ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author

PhD Dissertation

Randomized Space-Time Block Coding for the Multiple-Relay Channel

Author: David Gregoratti

Advisor: Dr. Xavier Mestre (CTTC)

Tutor: Dr. Miguel Ángel Lagunas Hernández (UPC, CTTC)

Departament de Teoria del Senyal i Comunicacions
Universitat Politècnica de Catalunya

June 22, 2010

This work was partially supported by the Catalan Government under grants 2006FI 00879, 2007FIC 00028, 2008FIC 00373, 2009FIC 00014, 2009 BE-1 00112 and 2009 SGR 1046.

A mê mari e a gno pari.

Abstract

In the last decade, cooperation among multiple terminals has been seen as one of the more promising strategies to improve transmission speed in wireless communications networks. Basically, the idea is to mimic an antenna array and apply distributed versions of well-known space-diversity techniques. In this context, the simplest cooperative scheme is the relay channel: all the terminals (relays) that overhear a point-to-point communication between a source and a destination may decide to aid the source by forwarding (relaying) its message.

In a mobile system, it is common to assume that the relays do not have any information about the channel between them and the destination. Under this hypothesis, the best solution to exploit the diversity offered by multiple transmitting antennas is to use space-time coding (STC). However, classical STC's are designed for systems with a fixed and usually low number of antennas. Thus, they are not suitable for relaying in most mobile communications systems where the number of terminals is potentially large and may vary as users join or leave the network. For each new configuration, a new code has to be chosen and notified to the relays, introducing a set-up overhead of signaling traffic.

In this dissertation we will propose and analyze a randomized distributed linear-dispersion space-time block code (LD-STBC): each relay is assigned a specific matrix which linearly transforms (left-multiplies) the column vector of source symbols. Each matrix is independently generated and does not depend on the total number of transmitters, which can thus change without interrupting data transmission for a new code–relay assignment.

The more intuitive way to build independent linear-dispersion matrices is to fill them with independent and identically distributed (i.i.d.) random variables. Therefore, we will first consider these i.i.d. codes and characterize the resulting spectral efficiency. In order to analyze the performance achieved by the system, we consider a large-system analysis based on random matrix theory. We will show that the random spectral efficiency (function of the random linear-dispersion matrices) converges almost surely to a deterministic quantity when the dimensions of the code grow indefinitely while keeping constant the coding rate. Since convergence is very fast, the random spectral efficiency will be approximated by the deterministic limit

in the subsequent analysis. By comparison with the direct link, sufficient conditions are derived for the superiority of relaying.

Next, we will analyze the outage probability of the system, that is the probability that the spectral efficiency falls below a given target rate due to channel fading. The main purpose of diversity techniques is to introduce alternative paths from the source to the destination, so that data transmission does not fail when the direct link undergoes deep fading. We will show that the diversity behavior of LD-STBC relaying mainly depends on both the coding rate and the relaying strategy (amplify and forward or decode and forward). It is then important to choose the coding rate that maximizes the diversity order without wasting too many resources.

To conclude the dissertation, we will consider a different code based on independent isometric Haar-distributed random linear-dispersion matrices. The new code maintains the flexibility of the previous one with respect to variations in the number of relays. However, the more complex structure of the codes allows a noticeable reduction of the interference generated by the relays. Unfortunately, isometric codes also require more sophisticated mathematical tools for their asymptotic analysis. For this reason, we simply introduce the problem by showing that it is possible to have some spectral-efficiency gain with respect to i.i.d. codes. The outage-probability analysis requires a more thorough understanding and will be the subject of future work.

Resumen

En la última década, la cooperación entre usuarios ha generado un gran interés por la posibilidad de mejorar la velocidad de transmisión en las redes de comunicaciones inalámbricas. El objetivo es formar un *array* con las antenas de todos los dispositivos y, de esta forma, aplicar técnicas de procesado espacio-temporal. El esquema de cooperación más sencillo es el canal con *relays*: todos los terminales que escuchen una comunicación entre dos puntos pueden ayudar a la fuente retransmitiendo lo que hayan recibido.

En un sistema realista, los *relays* no disponen de información sobre el canal en trasmisión. En este escenario, los códigos espacio-temporales (STC, del inglés *space-time coding*) son la alternativa más eficiente para aprovechar la diversidad introducida por los relays. Sin embargo, los STC clásicos están diseñados para un número limitado y fijo de antenas transmisoras y no se adaptan bien a sistemas cooperativos donde el número de *relays* puede ser elevado y, sobretodo, puede variar en el tiempo, según los usuarios entren o salgan de la red. El problema principal es la necesidad de usar un código nuevo cada vez que cambie la configuración de la red, generando un importante tráfico de señalización.

Esta tesis analiza un código espacio-temporal a bloques de dispersión lineal (LD-STBC, del inglés *linear-dispersion space-time block coding*), aleatorio y distribuido: a cada *relay* se le asigna una matriz aleatoria que aplica una transformación lineal al vector que contiene los símbolos de la fuente. Cada matriz se genera de forma independiente y sin ninguna relación con el número de usuarios involucrados. De esta manera, el número de nodos puede variar sin necesidad de modificar los códigos existentes.

La forma más intuitiva de construir matrices de dispersión lineal independientes es que sus elementos sean variables aleatorias independientes e idénticamente distribuidas (i.i.d.). Por esta razón, se estudia primero la eficiencia espectral obtenida por este tipo de LD-STBC. Es importante remarcar que la eficiencia espectral es una cantidad aleatoria, ya que es una función de los códigos aleatorios anteriormente descritos. Sin embargo, cuando las dimensiones de las matrices crecen infinitamente pero manteniendo constante la tasa del código (relación entre número de símbolos de la fuente sobre el número de símbolos de los *relays*), la eficiencia espectral converge rápidamente hacia una cantidad determinista. Este resultado se demues-

tra usando la teoría de las matrices aleatorias. Por esta razón, el sistema se analiza aproximando la eficiencia espectral con su limite. Por ejemplo, la comparación con el canal directo entre fuente y destino permite definir unas condiciones suficientes en donde el sistema con *relays* es superior a la comunicación punto a punto.

Posteriormente se debe analizar la probabilidad de *outage*, es decir la probabilidad de que, debido a la baja calidad del canal, la eficiencia espectral sea menor que la velocidad de transmisión solicitada por el sistema. Como ya se ha mencionado anteriormente, los *relays* se introducen para aumentar la diversidad del canal y, con ella, el número de caminos independientes entre la fuente y el receptor, reduciendo la probabilidad de *outage*. Para los LD-STBC i.i.d. las prestaciones en términos de *outage* dependen del tipo de *relay* (*amplify and forward* o *decode and forward*) y son función de la tasa del código, que debe ser cuidadosamente elegida para maximizar el orden de diversidad sin desperdiciar demasiados recursos.

Finalmente, en el último capítulo de la tesis se considera otro tipo de LD-STBC, distinto del i.i.d. analizado hasta ahora. En este caso, las matrices de dispersión lineal siguen siendo independientes la una de la otra pero se añade la restricción de que cada una tenga columnas (o filas, según la tasa del código) ortogonales. Así, se consigue que el código siga siendo flexible con respecto a las variaciones en el número de usuarios, pero su estructura permite reducir la interferencia generada por cada *relay*, como se puede notar comparando su eficiencia espectral con la eficiencia espectral obtenida por el código i.i.d. Cabe destacar que el análisis asintótico de estos códigos (llamados isométricos) se basa en herramientas matemáticas más sofisticadas que las anteriores y, por lo tanto, es necesario un estudio más profundo para poder entender cómo se comporta en términos de *outage*.

Struc

L'interès par lis comunicazions cooperativis al è une vore cressût intai ultins dîs agns. La idee e je chê di meti in comun lis risorsis dai utents par miorâ la cualitât dai leams e, duncje, la velocitât di trasmission. Par esempli, cu lis antenis dai terminâi si puedin formâ des schiriis e doprâ cussì lis tecnichis di tratament spazi-temporâl dal segnâl. Il scheme plui sempliç di cooperazion al è il canâl cun ripetidôrs (*relay* in inglês), ven a stâi un sisteme li che ducj i utents che a ricevin un messaç che nol è par lôr lu tornin a trasmeti viers la sô destinazion, judant il passaç de informazion.

Cheste tesi e studie un pussibil esempli des operazions che si àn di fâ intai ripetidôrs par impedî che si fasedin interference tra di lôr. La soluzion proponude e je particolarmentri interessante cuant che il numar di utents al cambie intal timp, stant che si doprin dai codiçs spazi-temporâi che no dipindin dal numar di trasmetidôrs. Ancje se il sisteme al è une vore sempliç, parcè che si base su des trasformazions lineârs, il ricevidôr al pues ricognossi il contribût di ogni ripetidôr cence confusion. In cheste maniere, al baste che un ripetidôr (e no ducj) al formedi un leam di buine cualitât tra la sorzint e la destinazion par che la informazion e sedi ricevude coretementri.

