilde{\hat{g}}_1)}}{\|\hat{\mathbf{h}}\|}$ and $\mathbf{y_{-1}}^{\mathsf{opt}} = \frac{\tilde{\hat{\mathbf{g}}}_{-1}^*}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|}\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}$. Therefore, replacing $\mathbf{y}^{\mathsf{opt}}$ in

(2.60), the actual received power at the Rx can be written as

$$G_2^{\text{th, Im}} = \left(\frac{|\tilde{\hat{g}}_1|\sqrt{\epsilon}}{\|\hat{\mathbf{h}}\|} + \sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}} \|\tilde{\hat{\mathbf{g}}}_{-1}\| \right)^2$$

$$+ \left| \frac{\tilde{w}_1\sqrt{\epsilon}e^{-j\arg(\tilde{\hat{g}}_1)}}{\|\tilde{\hat{\mathbf{h}}}\|} + \frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|} \sum_{i=2}^{n} \tilde{w}_i\tilde{\hat{g}}_i^{\,*} \right|^2$$

$$- \left(\frac{\tilde{\hat{g}}_1\sqrt{\epsilon}e^{-j\arg(\tilde{\hat{g}}_1)}}{\|\tilde{\hat{\mathbf{h}}}\|} + \sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}} \|\tilde{\hat{\mathbf{g}}}_{-1}\| \right)$$

$$\left(\frac{\tilde{w}_1^*\sqrt{\epsilon}e^{j\arg(\tilde{\hat{g}}_1)}}{\|\tilde{\hat{\mathbf{h}}}\|} + \frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|} \sum_{i=2}^{n} \tilde{w}_i^*\tilde{\hat{g}}_i \right)$$

$$- \left(\frac{\tilde{\hat{g}}_1^*\sqrt{\epsilon}e^{j\arg(\tilde{\hat{g}}_1)}}{\|\tilde{\hat{\mathbf{h}}}\|} + \sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}} \|\tilde{\hat{\mathbf{g}}}_{-1}\| \right)$$

$$\left(\frac{\tilde{w}_1\sqrt{\epsilon}e^{-j\arg(\tilde{\hat{g}}_1)}}{\|\tilde{\hat{\mathbf{h}}}\|} + \frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|} \sum_{i=2}^{n} \tilde{w}_i\tilde{\hat{g}}_i^{\,*} \right). \tag{C.2}$$

Denote the first term in (C.2) as A. Thus, we obtain

$$\mathsf{E}\left[A \cdot \mathbf{1}_{\left\{ \|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\tilde{\hat{g}}_1|^2} \right\}} \right] = 2P(n-1)\left(\sigma^2 + \sigma_e^2 \right)$$

$$+ 2\sqrt{2P\epsilon(\sigma^2 + \sigma_e^2)} \frac{\Gamma\left(\frac{3}{2}\right)\Gamma(n - \frac{1}{2})\Gamma\left(n + \frac{1}{2}\right)}{\Gamma(n)\Gamma(n-1)\left(n - \frac{1}{2}\right)}$$

$$- \epsilon\left(\frac{n-2}{n-1} \right) - O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}} \right), \tag{C.3}$$

which is derived following the same approach as in Section 2.4.1 with $\hat{\mathbf{h}}$ in place of $\mathbf{h}$ and knowing

$$\mathsf{E}\left[\|\tilde{\hat{\mathbf{g}}}_{-1}\|^2 \right] = \mathsf{E}\left[\sum_{i=2}^{n} |\tilde{g}_i + \tilde{w}_i|^2 \right] = \mathsf{E}\left[\sum_{i=2}^{n} |\tilde{g}_i|^2 \right] + \mathsf{E}\left[\sum_{i=2}^{n} |\tilde{w}_i|^2 \right] = 2(n-1)\left(\sigma^2 + \sigma_e^2 \right).$$