Acknowledgments

Why can't we express how thankful we are by a mathematical equation? Direct, concise, efficient and with no risk of sounding tedious or rhetoric. Unfortunately, no such theory has been developed yet, so I'll try to do my best with simple words.

There's no doubt I have to start with CTTC, as a whole. And not only for funding my Ph.D., but also for smoothing the path through this four-year experience and for the friendly working atmosphere. *Visca el CTTC!*

Xavi, it's your turn. Half of the results that follow won't be there without your hints, your patience and your time. I've really learned a lot from you and hope I'll still have the chance of working with you in the future. Thanks!

Another great contribution to this dissertation, in particular to the last chapter, comes from Walid. I sincerely enjoyed my stay at Télécom ParisTech, your interest in my work, your technical explanations and our collaboration. *Merci !*

Next, I'd like to thank all the members of the defense evaluation board to accept the invitation. Of course, I must also mention Aitor, Christian and Mari, who probably are the quickest reviewers in the world... even if there is no reward dinner for them. I owe you a beer, at least.

Luckily, getting a Ph.D. doesn't only mean publications and writing the dissertation. During these four years I had the opportunity to meet some good friends, who have been my personal Spanish teachers, too. Empezaré por Javi-bi el sabadellense, Jesús o, mejor dicho, Jesus y su pesimismo ("David, esto no te va a funcionar. Ya lo sabes, ¿verdad?"), Ana María (Miss Hiperactividad) y Dani: sin los innumerables cafés, esta tesis se habría acabado un año antes, pero ¡qué aburrimiento! Aitor, Marc, Fermín y toda la panda de las noches pasadas entre "El Cangrejo" y "El Marsella". Jalonso, NZ: mis primeros compañeros de despacho. Ricardo y los demás miembros del club del tupper. Miguel Ángel, con que he probado todos los restaurantes de Castelldefels durante los findes pasados escribiendo la tesis. Paolo, que fundó conmigo la colonia italiana del CTTC: *desafortunadamente ;)* se nos ha descontrolado. ¡Un abrazo a todos!

Para los que me han pillado por los pasillos cantando temas de *Lola Flores* o de *Los Romeos*, que sepan que la culpa es de mis compañeras de piso, Amaia y Su, que cuando arrancan el *chinguestar* no hay manera de

callarlas. A pesar de eso, aguantaros no es tan duro. Y con vosotras, a la horda atroz y espeluznante que a menudo os acompaña.

Los latinos decían *"mens sana in corpore sano"*: pues yo lo he intentado y conmigo los chicos y chicas del equipo de natación máster del Medi. Para olvidarme de integrales y convergencias, nada me ha ido mejor que compartir con vosotros uno de los entrenos rompe-huesos de nuestro coach Ángel ("10 de 100 cada 1'45", ¡¡y no digas que no puedes!!"). ¡Nos vemos en la pecera!

Cheste volte no pues dismenteâmi da la çusse di me sûr, Ambra. O sin fradis e si volìn ben, però nissun dai doi si strussie masse par dimostrâlu... no impuarte, prime o dopo o cambiarìn.

O ai lassât par ultins i mie viei, Gigi e Maria. O volarès dîus un grum di robis, parcè che di pinsîrs us ind ai dâts avonde intai ultins agns. Par disfortune, chest mus al fâs fadie a tabaiâ, aromai o cognossês la bestie e il so cjavat dûr. E je almancul une robe che o vês di savê: ancje se a son dibot trente agns che o studii e mi pâr che o larai indenant cussì ancjemò un biel toc, o varai simpri alc di imparâ di voaltris doi, i miôr mestris che o vedi mai vût.

David
May 2010

Contents

List of Figures

Chapter 1

Introduction and Background

Mobile communications systems are widespread in day-to-day life: probably the most impressive example is the number of mobile cellular phones, which has increased by an order of magnitude in the last decade (from 490 million subscribers in 1999 to 4,100 million subscribers in 2008[1]). Another example is the wireless "Wi-Fi" access to the Internet, offered by an increasing number of private enterprises and public administrations. For both users and operators, it is hence of great interest to implement communications services typical of wired systems like, for instance, TV on demand, voice over IP or file sharing. Thus, the resulting demand for transmission speed and reliability still motivates lots of interest in the research community.

In all wireless communication systems, the main issues concern the fluctuations of the channel. Indeed, different physical agents influence the quality of the wireless link. In the next few pages, these problems are briefly presented together with some possible solutions, with particular interest in mobile systems. Then, we will see how this dissertation fits into this subject.

1.1 Multipath fading and diversity

Typical mobile communications users are not aware of the position of the receiver (base stations and access points in the mentioned examples). Hence, terminals are usually equipped with non-directive antennas and radio waves are scattered over the surrounding space. As a result, signals may be received from multiple paths, due to reflecting objects in the environment. In a mobile system, phases and delays of the different replicas vary continuously, due to changes in the relative position of transmitters, receivers and reflecting objects. Therefore, the quality of the equivalent channel is

[1] Source: International Telecommunication Union
<http://www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom99.html>.

a function of both time and frequency. This phenomenon is referred to as *small-scale fading* or simply *fading*[2].

Diversity is a simple solution to this issue [1–3]. Multiple copies of the same message are sent at different time instants (*time diversity*) or on different frequency bands (*frequency diversity*). According to the previous comments, each replica undergoes a different equivalent channel. If the replicas are sufficiently spaced (in time or frequency, respectively), the respective channels can be modeled as independent random variables. Intuitively, as the number of replicas increases, we reduce the probability that no channel is good enough to convey information from the source to the receiver. The price to pay is an increment in the total transmission delay or bandwidth.

More recently, *space diversity* has offered a different approach to the problem. Indeed, multipath fading is now seen as a resource to exploit as opposed to a penalizing characteristic of wireless channels. Briefly, by considering the signal paths separately, it may happen that some of them can support the transmission rate required by the system. In this perspective, whenever there are no multiple paths (or, better, when they cannot be resolved), they can be artificially generated by placing multiples antennas at the transmitter and/or at the receiver. The resulting channel is commonly called the multiple-input-multiple-output (MIMO) channel.

1.1.1 From space diversity to cooperative diversity

The idea of transmitting information over multiple channels to improve the total link quality has been known for quite a long time: the first works by A. Kaye and D. George [4] and W. van Etten [5, 6] date back to the seventies. It is in the last twenty years, however, that MIMO channels have become very popular and found applications in a variety of communications systems. Landmark contributions by S. M. Alamouti [7], V. Tarokh *et al.* [8, 9] and İ. E. Telatar [10] showed that space-diversity techniques are diversity achieving and, at the same time, offer interesting multiplexing properties.

Note that space diversity introduces less delay and spectrum dilation than time- and frequency-diversity techniques. Furthermore, it can be used in conjunction with them, if needed. To generate independent uncorrelated resolvable paths, however, antennas should be placed sufficiently far to one another. More specifically, the minimum relative distance is in the order of half the wavelength (e.g. around ten centimeters for systems with a carrier frequency of 2 GHz such as UMTS). This fact limits the applicability of multiple antennas when dealing with reduced-dimension portable devices. Nevertheless, continuous requests for higher data rates (as mentioned before)

[2]Other phenomena influence the quality of the wireless channel, such as the *path loss* and the *large-scale fading* (or *shadowing*). However, they are usually considered in power-allocation problems and are not important for the topic of this thesis. The interested reader can refer, e.g., to [1, Chapter 2].

push for the introduction of space-diversity techniques in mobile communications systems.

The underlying idea behind *cooperative diversity* is that (possibly idle) terminals in the system overhear other users' communication and may hence relay this information to the receiver. In other words, these terminals mimic a virtual antenna array and distributed versions of space-diversity transmission techniques may be applied. A brief overview of *cooperative communications* and the relay channel is given in the next section. A more detailed introduction can be found in, e.g., [11].

1.2 Cooperative communications and the relay channel

The design of a cooperative communications system involves problems at different communications levels. For instance, one should define (i) the signal processing at the relays in order to achieve diversity and maximize the spectral efficiency (physical layer), (ii) how to share resources between source terminals and relaying ones (MAC layer) or even (iii) how to route messages when multiple relaying hops are allowed (a routing problem).

The following overview reports some important results regarding the physical layer. Even though the other aspects mentioned above offer interesting research lines, they fall outside the topic of this dissertation and are not treated here.

1.2.1 Some history

The first publications on the relay channel are probably those due to E. C. van der Meulen. In his Ph.D. dissertation [12] (1968) and the following article [13] (1971), van der Meulen describes a three-terminal model (a source, a destination and a relay) which is later deeply analyzed by T. M. Cover and A. El Gamal in 1979. In their landmark article [14], they derive upper- and lower-bounds for the capacity for the general single-relay channel and give an exact expression for the Gaussian degraded case. Unfortunately, there are no straightforward applications for this latter result: roughly speaking, it implies that the destination has some knowledge about the relay input.