Denoting the second term in (C.2) as B, we obtain

$$
\mathsf{E}\left[B \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\tilde{g}_1|^2}\right\}}\right] = \epsilon\, \mathsf{E}_{\tilde{\mathbf{g}}}\left[|\tilde{w}_1|^2 \mathsf{E}_{\hat{\mathbf{h}}}\left[\frac{\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\tilde{g}_1|^2}\right\}}}{\|\hat{\mathbf{h}}\|^2}\Big|\tilde{\tilde{\mathbf{g}}}\right]\right]
$$

$$
+ \mathsf{E}_{\tilde{\tilde{\mathbf{g}}}}\left[\frac{\sum_{i=2}^{n}|\tilde{w}_i|^2|\tilde{\tilde{g}}_i|^2}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|^2}\mathsf{E}_{\hat{\mathbf{h}}}\left[\left(P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}\right)\cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\tilde{g}_1|^2}\right\}}\Big|\tilde{\tilde{\mathbf{g}}}\right]\right]
$$

$$
+ \mathsf{E}_{\tilde{\mathbf{g}}}\left[\frac{\sum_{i=2, i\neq j}^{n}\sum_{j=2}^{n}\tilde{w}_i^*\tilde{w}_j\tilde{\tilde{g}}_i\tilde{\tilde{g}}_j^*}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|^2}\mathsf{E}_{\hat{\mathbf{h}}}\left[\left(P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}\right)\cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\tilde{g}_1|^2}\right\}}\Big|\tilde{\tilde{\mathbf{g}}}\right]\right]
$$

$$
+ \sqrt{\epsilon}\, \mathsf{E}_{\tilde{\mathbf{g}}}\left[\frac{\tilde{w}_1^* e^{j\arg(\tilde{g}_1)}\sum_{i=2}^{n}\tilde{w}_i\tilde{\tilde{g}}_i^*}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|}\mathsf{E}_{\hat{\mathbf{h}}}\left[\frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\hat{\mathbf{h}}\|}\cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\tilde{g}_1|^2}\right\}}\Big|\tilde{\tilde{\mathbf{g}}}\right]\right]
$$

$$
+ \sqrt{\epsilon}\, \mathsf{E}_{\tilde{\mathbf{g}}}\left[\frac{\tilde{w}_1 e^{-j\arg(\tilde{g}_1)}\sum_{i=2}^{n}\tilde{w}_i^*\tilde{\tilde{g}}_i}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|}\mathsf{E}_{\hat{\mathbf{h}}}\left[\frac{\sqrt{P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\hat{\mathbf{h}}\|}\cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\tilde{g}_1|^2}\right\}}\Big|\tilde{\tilde{\mathbf{g}}}\right]\right]. \quad \text{(C.4)}
$$

Therefore, using the same argument as in (2.39), (2.48), and (2.49) with $\hat{\mathbf{h}}$ in place of $\mathbf{h}$, and knowing that the last two terms in (C.4) are equivalent due to spherical symmetry, we find

$$
\mathsf{E}\left[B \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\tilde{\mathbf{g}}}\|^2}{|\tilde{g}_1|^2}\right\}}\right] = \epsilon\mathsf{E}\left[|\tilde{w}_1|^2\right]\mathsf{E}\left[\frac{1}{\|\hat{\mathbf{h}}\|^2}\right]
$$

$$
+ \mathsf{E}\left[\frac{\sum_{i=2}^{n}|\tilde{w}_i|^2|\tilde{\tilde{g}}_i|^2}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|^2}\right]\mathsf{E}\left[\left(P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}\right)\right]
$$

$$
+ \mathsf{E}\left[\frac{\sum_{i=2, i\neq j}^{n}\sum_{j=2}^{n}\tilde{w}_i^*\tilde{w}_j\tilde{\tilde{g}}_i\tilde{\tilde{g}}_j^*}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|^2}\right]\mathsf{E}\left[\left(P - \frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}\right)\right]
$$

$$
+ 2\sqrt{\epsilon}\,\mathsf{E}\left[\tilde{w}_1^* e^{j\arg(\tilde{g}_1)}\right]\mathsf{E}\left[\frac{\sum_{i=2}^{n}\tilde{w}_i\tilde{\tilde{g}}_i^*}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|}\right]\mathsf{E}\left[\frac{\sqrt{P}}{\|\hat{\mathbf{h}}\|}\right]
$$

$$
+ O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right).
$$

Let

$$
W = \frac{\sum_{i=2}^{n}|\tilde{w}_i|^2|\tilde{\tilde{g}}_i|^2}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|^2}, \quad V = \frac{\sum_{i=2, i\neq j}^{n}\sum_{j=2}^{n}\tilde{w}_i^*\tilde{w}_j\tilde{\tilde{g}}_i\tilde{\tilde{g}}_j^*}{\|\tilde{\tilde{\mathbf{g}}}_{-1}\|^2}.
$$

102

Following the same approach as in Appendix A to find (A.2) and (A.3) (with $\|\boldsymbol{\mu}\|^2 = 0$ and $m = 1$), we can find

$$\mathsf{E}[W + V] = \frac{2\sigma_e^4}{\sigma^2 + \sigma_e^2}(n-1) + \frac{2\sigma^2\sigma_e^2}{\sigma^2 + \sigma_e^2}. \tag{C.5}$$

Furthermore, recall from Appendix A that $Y = -\frac{\sigma_e^2}{\sigma^2}\tilde{g}_i + \tilde{w}_i$ is independent of $\tilde{\hat{g}}_i = \tilde{g}_i + \tilde{w}_i$ for any i. Thus, writing $\tilde{w}_i$ in terms of $\tilde{\hat{g}}_i$ and Y_i as $\tilde{w}_i = \frac{\sigma_e^2}{\sigma^2 + \sigma_e^2}\tilde{\hat{g}}_i + \frac{\sigma^2}{\sigma^2 + \sigma_e^2}Y_i$, we obtain

$$
\begin{aligned}
\mathsf{E}\left[\tilde{w}_1^* e^{j \arg(\tilde{\hat{g}}_1)}\right] &= \mathsf{E}\left[\frac{\tilde{w}_1^* \tilde{\hat{g}}_1}{|\tilde{\hat{g}}_1|}\right] = \mathsf{E}_{\tilde{\hat{g}}_1}\left[\mathsf{E}\left[\frac{\tilde{w}_1^* \tilde{\hat{g}}_1}{|\tilde{\hat{g}}_1|}\Big| \tilde{\hat{g}}_1\right]\right] = \mathsf{E}_{\tilde{\hat{g}}_1}\left[\frac{\tilde{\hat{g}}_1}{|\tilde{\hat{g}}_1|}\mathsf{E}\left[\tilde{w}_1^*|\tilde{\hat{g}}_1\right]\right] \\
&= \mathsf{E}\left[\frac{\tilde{\hat{g}}_1}{|\tilde{\hat{g}}_1|}\left(\frac{\sigma_e^2}{\sigma^2 + \sigma_e^2}\tilde{\hat{g}}_1^* + \frac{\sigma^2}{\sigma^2 + \sigma_e^2}Y_1^*\right)\right] \\
&= \frac{\sigma_e^2}{\sigma^2 + \sigma_e^2}\mathsf{E}\left[|\tilde{\hat{g}}_1|\right] = \sqrt{2}\,\Gamma\left(\frac{3}{2}\right)\frac{\sigma_e^2}{\sqrt{\sigma^2 + \sigma_e^2}},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathsf{E}\left[\frac{\sum_{i=2}^n \tilde{w}_i \tilde{\hat{g}}_i^*}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|}\right] &= \mathsf{E}\left[\frac{\sum_{i=2}^n \tilde{\hat{g}}_i^* \mathsf{E}\left[\tilde{w}_i|\tilde{\hat{g}}_i\right]}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|}\right] \\
&= \mathsf{E}\left[\frac{\sum_{i=2}^n \mathsf{E}\left[\left(\frac{\sigma_e^2}{\sigma^2 + \sigma_e^2}\tilde{\hat{g}}_i + \frac{\sigma^2}{\sigma^2 + \sigma_e^2}Y_i\right)\tilde{\hat{g}}_i^*\right]}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|}\right] \\
&= \frac{\sigma_e^2}{\sigma^2 + \sigma_e^2}\mathsf{E}\left[\frac{\sum_{i=2}^n |\tilde{\hat{g}}_i|^2}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|}\right] = \frac{\sigma_e^2}{\sigma^2 + \sigma_e^2}\mathsf{E}\left[\|\tilde{\hat{\mathbf{g}}}_{-1}\|\right] \\
&= \frac{\sqrt{2}\,\sigma_e^2\,\Gamma(n - \frac{1}{2})}{\sqrt{\sigma^2 + \sigma_e^2}\,\Gamma(n-1)}.
\end{aligned}
$$