At this point, the topic showed to be complex and lacked of practical applications. Consequently, there were no striking contributions in the following twenty years. Then, between the years 2002 and 2003, the subject gained new interest after some new works [15–18] suggested that relay networks might have benefited from the powerful results on MIMO channels mentioned in Section 1.1.1. Since then, countless contributions have been published on the topic, proposing different relaying strategies and approaches to the problem.

1.2.2 Half-duplex vs. full-duplex

In most practical applications, relays are half-duplex terminals, i.e. they receive and transmit signals at different times or on different frequency bands. Full-duplex relays, capable of transmitting and receiving on the same time–frequency channel, also exist. Lots of theoretical works (e.g. [14, 17–23]) introduce full-duplex relays to characterize capacity bounds. Furthermore, intuition suggests that they better exploit the system degrees of freedom. However, the interference-cancellation techniques they imply are computationally expensive, extremely sensitive to errors and, thus, less suitable for practical systems.

1.2.3 Forwarding techniques

In the previous sections we have seen that relays overhear the information broadcasted by the source and, then, forward it to the destination. According to how the information is processed at the relay before being retransmitted, we can identify different forwarding strategies. The three most significant ones probably are:

Amplify and Forward (AF). In its original meaning, an AF relay simply applies a power gain to the received noisy message. More generally, many authors extend the definition to any relay that linearly transforms the incoming signal (e.g., [24–26]). AF relays are very simple devices but the total noise at the destination may increase considerably;

Decode and Forward (DF). DF relays try to decode the source information and forward it only if they are successful. In this way, the noise at the receiver is much lower than in the AF case. However the terminals are more complex and power consuming;

Compress and Forward (CF). This is the more general strategy, since it includes any non-linear relaying function. Sometimes, to identify more exactly this function, names like "estimate and forward", "quantize and forward" or "observe and forward" are used. Note that we can even think to systems where the compression level varies according to the channel qualities: it is intuitive that the receiver needs less information from the relays when the direct link from the source is good enough [27, 28].

1.2.4 Channel model, channel state information and channel access

Source–destination, source–relay and relay–destination channels may be modeled in many different ways, according to the underlying application. Let

us point out here that the technical chapters of this dissertation assume frequency flat, slow (quasi-static) fading. This is a common assumption in the literature (see, e.g., [18–20,24,26,29–33], just to cite a few). Mentioning briefly other possible models, the additive-white-Gaussian-noise (AWGN) channel is sometimes assumed to simplify the analysis [25, 34] or to model situations where two terminals are very close to one another [23]. Cover and El Gamal's results on the degraded channel [14] have already been cited in Section 1.2.1. A generalization to the multiple-relay case can be found in [35].

Hypothesis on the channel state information (CSI) available at each terminal are other important differences among relaying schemes. In order to minimize complexity and power consumption, we can think about very simple AF relays without any CSI [24, 30, 34]. This assumption usually implies that all the channels are known at the receiver.

When some more complexity is allowed, relays are assumed to have information about the channel from the source. For example, DF relays usually need this information to decode source messages.

CSI about the relay–destination channel, instead, is important when designing the relay transmission strategy, especially when considering multiple-relay schemes. Based on MIMO experience, when CSI is available at the transmitter side, the optimal solution is beamforming [2, Chapters 5 and 7]. Indeed, by performing a singular value decomposition of the channel matrix, the transmitted power can be concentrated into the main channel mode(s). Distributed versions of beamforming are proposed, e.g., in [25, 30].

Note that transmitter-side CSI requires some feedback from the receiver. Furthermore, beamforming techniques require a distributed knowledge of the channels. In other words, relays should have information not only about their own channels, but also about the channels seen by the other relays. Consequently, a lot of works assume no transmitter-side CSI at the relays and employ space-time coding (STC) [29,32,36,37]. More details about this solution are given in Section 1.3.

To conclude, we simply mention here that other channel-access solutions, not explicitly designed for MIMO channels, appear in the literature. Some examples are time division multiple access (TDMA) [32], code division multiple access (CDMA) [17,18,38] and time-variant relay-specific phase rotations [39].

1.2.5 Capacity and outage probability

In spite of the great amount of publications on the subject, the general expression for the relay channel capacity has not been found yet. There exist, however, capacity expressions for some special cases as, e.g., the Gaussian degraded relay channel (both single-relay [14] and multiple-relay [35]) and the single-relay case with orthogonal components [40].

For more general cases, numerous bounds are available. Relay capacity with AWGN channels and relay ergodic capacity under the Rayleigh-fading assumption are considered and bounded in [41], for single-antenna terminals, and [20], for multiple-antenna ones.

Other papers (see, e.g., [21, 30, 42–45]) consider the global capacity of systems where the number of relays tends to infinity. For instance, M. Gastpar and M. Vetterli show in [21, 42] that the capacity of a Gaussian relay network increases as the logarithm of the number of terminals when the latter grows large.

Instead of dealing with ergodic capacity, other publications [29, 32, 45] focus on outage probability, which is the probability that a target rate cannot be supported by the system because of fading. This performance metric is simpler to analyze (at least in principle). More specifically, when assuming high transmitted power, the study of the outage probability allows us to identify the diversity order of the system and how it is obtained at the expense of multiplexing capabilities. This is the so called diversity–multiplexing tradeoff introduced by L. Zheng and D. N. C. Tse in [46] and applied to the relay channel in, e.g., [26, 32]. More precise definitions of outage probability, diversity order and diversity–multiplexing tradeoff are given in the introduction to Chapter 4.

1.3 Motivation

Being interested in a reasonably feasible communications system, this dissertation deals with a basic multiple-relay channel. A point-to-point communication between a source and a destination is aided by a set of L relays. All terminals are half-duplex and are equipped with a single antenna. The channel state information is assumed to be concentrated at the destination. On the contrary, relays are intended to be low-complexity devices and, hence, do not have any knowledge about their channel to the destination. (When considering the DF relaying strategy, Chapters 4 and 5, some information about the source–relay channels is needed at the relays. This is not the case with the AF relaying strategy, Chapters 3 and 4.)

As mentioned in Section 1.2.4, space-time coding is the best solution when trying to exploit the degrees of freedom offered by a multiple-input channel with no CSI at the transmitter side. In general terms, each antenna transmits a different message containing all the information. By properly designing the messages, the antennas can transmit on the same time–frequency channel without interfering with each other [7–9, 47, 48]. This approach is much more efficient than repeating the same message by one antenna at a time in a TDMA fashion.

In the attempt to obtain similar benefits from a set of relays mimicking an antenna array, an important number of works propose distributed versions

of space-time codes (see, e.g., [49–54]). Classical space-time codes, however, are not suitable for the multiple-relay channel because of different reasons.

First, to achieve diversity, the code must be jointly designed for all transmitting antennas. This means that the assignment of each code section to the corresponding relay must be done by a central entity, introducing a set-up overhead of signaling traffic.

Second, codes are generally designed for a specific number of transmitters. In dynamic networks, this fact implies that a new code–relay assignment must be carried out whenever one or more terminals drop in and out of the system, thus generating more overhead.

Finally, the design complexity of classical space-time codes increases with the number of transmitters, as opposed to their efficiency (for example, space-time codes from orthogonal designs [8] have rate 1 with two transmitters, rate 3/4 with three or four transmitters and rate 1/2 with five or more transmitters). Once again, this is a limitation for typical mobile communications systems (just think to the number of mobile phones or devices with a wireless-LAN interface that we can count in most public places).

Summarizing, an ideal space-time code for relay networks should work with any number of transmitters. Moreover, it should be distributed and dynamic, in the sense that each relay should be assigned a section of the code which is independent of the other relays and does not change with the number of terminals in the system.

1.3.1 Previous work

The STC flexibility problem presented above has long been known. In [32], J. N. Laneman and G. W. Wornell suggest employing space-time codes from orthogonal designs as a possible solution. These space-time block codes, proposed by V. Tarokh *et al.* [8], are designed for a given number L of transmitters but maintain their orthogonal properties when some of the antennas are shut down. This implies that the maximum number of relays in the system must be known a priori. Moreover, for more than four transmitters, the coding rate is only 1/2, thus limiting the spectral efficiency.

The solution proposed in [37] is based on linear-dispersion space-time block coding (LD-STBC): each relay is assigned a specific unitary matrix which produces a linear transformation of the vector of source symbols. The system is quite flexible, since no particular relation is assumed among the different coding matrices: when a new terminal joins the network, a new matrix is generated without modifying the existent ones. In this work, however, no direct link between the transmission source and its destination is considered. Furthermore, the choice of unitary matrices constrains the coding rate (equal to the number of columns divided by the number of rows) to one. As shown in [55, 56], this is not always the best choice for half-duplex relays: it may be enough for the relays to send a compressed

version of the message (i.e. coding rate larger than one), thus reducing the time where the source remains silent.