Therefore, knowing

$$\mathsf{E}\left[\frac{1}{\|\hat{\mathbf{h}}\|}\right] = \frac{\sqrt{2}\,\Gamma\left(n + \frac{1}{2}\right)}{\sqrt{\sigma^2 + \sigma_e^2}\,\Gamma(n)\,(2n - 1)},$$

and

$$\mathsf{E}\left[\frac{1}{\|\hat{\mathbf{h}}\|^2}\right] = \frac{1}{2(n-1)\,(\sigma^2 + \sigma_e^2)},$$

by using (2.48) and (2.49) with $\hat{\mathbf{h}}$ in place of $\mathbf{h}$, we obtain

$$
\begin{aligned}
\mathsf{E}\left[B \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P} \frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}}\right] = {} & \frac{2P\sigma_e^2}{\sigma^2 + \sigma_e^2}\left(\sigma_e^2(n-1) + \sigma^2\right) \\
& + 2\sqrt{2P\epsilon}\,\Gamma\left(\frac{3}{2}\right)\frac{\sigma_e^4\,\Gamma\left(n-\frac{1}{2}\right)\Gamma\left(n+\frac{1}{2}\right)}{(\sigma^2+\sigma_e^2)^{\frac{3}{2}}\,\Gamma(n-1)\Gamma(n)\left(n-\frac{1}{2}\right)} \\
& + \frac{\epsilon\sigma_e^2}{\sigma^2+\sigma_e^2}\left(\frac{1}{n-1} - \frac{\sigma_e^2}{\sigma^2+\sigma_e^2} - \frac{\sigma^2}{(n-1)\left(\sigma^2+\sigma_e^2\right)}\right) \\
& + O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right).
\end{aligned} \tag{C.6}
$$

Denoting the third term in (C.2) as C, we find

$$
\begin{aligned}
-\mathsf{E}\left[C \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}}\right] = {} & \epsilon\,\mathsf{E}_{\tilde{\mathbf{g}}}\left[|\tilde{w}_1|^2\mathsf{E}_{\hat{\mathbf{h}}}\left[\frac{\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}}}{\|\hat{\mathbf{h}}\|^2}\Big|\tilde{\mathbf{g}}\right]\right] \\
& + \sqrt{\epsilon}\,\mathsf{E}_{\tilde{\mathbf{g}}}\left[\frac{|\hat{\tilde{g}}_1|\sum_{i=2}^n \tilde{w}_i^* \tilde{g}_i}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|}\mathsf{E}_{\hat{\mathbf{h}}}\left[\frac{\sqrt{P-\frac{\epsilon}{\|\hat{\mathbf{h}}\|}}}{\|\hat{\mathbf{h}}\|}\cdot\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}}\Big|\tilde{\mathbf{g}}\right]\right] \\
& + \sqrt{\epsilon}\,\mathsf{E}_{\tilde{\mathbf{g}}}\left[\tilde{w}_1^* e^{j\arg(\tilde{g}_1)}\|\tilde{\hat{\mathbf{g}}}_{-1}\|\mathsf{E}_{\hat{\mathbf{h}}}\left[\frac{\sqrt{P-\frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}}}{\|\hat{\mathbf{h}}\|}\cdot\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}}\Big|\tilde{\mathbf{g}}\right]\right] \\
& + \mathsf{E}_{\tilde{\mathbf{g}}}\left[\sum_{i=2}^n \tilde{g}_i\tilde{w}_i^*\mathsf{E}_{\hat{\mathbf{h}}}\left[\left(P-\frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}\right)\cdot\mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}}\Big|\tilde{\mathbf{g}}\right]\right],
\end{aligned}
$$

which can be written as

$$
\begin{aligned}
-\mathsf{E}\left[C \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{\tilde{g}}_1|^2}\right\}}\right] = {} & \epsilon\,\mathsf{E}\left[|\tilde{w}_1|^2\right]\mathsf{E}\left[\frac{1}{\|\hat{\mathbf{h}}\|^2}\right] \\
& + \sqrt{\epsilon}\,\mathsf{E}\left[|\tilde{g}_1|\right]\mathsf{E}\left[\frac{\sum_{i=2}^n \tilde{w}_i^*\tilde{g}_i}{\|\tilde{\hat{\mathbf{g}}}_{-1}\|}\right]\mathsf{E}\left[\frac{\sqrt{P}}{\|\hat{\mathbf{h}}\|}\right] \\
& + \sqrt{\epsilon}\,\mathsf{E}\left[\tilde{w}_1 e^{-j\arg(\tilde{g}_1)}\right]\mathsf{E}\left[\|\tilde{\hat{\mathbf{g}}}_{-1}\|^2\right]\mathsf{E}\left[\frac{\sqrt{P}}{\|\hat{\mathbf{h}}\|}\right] \\
& + \mathsf{E}\left[\sum_{i=2}^n \tilde{g}_i\tilde{w}_i^*\right]\mathsf{E}\left[P-\frac{\epsilon}{\|\hat{\mathbf{h}}\|^2}\right] \\
& + O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right).
\end{aligned}
$$

Therefore, we derive

$$-\mathsf{E}\left[C \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right] = 2P\sigma_e^2(n-1) - \frac{\epsilon\,(n-2)\sigma_e^2}{(n-1)\,(\sigma^2 + \sigma_e^2)}$$

$$+ \frac{2\sqrt{2P\epsilon}\,\sigma\Gamma\left(\frac{3}{2}\right)\Gamma\left(n+\frac{1}{2}\right)\Gamma\left(n-\frac{1}{2}\right)}{\sqrt{\gamma(1+\gamma)}\Gamma(n)\Gamma\left(n-1\right)\left(n-\frac{1}{2}\right)} + O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right). \quad \text{(C.7)}$$

Thus, combining (C.3), (C.6), and (C.7) and knowing that the fourth term in (C.2) is equal in distribution with C, we can write

$$\mathsf{E}\left[G_2^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right] = \mathsf{E}\left[A \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right]$$

$$+ \mathsf{E}\left[B \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right]$$

$$+ 2\mathsf{E}\left[C \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right]$$

$$= Q + \sqrt{\frac{\epsilon}{P}}R + \frac{\epsilon}{P}T + O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right), \quad \text{(C.8)}$$

where having $\gamma = \sigma^2/\sigma_e^2$,

$$Q = 2P(n-1)\sigma^2\left(1 - \frac{1}{\gamma} + \frac{1}{\gamma(1+\gamma)} + \frac{1}{(n-1)(1+\gamma)}\right),$$

$$R = 2\sqrt{2}P\sigma\frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(n-\frac{1}{2}\right)\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(n)\Gamma(n-1)\left(n-\frac{1}{2}\right)}\left(\frac{\gamma}{1+\gamma}\right)^{\frac{3}{2}},$$

$$T = P\left(-1 + \frac{1}{n-1} - \frac{1}{(1+\gamma)^2} - \frac{\gamma}{(n-1)(1+\gamma)^2} + \frac{2n-3}{(n-1)(1+\gamma)}\right).$$