A completely different approach appears in [36]. From the source message, each relay generates a new vector of symbols by doing a random linear combination of the columns of a matrix codeword, which is obtained from a common deterministic space-time mapping. It turns out that the system performance is limited by the minimum between the number of relays and the number of transmitters of the underlying deterministic STC.

1.3.2 The proposed scheme

The solution proposed in this dissertation employs distributed randomized LD-STBC to exploit the intrinsic diversity of the multiple-relay channel. More specifically, each relay is assigned an independent random matrix which linearly transforms (left-multiplies) the column vector of source symbols. These linear-dispersion matrices are either filled with independent and identically distributed (i.i.d.) entries (see Chapters 3 and 4) or generated according to an isometric bi-unitarily invariant distribution (see Chapter 5). Note neither distribution depends on the number of generated matrices (that is the number of relays): in this way we guarantee a flexible coding scheme according to which the active terminals are not affected by those that enter or leave the network. Furthermore, at least in the i.i.d. case, direct analogies with direct-sequence CDMA (DS/CDMA) systems [57] suggest that the coding scheme is robust to little relative delays between any two relay messages (see Chapter 5).

All the channels are assumed to undergo frequency-flat quasi-static fading (i.e. the channel gains do not vary during the transmission of a message). The CSI is concentrated at the destination: the receiver knows the gains of all source–destination, source–relay and relay–destination channels, together with the coding matrices assigned to the relays. Conversely, each relay knows its specific matrix only. Under these hypothesis, the system performances are analyzed in terms of spectral efficiency and outage probability for two different receivers, namely the optimum maximum-likelihood receiver and the linear minimum-mean-square-error (LMMSE) receiver.

Note that no assumptions have been made on the dimensions of the linear-dispersion matrices. By varying their ratio (i.e. the coding rate, ratio between the number of source symbols and the number of relay symbols), different weights can be assigned to source and relay contributions to the message estimation at the destination. In many situations, a coding rate larger than one (relays compress the information) is sufficient to transfer information properly and, meanwhile, exploits system resources in a more efficient manner.

To conclude, let as mention that the analysis of this coding scheme is also very appealing from a mathematical point of view. Indeed, observe that the

spectral efficiency we derive is a function of the random linear-dispersion matrices and, thus, is itself a random quantity. However, by means of random matrix theory and free-probability theory, we will prove that the random spectral efficiency tends to a deterministic quantity when the dimensions of the matrices grow indefinitely but at the same rate, meaning that the coding rate is kept constant. Furthermore, the resulting limit is an excellent approximation of systems with finite-dimensional codes, even for not-so-large linear-dispersion matrices.

1.4 Thesis outline

This section gives a brief overview of the contents of the following chapters.

Chapter 2 In this chapter we review some basic results about the convergence of sequences of random variables, about random matrix theory and about free-probability theory. The proofs are omitted whenever they are too long or complicated for the introductory level of this summary. All the details can be found in the referenced literature.

Most of these mathematical tools are employed to derive the technical results of this dissertation. Others, though, are reported with the mere intention of giving a self-contained aspect to the chapter.

Chapter 3 Here we define the details of the reference system. More specifically, we focus on the AF relaying strategy, which is more general, and on i.i.d. LD-STBC, since this is probably the first solution one can think of when looking for independent linear-dispersion matrices. The resulting spectral efficiency is shown to converge to a deterministic quantity when both matrix dimensions grow large while keeping constant the coding rate. This limit is used to derive conditions for the superiority of relaying over the direct link.

The work of this chapter has been published in the following papers:

- D. Gregoratti and X. Mestre, "Random DS/CDMA for the amplify and forward relay channel," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 1017–1027, Feb. 2009;

- D. Gregoratti and X. Mestre, "Asymptotic spectral efficiency analysis of the DS/CDMA amplify and forward relay channel," in *Proc. Asilomar Conference on Signals, Systems, and Computers 2007*, Pacific Grove, CA, USA, Nov. 4–7 2007 (invited paper);

- D. Gregoratti and X. Mestre, "About asymptotic spectral efficiency in the DS/CDMA amplify and forward relay channel," in *Proc. IEEE ISWPC 2008*, Santorini, Greece, May 7–9 2008 (invited paper);

- D. Gregoratti and X. Mestre, "The DS/CDMA amplify and forward relay channel: Asymptotic spectral efficiency," in *Proc. ICT-MobileSummit 2008*, Stockholm, Sweden, Jun. 10–12 2008.

Other work related to the topic appears in

- D. Gregoratti and X. Mestre, "On the low source power regime of the DS/CDMA relay channel," in *Proc. IEEE SPAWC 2008*, Recife, Brazil, Jul. 6–9 2008.

Chapter 4 The analysis of i.i.d. LD-STBC is extended to the characterization of the outage probability in the high-SNR regime. As a first approach to the problem, we introduce the DF relaying strategy, which is simpler since relays do not forward any noise. Based on the asymptotic spectral efficiency, the diversity order, the outage gain and the diversity–multiplexing tradeoff are computed and compared to existing relaying strategies. Similar results are then derived for the AF relaying scheme.

The following articles deal with the contents of this chapter:

- D. Gregoratti and X. Mestre, "Large-SNR Outage Analysis for the DF Relay Channel with Randomized Space-Time Block Coding," submitted to *IEEE Trans. Wireless Commun.*;

- D. Gregoratti and X. Mestre, "The single relay channel: Does randomized coding increase diversity?" in *Proc. ICT-MobileSummit 2009*, Santander, Spain, Jun. 10–12 2009;

- D. Gregoratti and X. Mestre, "Diversity analysis of a randomized distributed space-time coding in an amplify and forward relay channel," in *Proc. IEEE ICC 2009*, Dresden, Germany, Jun. 14–18 2009;

- D. Gregoratti and X. Mestre, "Decode and forward relays: Full diversity with randomized distributed space-time coding," in *Proc. IEEE ISIT 2009*, Seoul, Korea, Jun. 28–Jul. 3 2009;

- D. Gregoratti and X. Mestre, "Diversity order for the amplify-and-forward multiple-relay channel with randomized distributed space-time coding," in *Proc. EUSIPCO 2009*, Glasgow, Scotland, Aug. 24–28 2009.

Chapter 5 In this chapter we investigate whether some gain could be introduced by adding some structure to the LD-STBC. More specifically, the i.i.d. linear-dispersion matrices are replaced by isometric linear-dispersion matrices, that is each random matrix has orthogonal columns (or rows, according to the coding rate). Intuitively, by analogy with DS/CDMA systems, this choice reduces interference. Since the asymptotic analysis is quite

more complex than before, we limit our interest to the spectral efficiency of the DF relaying strategy.

These results have been published in:

- D. Gregoratti, W. Hachem, and X. Mestre, "Randomized isometric linear-dispersion space-time block coding for the DF relay channel," submitted to *IEEE Trans. Signal Process.*
- D. Gregoratti, W. Hachem, and X. Mestre, "Orthogonal matrix precoding for relay networks," in *in Proc. IEEE ISWPC 2010*, Modena, Italy, May 5–7 2010 (invited paper);

Chapter 6 This chapter offers a brief summary of the results of this dissertation and outlines some of the research lines that arise as a direct extension of this work.

Chapter 2

Mathematical Background

This chapter is a brief review of the main mathematical tools that are employed to prove technical results later in this thesis. First, for a sequence of random variables, the different types of convergence are defined and discussed. Second, we introduce random matrix theory, a branch of multivariate statistics that deals with the characterization of the eigenvalues of random matrices whose dimensions grow without bound but at the same rate. The main concepts are clarified with some basic and simple examples before dealing with the Stieltjes transform, which allows to address more complex problems. Finally, the last section is devoted to free probability: under some assumptions, indeed, large random matrices can be treated as free non-commutative random variables.

2.1 Convergence of sequences of random variables

Sequences of random variables $\{x_n\} = \{x_1, x_2, \dots\}$ appear quite often in statistical models: the index can refer, for instance, to the sample size or to the number of experiments. In this contest, the concept of convergence for deterministic sequences can be extended to random variables, as thoroughly described in the literature. The interested reader should refer to any book on probability theory, e.g. [58, 59], for a detailed exposition of the subject. Here, as an introduction to the following sections, we will only recall the definitions and properties of the four main modes of convergence for sequences of random variables.

Let x and the sequence $\{x_n\} = \{x_1, x_2, \dots\}$ be real- or complex-valued random variables defined on a common probability space. Furthermore, we denote with $\Pr[\mathcal{A}]$ the probability of the event $\mathcal{A}$ and with $\mathbb{E}[\cdot]$ the expected value of any random quantity. Then, one can define the convergence of the sequence $\{x_n\}$ towards the random variable x according to one of the four following modes:

Almost sure convergence (also convergence almost everywhere or con-

vergence with probability 1). The sequence $\{x_n\}$ is said to converge almost surely to x, and we write $x_n \xrightarrow{\text{a.s.}} x$, if

$$\Pr\Big[\lim_{n\to+\infty} x_n = x\Big] = 1.$$

Convergence in the r-th mean. The sequence $\{x_n\}$ is said to converge in the r-th mean to x if

$$\lim_{n\to+\infty} \mathbb{E}\Big[|x_n - x|^r\Big] = 0.$$

We just mention here that convergence in the r-th mean implies convergence in the s-th mean, for any $s < r$. The notation $x_n \xrightarrow{\text{m.s.}} x$ is common in the literature and denotes convergence in mean square, i.e. for $r = 2$.

Convergence in probability. The sequence $\{x_n\}$ is said to converge in probability to x, and we write $x_n \xrightarrow{\text{P}} x$, if

$$\lim_{n\to+\infty} \Pr[|x_n - x| < \epsilon] = 1, \qquad \forall \epsilon > 0.$$

Convergence in distribution (also convergence in law). Let $F(t)$ and $\{F_n(t)\}$ denote the probability distribution functions of the random variables x and $\{x_n\}$[1], respectively. The sequence $\{x_n\}$ is said to converge in distribution to x, and we write $x_n \xrightarrow{\text{D}} x$, if

$$\lim_{n\to+\infty} F_n(t) = F(t)$$

for any continuity point t of $F(t)$.

There exist some important relations among the convergence modes. In particular, both almost sure and r-th mean convergence imply convergence in probability which, in turn, implies convergence in distribution. Conversely, there are no general implications in the opposite sense or between almost sure and r-th mean convergence. Figure 2.1 depicts these relations, which also define somehow the strength of the convergence modes. Usually, stronger modes are also harder to prove. However, weak types of convergence can offer important insights, as in the following example.

Example 2.1. Assume that we are given the sequence of random variables $\{x_n\}$, and the relative distributions $\{F_n\}$, and we want to check whether it converges in probability. The first thing to do is certainly to compute the pointwise limit $F(t) = \lim_{n\to+\infty} F_n(t)$. Indeed, if it exist a random variable x such that $x_n \xrightarrow{\text{P}} x$, its probability distribution function must be F.

[1]For real-valued random variables, the probability distribution function coincides with the cumulative distribution function, i.e. $F(t) = \Pr[x \leq t]$, for any real t.

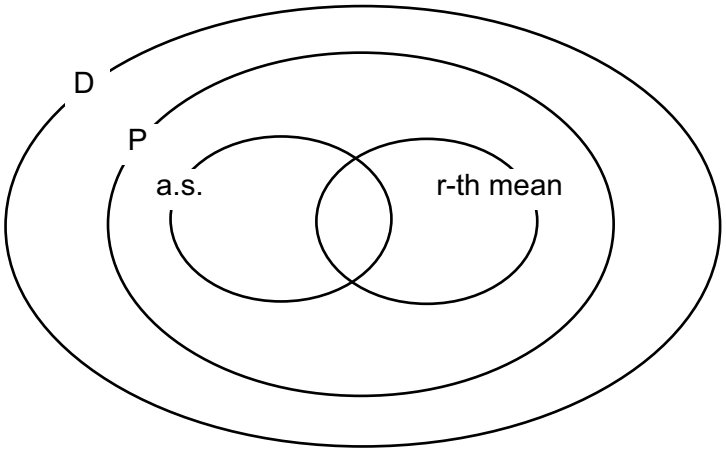


Figure 2.1: Relations among the main four convergence modes.

The previous concepts can be readily extended to the multivariate case. Specifically, we say that the sequence of random vectors $\{\mathbf{x}_n\}$ converges almost surely to the random vector $\mathbf{x}$ if the same property holds true between corresponding entries of the vectors. The same can be said for convergence in the r-th mean and in probability (in these cases, however, we can equivalently adapt the original definition by replacing the distance in $\mathbb{R}$, or $\mathbb{C}$, by the distance in $\mathbb{R}^N$, or $\mathbb{C}^N$) and for convergence in distribution (where, equivalently, the single-variate probability distribution function can be replaced by the multivariate one).

To conclude this overview, we recall a property that will be largely employed in the following chapters and which is about convergence preservation after continuous mapping. Specifically, assume that $\{\mathbf{x}_n\}$ and $\mathbf{x}$ are random vectors with values in $\mathbb{R}^N$ ($\mathbb{C}^N$) and that $\mathbf{x}_n \xrightarrow{\text{a.s.}} \mathbf{x}$. Then, for any continuous mapping $\mathbf{g}(\cdot) : \mathbb{R}^N \to \mathbb{R}^M$ ($\mathbb{C}^N \to \mathbb{C}^M$), the sequence of random vectors $\{\mathbf{g}(\mathbf{x}_n)\}$ converges to $\mathbf{g}(\mathbf{x})$ almost surely[2]. The same thing can be stated also for convergence in probability and in distribution, but not for convergence in the r-th mean.

2.2 Random matrix theory

As mentioned in the introduction, random matrix theory (RMT) is a set of powerful tools that allows to characterize the spectrum (i.e. the eigenvalues) of large matrices with random entries, when both dimensions grow at the same rate. In the following overview, which is not intended to be exhaustive, the very basic concepts are recalled, focusing on the tools that will be applied

[2]More generally, the random variables $\{x_n\}$ and x can take values on any metric space S and $\mathbf{g}(\cdot) : S \to S'$, with S' another, possibly different, metric space.

in the analysis of our transmission problem. We start with an example which shows the importance of being able to deal with matrices whose dimensions are both large.

Example 2.2. Consider the uplink of a direct-sequence code division multiple access (DS/CDMA) system and assume that K users are received with the same power P. The received signal can be modeled as

$$\mathbf{y} = \sqrt{\frac{P}{N}} \sum_{k=1}^{K} x_k \mathbf{s}_k + \mathbf{n}$$

where x_k ($\mathbb{E}[x_k] = 0$ and $\mathbb{E}[x_k x_i^*] = \delta_{i,k}$)[3] is the symbol of user k and $\frac{1}{\sqrt{N}}\mathbf{s}_k$ ($\mathbb{E}[\mathbf{s}_k] = \mathbf{0}$ and $\mathbb{E}[\mathbf{s}_k \mathbf{s}_i^H] = \delta_{i,k}\mathbf{I}_N$) its randomly generated spreading sequence with spreading factor N. The vector $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ is the normalized white Gaussian noise. The random model for the spreading sequences is quite common in the literature, either to represent practical pseudo-noise signatures (as, e.g., in IS-95 [60]) or to take into account the loss of orthogonality due to asynchronous transmissions.

Since the spreading sequences are neither orthogonal nor deterministic, the performance analysis of the system is not an easy task. To simplify the problem, it can be tempting to let either one of the two parameters K and N grow very large and apply the law of large numbers. The resulting white problem, however, loses a lot of insight. Indeed, besides disregarding the interference structure, we are either over-loading the system (case $K \to +\infty$) or wasting most of the degrees of freedom (case $N \to +\infty$). Furthermore, considering number of users and spreading factor of the same order of magnitude makes much more sense in a practical perspective. In the following sections, we will see how to get an excellent approximation of the finite reality by letting both K and N tend to infinity at the same rate, i.e. keeping constant their ratio K/N.

This example in the CDMA framework is just one of the numerous applications of RMT in communications problems. Basically, these mathematical tools can be employed each time the behavior of the system is intrinsically dependent on the ratio between two quantities that are allowed to grow indefinitely. Two other classical examples are (i) capacity computation for MIMO channels with a large, but similar, number of transmitter and receiver antennas and (ii) evaluation of parameter estimators, where the sample size is of the same order of magnitude of the number of estimated parameters.

A detailed overview of these communications problems (at least, the CDMA and MIMO ones) and their RMT solutions can be found in the monograph [61]. For a more mathematical approach, the interested reader can refer to the tutorials in [62,63] or to the book by V. L. Girko [64]. As for

[3] $\delta_{i,k}$ is the Kronecker delta, i.e. $\delta_{i,k} = 1$ if and only if $i = k$ and zero otherwise.

this dissertation, the following sections first report the asymptotic eigenvalue distribution of some simple matrix ensembles. Then, the Stieltjes transform will be introduced to deal with more complex matrix models. These basic concepts, together with some useful results in Section 2.2.3 and the theory in Section 2.3, should be sufficient to understand the following technical chapters.

2.2.1 The eigenvalue distribution: basic results

Before delving into asymptotic properties of large matrices, let us define the *empirical eigenvalue distribution* of a generic matrix.

Definition 2.1. Let $\mathbf{A}_N$ be a square matrix with N real eigenvalues denoted by $\lambda_1, \ldots, \lambda_N$. Then, the empirical eigenvalue distribution of $\mathbf{A}_N$, evaluated at $x \in \mathbb{R}$, is

$$F_N(x) = \frac{1}{N}|\{\lambda_n, n = 1 \ldots N : \lambda_n \leq x\}|, \tag{2.1}$$

where $|\cdot|$ is the cardinality of the set. In other words, $F_N(x)$ is the fraction of eigenvalues which are lower than or equal to x.