Therefore, combining (C.1) and (C.8), we find that the actual received power at the Rx has the expected value given by

$$\mathsf{E}[G^{\text{th, Im}}] = \mathsf{E}\left[G_1^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 \le \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right] + \mathsf{E}\left[G_2^{\text{th, Im}} \cdot \mathbf{1}_{\left\{\|\hat{\mathbf{h}}\|^2 > \frac{\epsilon}{P}\frac{\|\tilde{\hat{\mathbf{g}}}\|^2}{|\hat{g}_1|^2}\right\}}\right]$$

$$= Q + \sqrt{\frac{\epsilon}{P}}R + \frac{\epsilon}{P}T + O\left(\left(\frac{\epsilon}{P}\right)^{\frac{3}{2}}\right)$$

$$= Q + \sqrt{\alpha}R + \alpha T + O\left(\alpha^{\frac{3}{2}}\right),$$

for $\epsilon = \alpha P$.

REFERENCES

[1] "ICT facts and figures," *International Telecommunication Union, ICT Data and Statistics Division*, February 2013.

[2] *http://www.ntt.co.jp/inlab/e/org/as.html*.

[3] A. Molisch, *Wireless Communications*. Wiley, IEEE Press, 2011.

[4] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[5] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[6] W. Jakes, *Microwave Mobile communications*. Wiley, New York, 1974.

[7] A. Wittneben, "Basestation modulation diversity for digital simulcast," in *41st IEEE Vehicular Technology Conference, Gateway to the Future Technology in Motion.*, May 1991, pp. 848–853.

[8] D. G. Brennan, "Linear diversity combining techniques," *Proceedings of the IEEE*, vol. 91, no. 2, pp. 331–356, February 2003.

[9] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[10] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.

[11] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[12] T. Lo, "Maximal ratio transmission," *IEEE Transactions on Communications*, vol. 47, no. 10, pp. 1458–1461, October 1999.

[13] X. Feng and C. Leung, "A new optimal transmit and receive diversity scheme," in *IEEE Pacific Rim Conference on Communications, Computers and signal Processing (PACRIM. 2001)*, vol. 2, 2001, pp. 538–541.

[14] S. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. Calderbank, "Great expectations: the value of spatial diversity in wireless networks," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 219–270, February 2004.

[15] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[16] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Communications*, vol. 10, pp. 585–598, November 1999.

[17] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, "Integer-forcing linear receivers: A new low-complexity MIMO architecture," in *2010 IEEE 72nd Vehicular Technology Conference*, September 2010, pp. 1–5.

[18] ——, "Integer-forcing linear receivers," *IEEE Transactions on Information Theory*, no. 99, 2014.

[19] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 637–650, March 2002.

[20] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[21] M. Yuksel and E. Erkip, "Multiple-antenna cooperative wireless systems: A diversity multiplexing tradeoff perspective," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3371–3393, October 2007.

[22] N. Prasad and M. K. Varanasi, "Diversity and multiplexing tradeoff bounds for cooperative diversity schemes," in *Proceedings of International Symposium on Information Theory (ISIT 2004)*, Chicago, IL, June 2004.

[23] T. Marzetta, "BLAST training: estimating channel characteristics for high-capacity space-time wireless," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 1999.

[24] C. Budianu and L. Tong, "Channel estimation for space-time orthogonal block codes," in *IEEE International Conference on Communications (ICC 2001)*, vol. 4, 2001, pp. 1127–1131.

[25] J. Yang, "Channel state information in multiple antenna systems," Ph.D. dissertation, Georgia Institute of Technology, 2006.

[26] H. Meyer, M. Moeneclaey, and S. A. Fletchel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing.* Wiley, New York, 1997.