Note that when $\mathbf{A}_N$ is a random matrix, its eigenvalues are random quantities and, hence, the empirical eigenvalue distribution $F_N(x)$ is a random function.

The basic concept behind random matrix theory can be outlined as follows: for some models, or ensembles, of random matrices, the empirical eigenvalue distribution tends to a deterministic distribution $F(x)$ when the dimensions of the matrix tend to infinity.

Example 2.3 (Wigner-type matrices)**.** Consider a $N \times N$ Hermitian (self-adjoint) Gaussian random matrix $\mathbf{W}_N$, with $\frac{1}{2}N(N+1)$ independent Gaussian entries with zero mean and variance $\frac{1}{N}$. Note that the elements on the diagonal will be real, while all the others are complex with variance $\frac{1}{2N}$ per dimension. It can be proven that, as the dimension N tends to infinity, the empirical eigenvalue distribution $F_N(x)$ of $\mathbf{W}_N$ converges to the absolutely continuous deterministic function $F(x)$ corresponding to the density

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = \begin{cases} \dfrac{1}{2\pi}\sqrt{4-x^2} & \text{for } x \in [-2, 2], \\ 0 & \text{elsewhere,} \end{cases}$$

known as the *Wigner's semicircular law* and depicted in Figure 2.2.

This result was first proven by E. P. Wigner in 1958 [65] (thus the name of this ensemble) for convergence in expectation and then, in 1967, T. W. Arnold [66] extended it to almost sure convergence, i.e.

$$\Pr\Big[\lim_{N\to+\infty} F_N(x) = F(x)\Big] = 1, \quad \forall x \in \mathbb{R}.$$

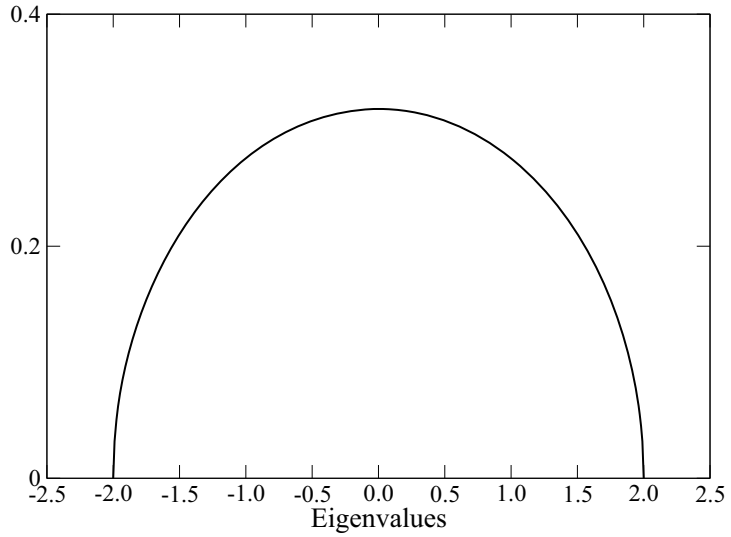


Figure 2.2: Wigner's semicircular law.

Similarly, also the Gaussian constraint has been dropped: the empirical eigenvalue distribution $F_N(x)$ will always converge to the Wigner's semicircular law as long as the entries of $N\mathbf{W}_N$ have zero mean and bounded moments.

Example 2.4 (Wishart matrices)**.** Consider a matrix $\mathbf{N} \in \mathbb{C}^{N\times M}$ with independent circularly symmetric Gaussian distributed random entries with zero mean and variance $\frac{1}{N}$. Then, the matrix $\mathbf{M}_N = \mathbf{N}\mathbf{N}^H$ is referred to as a *Wishart matrix*. (To be precise, this is a special case of Wishart matrix. The general one admits outer correlation and is treated in Example 2.6.) Matrices with this structure are very common models in communications theory, e.g. sample correlation matrices or CDMA interference covariance matrices (cf. Example 2.2).

For the Wishart matrices, it has been proven [67] that the empirical eigenvalue distribution $G_N(x)$ tends almost surely to the so-called *Marčenko-Pastur law* $G(x)$ when both M and N grow without bound while their ratio converges to a constant value c, i.e. $\frac{M}{N} \to c$. The density associated with $G(x)$ is

$$\frac{\mathrm{d}G(x)}{\mathrm{d}x} = [1-c]^+\delta_0(x) + \\ + \frac{\sqrt{4c-(x-(1-c))^2}}{2\pi x}\mathbb{1}\left\{x \in \left[(\sqrt{c}-1)^2, (\sqrt{c}-1)^2\right]\right\}, \quad (2.2)$$

where $[x]^+ = \max\{0, x\}$, $\delta_0(x)$ is the Dirac delta centered in zero and the indicator function for the event ω, $\mathbb{1}\{\omega\}$, is equal to one if ω is true and

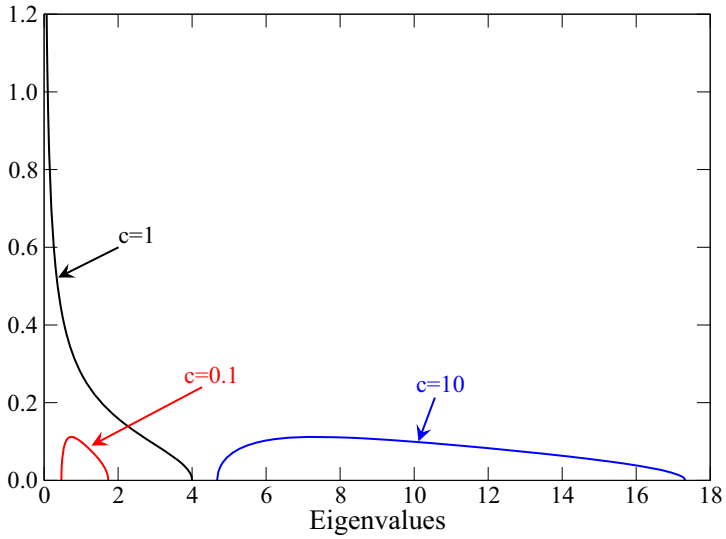


Figure 2.3: The absolutely continuous part of the Marčenko-Pastur law.

to zero otherwise. With some abuse of notation, the Dirac delta has been introduced to account for a possible point mass in zero. Indeed, when $M < N$ and $c < 1$, the matrix $\mathbf{M}_N$ is rank deficient with a fraction $(N-M)/N = 1-c$ of null eigenvalues. In this case, the asymptotic eigenvalue distribution is not derivable in $x = 0$.

It also makes sense to ask ourselves about the behavior when only one dimension of $\mathbf{N}$ tends to infinity. By means of the central limit theorem, it is simple to show that, when M tends to infinity with N constant ($c \to +\infty$), the normalized Wishart matrix $\frac{N}{M}\mathbf{M}_N$ tends in law to a Hermitian Gaussian random matrix plus an identity, i.e. $\sqrt{M}(\frac{1}{c}\mathbf{M}_N - \mathbf{I}_N) \xrightarrow{D} \mathbf{W}_N$, with $\mathbf{W}_N$ a Wigner matrix as in the previous example. A similar result can be found when $N \to +\infty$ and M constant ($c \to 0$), by noting that the non-zero eigenvalues of $\mathbf{N}\mathbf{N}^H$ are equal to those of $\mathbf{N}^H\mathbf{N}$. Figure 2.3 depicts the absolutely continuous second term of (2.2), where it can be noticed that the Marčenko-Pastur law tends to be shaped as a shifted version of the Wigner's semicircular law both for $c \to 0$ and $c \to +\infty$.

To conclude the example, let us mention that some constraints on the distribution of $\mathbf{N}$ can be relaxed: it is enough that the independent entries of $\sqrt{N}\mathbf{N}$ have zero mean and bounded moments of all orders for the eigenvalue distribution of $\mathbf{M}_N$ to converge to the Marčenko-Pastur law.

2.2.2 Further results: the Stieltjes transform

The last example gives an explicit analytic expression of the asymptotic eigenvalue distribution of a Wishart matrix. The model $\mathbf{M}_N = \mathbf{N}\mathbf{N}^H$ with independent entries for $\mathbf{N}$, however, is not very practical. In real cases,

it is much more common to identify some correlation along the rows or the columns (or both) of the matrix $\mathbf{N}$. For the resulting matrices, it is usually still possible to show the convergence of their eigenvalue distributions as $M, N \to +\infty$ and $M/N \to c$. Unfortunately, the limiting distribution cannot generally be expressed in an explicit analytic form. Before giving more details and some examples, let us introduce the following operation on distributions.