[27] B. Hassibi and B. Hochwald, "Optimal training in space-time systems," in *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*, vol. 1, October 2000, pp. 743–747.

[28] G. X. L. Tong and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 340–349, March 1994.

[29] A. Paulraj, C. Papadias, V. Reddy, and A. van der Even, *Blind Space-Time Signal Processing: Wireless Communications*. Prentice Hall, 1998.

[30] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[31] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, September 2001.

[32] J. Leinonen, J. Hamalainen, and M. Juntti, "Capacity analysis of downlink MIMO-OFDMA frequency allocation with imperfect feedback information," in *IEEE 10th Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2009)*, June 2009, pp. 141–145.

[33] A. Heidari and A. K. Khandani, "Closed-loop transmit diversity with imperfect feedback," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 9, pp. 2737–2741, September 2010.

[34] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341–1365, October 2008.

[35] D. Love, R. Heath, W. Santipach, and M. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, October 2004.

[36] D. Gerlach, "Adaptive transmitting antenna arrays at the base station in mobile radio networks," Ph.D. dissertation, Stanford University, 1995.

[37] P. Chevillat, J. Jelitto, and H. Truong, "Dynamic data rate and transmit power adjustment in IEEE 802.11 wireless LANs," *International Journal of Wireless Information Networks*, vol. 12, no. 3, pp. 123–145, July 2005.

[38] D. Senthilkumar and A. Krishnan, "Throughput analysis of IEEE 802.11 multirate WLANs with collision aware rate adaptation algorithm," *International Journal of Automation and Computing*, vol. 7, no. 4, pp. 571–577, November 2010.

[39] H. V. Trees, *Detection, Estimation, and Modulation Theory, Part I.* New York: John Wiley and Sons, 1968.

[40] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, February 2005.

[41] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[42] A. Khandekar, N. Bhushan, J. Tingfang, and V. Vanghi, "LTE-advanced: Heterogeneous networks," in *2010 European Wireless Conference (EW)*, April 2010, pp. 978–982.

[43] H. Dhillon, R. Ganti, F. Baccelli, and J. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, April 2012.

[44] "Spectrum policy task force report of the spectrum efficiency working group," *Federal Communications Commission, Technical Report 02-135*, November 2002.

[45] "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies, notice of proposed rule making and order," *Federal Communications Commission, Technical Report 03-322*, December 2003.

[46] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, KTH Royal Institute of Technology, 2000.

[47] "Notice of proposed rule making and order," *Federal Communications Commission, Technical Report 03-222*, December 2003.

[48] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[49] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[50] 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 Journal (Elsevier)*, September 2006.

[51] X. Shang, B. Chen, and H. V. Poor, "Multiuser MISO interference channels with single-user detection: Optimality of beamforming and the achievable rate region," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4255–4273, July 2011.

[52] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5450–5458, October 2010.

[53] R. Zhang and Y. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 88–102, February 2008.

[54] R. Mochaourab and E. Jorswieck, "Optimal beamforming in interference networks with perfect local channel information," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1128–1141, March 2011.

[55] Y. Pei, Y. Liang, L. Zhang, K. Teh, and K. H. Li, "Secure communication over MISO cognitive radio channels," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1494–1502, April 2010.

[56] K. Wang, T. Chang, W. Ma, A.-C. So, and C.-Y. Chi, "Probabilistic SINR constrained robust transmit beamforming: A bernstein-type inequality based conservative approach," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 3080–3083.

[57] R. Zhang, F. Gao, and Y.-C. Liang, "Cognitive beamforming made practical: Effective interference channel and learning-throughput tradeoff," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 706–718, February 2010.

[58] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, First 2009.

[59] G. Zheng, S. Ma, K.-K. Wong, and T.-S. Ng, "Robust beamforming in cognitive radio," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 570–576, February 2010.

[60] Y. Pei, Y.-C. Liang, K. C. Teh, and K. H. Li, "Secure communication in multiantenna cognitive radio networks with imperfect channel state information," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1683–1693, April 2011.