Definition 2.2. For a given distribution $F(x)$, its *Stieltjes transform* is defined as[4]

$$m(z) = \int \frac{1}{z - x} \, \mathrm{d}F(x). \tag{2.3}$$

When $F(x)$ corresponds to a probability measure, $m(z)$ is an analytic function in $\mathbb{C}^+ = \{z \in \mathbb{C} : \Im(z) > 0\}$. If, in addition, the distribution has compact support, the Stieltjes transform can be seen as a formal power series in z, i.e. $m(z) \in \mathbb{C}[[z^{-1}]]$:

$$m(z) = z^{-1} + \sum_{k=1}^{+\infty} \mu_k z^{-k-1},$$

where $\mu_k = \int x^k \, \mathrm{d}F(x)$ is the k-th order moment of $F(x)$.

Next, some examples are given of how the asymptotic eigenvalue distribution of some matrix models can be expressed as the solution to a functional equation in terms of the Stieltjes transform. It can happen that these functional equations admit more than one solution: the correct one can be usually identified by observing that a proper Stieltjes transform tends to zero when the module of z grows without bound.

Example 2.5. Let us consider matrices of the form $\mathbf{R}_N = \mathbf{A} + \mathbf{N}\mathbf{T}\mathbf{N}^H$, with $\mathbf{A} \in \mathbb{C}^{N \times N}$ a Hermitian deterministic matrix and $\mathbf{T} \in \mathbb{C}^{M \times M}$ a positive-definite diagonal matrix. Name $F_A(x)$ and $T(x)$, respectively, their limiting eigenvalue distribution for N and M, respectively, tending to infinity. Let $\mathbf{N} \in \mathbb{C}^{N \times M}$ be filled with complex random entries with zero mean, variance $\frac{1}{N}$ and independent of both one another and $\mathbf{A}$. Then, in the limit for M and N tending to infinity with $M/N \to c$, the eigenvalues of $\mathbf{R}_N$ are distributed according to a deterministic distribution whose Stieltjes transform $m(z)$ is a solution to the functional equation

$$m(z) = m_A\left(z - c \int \frac{x}{1 - x m(z)} \, \mathrm{d}T(x)\right),$$

where $m_A(z)$ is the Stieltjes transform of $F_A(x)$.

[4] A definition with a changed sign also exists.

This matrix model is quite common in practice since it can represent covariance matrices with a two-component structure: the first component, $\mathbf{A}$, is associated to noise and interference, while the second one, $\mathbf{NTN}^H$, is due to the desired signal. For instance, for the signal model in Example 2.2, the matrix $\mathbb{E}[\mathbf{y}\mathbf{y}^H]$ has the structure of $\mathbf{R}_N$, with $\mathbf{A} = \mathbf{I}_N$, $\mathbf{T} = P\mathbf{I}_K$ and $\mathbf{N} = \begin{bmatrix}\mathbf{s}_1 & \cdots & \mathbf{s}_K\end{bmatrix}$.

Example 2.6. Assume now that we are interested in a sample covariance matrix with outer correlations given by a deterministic matrix $\mathbf{T}$. A possible model is given by $\mathbf{C}_N = \mathbf{T}^{1/2}\mathbf{N}\mathbf{N}^H\mathbf{T}^{1/2}$, where $\mathbf{N}$ has the same statistical model as above and $\mathbf{T}^{1/2}$ denotes the Hermitian non-negative square root of the matrix $\mathbf{T}$. This is the general model for Wishart matrices (c.f. 2.4). Let $T(x)$ be the asymptotic eigenvalue distribution of $\mathbf{T}$, if it exists. Then, the empirical eigenvalue distribution of $\mathbf{C}_N$ tends almost surely to a deterministic distribution whose Stieltjes transform $m(z)$ is the unique solution to the equation

$$m(z) = \int \frac{1}{z - (c - 1 + zm(z))x} \,\mathrm{d}T(x).$$

The proofs to these results where given by J. W. Silverstein, respectively in [68] and [69]. The interested reader should be careful when comparing equations, since some slight modifications were necessary to match with the given definition of the Stieltjes transform.

Once the Stieltjes transform is known, the corresponding probability function can be recovered by means of the Stieltjes inversion formula, namely

$$\int_a^b \mathrm{d}F(x) = -\frac{1}{\pi} \lim_{y \to 0} \int_a^b \Im[m(x + jy)] \,\mathrm{d}x, \tag{2.4}$$

where a and b are two continuity points of F. This step, however, is seldom necessary since most of the quantities of interest in the communications field can be expressed directly in terms of the Stieltjes transform, evaluated at some particular value of z. An intuition for this property can be easily deduced: it is enough to observe that, if $F_N(x)$ is the empirical eigenvalue distribution of the $N \times N$ matrix $\mathbf{A}_N$, then the Stieltjes transform $m_N(z)$ is nothing else than the normalized trace of the resolvent of the matrix, i.e.

$$m_N(z) = \frac{1}{N}\mathrm{tr}\Big\{(z\mathbf{I}_N - \mathbf{A}_N)^{-1}\Big\}.$$

Thus, if we denote as $m(z)$ the Stieltjes transform of the distribution $F(x)$, the following limits are equivalent:

$$F(x) = \lim_{N \to +\infty} F_N(x), \tag{2.5a}$$

$$m(z) = \lim_{N \to +\infty} m_N(z), \tag{2.5b}$$

$$m(z) = \lim_{N \to +\infty} \frac{1}{N}\mathrm{tr}\Big\{(z\mathbf{I}_N - \mathbf{A}_N)^{-1}\Big\}, \tag{2.5c}$$

where the convergence mode is the same in the three cases. The proof is a straightforward application of the Portmanteau theorem[5] (see, e.g., [70, Chapter 1]). Some examples are given in the following section.

2.2.3 Additional results

According to the results of the previous section, the distribution of the eigenvalues of some large-enough random matrices can be approximated by a deterministic distribution. The knowledge of this distribution, or of its Stieltjes transform, is sufficient to solve many practical problems like those reported hereafter. First, let us recall this useful lemma which is a direct consequence of [71, Lemma 1].

Lemma 2.1. *Let* $\mathbf{A}$ *be an* $N \times N$ *complex random matrix and* $\mathbf{x}, \mathbf{y}$ *two* $N \times 1$ *complex random vectors, all independent of one another. Assume that the elements of* $\mathbf{x}$ *and* $\mathbf{y}$ *are centered i.i.d. random variables with unit variance and finite eighth-order moment. Also assume that the spectral norm (denoted by* $\|\cdot\|$*) of* $\mathbf{A}$ *is uniformly bounded, i.e.* $\sup_{N\in\mathbb{N}} \|\mathbf{A}\|$ *is finite. Then*

$$\frac{1}{N}(\mathbf{x}^H\mathbf{A}\mathbf{x} - \mathrm{tr}\{\mathbf{A}\}) \to 0 \qquad\qquad \frac{1}{N}\mathbf{x}^H\mathbf{A}\mathbf{y} \to 0$$

almost surely as N *tends to infinity.*

Example 2.7. For the signal model in Example 2.2, assume that the spreading sequences are known at the receiver and that we are interested in estimating the symbol sent by user 1 (without loss of generality), considering all other users as interference. The signal-to-interference-plus-noise ratio (SINR) at the output of a linear minimum-mean-square-error receiver can be expressed as

$$SINR = \frac{P}{N}\mathbf{s}_1^H\left(\mathbf{I}_N + \frac{P}{N}\mathbf{S}_1\mathbf{S}_1^H\right)^{-1}\mathbf{s}_1,$$

where $\mathbf{S}_1 = \begin{bmatrix}\mathbf{s}_2 & \cdots & \mathbf{s}_K\end{bmatrix}$. Assuming that the elements of the spreading sequences have finite eighth-order moment, Lemma 2.1 implies that

$$\lim SINR = P\lim\frac{1}{N}\mathrm{tr}\left\{\left(\mathbf{I}_N + \frac{P}{N}\mathbf{S}_1\mathbf{S}_1^H\right)^{-1}\right\},$$

where the limits are intended for $K, N \to +\infty$ and $K/N \to c < +\infty$ (all spreading sequences are independent of one another and, thus, so are $\mathbf{s}_1$ and $\mathbf{S}_1$). According to (2.5), the left-hand side is $-m(-1/P)$, being $m(z)$ the Stieltjes transform of the asymptotic eigenvalue distribution of $\frac{1}{N}\mathbf{S}_1\mathbf{S}_1^H$.

[5] To be rigorous, the function $F(x)$, defined as the pointwise limit of $F_N(x)$ for $N \to +\infty$, does not need to be a proper distribution, that is it can be $\lim_{x\to+\infty} F(x) \neq 1$. Tightness of the sequence $\{F_N\}$ should be verified.

Lemma 2.1 shows how to deal with quadratic forms. Another form that often appears in communications problems is the "log det" one, which is typical when computing capacities of, e.g., MIMO channels.