[61] L. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4143–4153, August 2009.

[62] H. Claussen, "Performance of macro- and co-channel femtocells in a hierarchical cell structure," in *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2007)*, September 2007, pp. 1–5.

[63] T. Qu, D. Xiao, D. Yang, W. Jin, and Y. He, "Cell selection analysis in outdoor heterogeneous networks," in *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, vol. 5, August 2010.

[64] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, September 2008.

[65] D. Hu and S. Mao, "Resource allocation for medium grain scalable videos over femtocell cognitive radio networks," in *2011 31st International Conference on Distributed Computing Systems (ICDCS)*, June 2011, pp. 258–267.

[66] F. Rusek, D. Persson, B. K.L., E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, January 2013.

[67] J. Hoydis, K. Hosseini, S. ten Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 5–21, September 2013.

[68] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, February 2013.

[69] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, March 2014.

[70] N. Jamal and P. Mitran, "Performance analysis of null-steering beamformers in cognitive radio systems," in *2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, December 2011, pp. 1–6.

[71] ——, "Performance tradeoffs offered by beamforming in cognitive radio systems: An analytic approach," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3766–3777, October 2012.

[72] ——, "Throughput scaling of MIMO channels with imperfect CSIT in the low-SNR regime," *IEEE Communications Letters*, vol. 18, no. 9, pp. 1563–1566, September 2014.

[73] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 2112–2120, April 2009.

[74] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Courier Dover Publications, 1964.

[75] J. Poland, "Three different algorithms for generating uniformly distributed random points on the n-sphere," *Unpublished note available online at http://www-alg.ist.hokudai.ac.jp/∼jan/randsphere.pdf.*

[76] M. R. Spiegel, *Mathematical Handbook of Formulas and Tables.* McGraw-Hill, 1964.

[77] A. Papoulis and S. Pillai, *Probabilily, random variables and stochastic processes.* McGraw-Hill, 2002.

[78] L. Greenstein, V. Erceg, and D. Michelson, "Modelling diversity reception over narrowband fixed wireless channels," *Electronics Letters*, vol. 34, no. 11, pp. 1146–1147, May 1998.

[79] A. Lozano, A. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2527–2544, October 2003.

[80] E. Jorswieck and H. Boche, "Multiple-antenna capacity in the low-power regime: channel knowledge and correlation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05).*, vol. 3, March 2005.

[81] S. Jin, M. McKay, W. Kai-Kit, and L. Xiao, "Low-SNR capacity of multiple-antenna systems with statistical channel-state information," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 2874–2884, July 2010.

[82] C. Rao and B. Hassibi, "Analysis of multiple-antenna wireless links at low SNR," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 2123–2130, September 2004.

[83] W. Xinzhou and R. Srikant, "MIMO channels in the low-SNR regime: Communication rate, error exponent, and signal peakiness," *IEEE Transactions on Information Theory*, vol. 53, no. 4, pp. 1290–1309, April 2007.

[84] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[85] T. Cover and J. Thomas, *Elements of Information Theory*. New York:Wiley, 2006.

[86] D. Gu and C. Leung, "Performance analysis of transmit diversity scheme with imperfect channel estimation," *Electronics Letters*, vol. 39, no. 4, pp. 402–403, February 2003.

[87] O. El Ayach, S. Peters, and J. Heath, R.W., "The practical challenges of interference alignment," *IEEE Wireless Communications*, vol. 20, no. 1, pp. 35–42, February 2013.

[88] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[89] C. Hansen, "WiGiG: Multi-gigabit wireless communications in the 60 GHz band," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 6–7, December 2011.

[90] Z. Shen, R. Heath, J. Andrews, and B. Evans, "Comparison of space-time water-filling and spatial water-filling for MIMO fading channels," in *IEEE Global Telecommunications Conference*, vol. 1, November 2004, pp. 431–435.