Proposition 2.1. *Let $F(x)$ be a probability distribution with real non-negative support and define the function*

$$\mathcal{V}(\gamma) = \int \ln(1+\gamma x)\,\mathrm{d}F(x), \tag{2.6}$$

where ln *is the natural logarithm and γ a real non-negative number. Moreover, let $m(z)$ be the Stieltjes transform associated to $F(x)$. Then, the following relation holds true:*

$$\gamma\frac{\mathrm{d}\mathcal{V}(\gamma)}{\mathrm{d}\gamma} = 1 + \frac{1}{\gamma}m\left(-\frac{1}{\gamma}\right).$$

$\mathcal{V}(\gamma)$ is sometimes referred to as the *Shannon transform* of the distribution $F(x)$. The proof follows directly from deriving both sides of (2.6) and comparing with (2.3).

Observe that, for a diagonalizable positive semi-definite $N \times N$ (random) matrix $\mathbf{A}_N$ with empirical eigenvalue distribution $F_N(x)$, it holds that

$$\frac{1}{N}\ln\det(\mathbf{I}_N + \gamma\mathbf{A}_N) = \int \ln(1+\gamma x)\,\mathrm{d}F_N(x).$$

Hence, Proposition 2.1 gives us a tool to compute $\lim_{N\to+\infty}\frac{1}{N}\ln\det(\mathbf{I}_N + \gamma\mathbf{A}_N)$.

Example 2.8. The maximum spectral efficiency for the system in Example 2.2 can be written as

$$I = \frac{1}{N}\ln\det\left(\mathbf{I}_N + \frac{P}{N}\mathbf{S}\mathbf{S}^H\right),$$

in nats per degree of freedom and where $\mathbf{S} = \begin{bmatrix}\mathbf{s}_1 & \cdots & \mathbf{s}_K\end{bmatrix}$. According to the theory presented so far, when the number of users K and the spreading factor N grow indefinitely at the same rate, the spectral efficiency tends to the deterministic quantity

$$I^* = \int_0^P \frac{1}{\gamma} + \frac{1}{\gamma^2}m\left(-\frac{1}{\gamma}\right)\mathrm{d}\gamma,$$

where $m(z)$ is the Stieltjes transform associated to the asymptotic distribution of the eigenvalues of $\frac{1}{N}\mathbf{S}\mathbf{S}^H$.

The last two examples are treated with plenty of details in [72], where explicit expressions are given for asymptotic SINR and spectral efficiency.

A more complex case, where each user is received with a different power, is analyzed in [73]. Basically, a diagonal matrix $\mathbf{P}$ containing the user powers is introduced, modifying the signal covariance matrix from $\mathbf{I}_N + \frac{P}{N}\mathbf{S}\mathbf{S}^H$ to $\mathbf{I}_N + \frac{1}{N}\mathbf{S}\mathbf{P}\mathbf{S}^H$. The last structure is still covered by the model in Example 2.5.

Summarizing in few words the theory introduced in this section, the Stieltjes transform and, more generally, random matrix theory supply an important tool to deal with matrices of random variables. Indeed, for some random matrix models covering an important set of practical cases, it has been shown that their eigenvalues tend to be distributed according to a deterministic function when the matrix dimensions grow indefinitely at the same rate. In other words, if the random linear system is large enough (but with dimensions that still make practical sense, as supported by simulation results), its performances can be accurately estimated by analyzing a deterministic problem.

Working with large matrices also presents another important advantage. Suppose that we are given two matrices $\mathbf{A}, \mathbf{B}$ and that we are interested in the eigenvalues of $\mathbf{A} + \mathbf{B}$ and $\mathbf{AB}$. For finite dimensions, this is an involved problem, since the spectrum of the sum (product) of matrices depends not only on the eigenvalues but also on the detailed structure (the eigenvectors) of $\mathbf{A}$ and $\mathbf{B}$ —see, e.g., [74] for the sum of Hermitian matrices. However, for infinite-dimension matrices like those discussed above, simple tools exist to express the eigenvalue distributions of $\mathbf{A} + \mathbf{B}$ and $\mathbf{AB}$ in terms of the spectra of $\mathbf{A}$ and $\mathbf{B}$. This is an application of free-probability theory for non-commutative random variables.

2.3 Free-probability for non-commutative random variables

The set of all possible $N \times N$ square matrices with complex entries forms an algebra over the field $\mathbb{C}$: roughly speaking, the sum or the product of any two matrices still belongs to the set. Therefore, the theory developed for these more abstract mathematical entities can be useful to solve matrix problems. In such a context, we are now going to introduce non-commutative probability theory and the concept of "freeness" or "free independence", which can be considered in analogy with the notion of independence in classic probability theory. This exposition will cover only basic ideas, especially those related to random matrices. The interested reader can also refer to the tutorials [75,76] or to the more thorough monographs [77] and [78]. The latter, in particular, is co-authored by D. Voiculescu, who first introduced free probability (in early 1980's) in order to study some problems related to operator algebras.

To start with, the following definitions are needed.

Definition 2.3. Let $\mathcal{A}$ be a unital algebra[6] over $\mathbb{C}$ (denote with $\mathbf{1}$ its unity) and $\varphi(\cdot)$ a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ such that $\varphi(\mathbf{1}) = 1$. Then, $(\mathcal{A}, \varphi)$ is called a *non-commutative probability space* and each element a of $\mathcal{A}$ is a *non-commutative random variable*. The functional φ, sometimes called "expectation", is said to be a *trace* if $\varphi(ab) = \varphi(ba)$, for any $a, b \in \mathcal{A}$. The number $\varphi(a^k)$ is the *k-th moment* of a.

If, moreover, $\mathcal{A}$ is a $*$-algebra[7], then φ is assumed to be a *state* on $\mathcal{A}$, i.e.: (i) $\varphi(\mathbf{1}) = 1$, (ii) $\varphi(a^*) = \varphi^*(a)$ and (iii) $\varphi(a^*a) \geq 0$, for every $a \in \mathcal{A}$. (With some abuse of notation, we denote by * both the algebra involution and the complex conjugation in $\mathbb{C}$. The context should be sufficient to discriminate between the two operations.)

In a few words, non-commutative probability is an approach to understanding of some problems in non-commutative algebras that is based on typical notations and ideas from classical probability theory. Furthermore, classical bounded random variables may be considered as non-commutative ones, even though of a degenerate kind (classical random variables do commute). [77, Example 1.2.2] and comments thereafter show a possible, non-unique, representation.

Returning to the original problem, it is straightforward to identify a $*$-algebra in the set of all $N \times N$ matrices with complex random entries having bounded moments of all orders. The sum and the product operations are the usual matrix ones, the identity matrix $\mathbf{I}_N$ is the product unity and the Hermitian transposition is the endowed involution $*$. Thus, we only need a functional $\varphi_N(\cdot)$ to define a non-commutative probability space. The choice $\varphi_N(\cdot) = \frac{1}{N}\mathbb{E}\mathrm{tr}[\cdot]$ (the expected value, in the classical sense, of the normalized trace of the matrix) satisfies the requirements in Definition 2.3: every matrix $\mathbf{A}$ is then associated to a set of moments $\varphi_N(\mathbf{A}^k) = \frac{1}{N}\mathbb{E}\mathrm{tr}[\mathbf{A}^k]$, for $k = 1, 2, \ldots$

Now, we focus our attention on the subalgebra of Hermitian matrices. Let $\lambda_1, \ldots, \lambda_N$ be the N eigenvalues, with multiplicities, of matrix $\mathbf{A}$. Since $\mathbf{A}$ is Hermitian, all the eigenvalues are real-valued and their empirical distribution $F_A(x)$ introduced in (2.1) exists and is well-defined. Considering $\mathbf{A}$ as a non-commutative random variable, its k-th order moment is given

[6] An algebra over a field $\mathbb{F}$ is a vector space V over $\mathbb{F}$ with a product operation $\cdot : V \times V \to V$ such that, for any $v, w, z \in V$ and $a, b \in \mathbb{F}$, (i) $(v + w) \cdot z = v \cdot z + w \cdot z$, (ii) $v \cdot (w + z) = v \cdot w + v \cdot z$ and (iii) $(av) \cdot (bw) = (ab)(v \cdot w)$. Observe that the first two properties are not redundant since, in general, the product is not a commutative operation.

To be unital, an algebra should present a neutral element $\mathbf{1}$ with respect to the product: $\mathbf{1} \cdot v = v \cdot \mathbf{1} = v$.

[7] A $*$-algebra (pronounced star-algebra) is a unital algebra endowed with an involution $*$ such that, for any two elements v, w of the algebra, (i) $v \to v^*$ is linear, (ii) $(v \cdot w)^* = w^* \cdot v^*$ and (iii) $(v^*)^* = v$.

by

$$\varphi_N(\mathbf{A}^k) = \frac{1}{N}\mathbb{E}\mathrm{tr}[\mathbf{A}^k] = \frac{1}{N}\mathbb{E}\left[\sum_{n=1}^{N}\lambda_n^k\right] = \mathbb{E}\left[\int x^k \,\mathrm{d}F_A(x)\right].$$

Observe that the right-hand side of the previous equation coincides with the k-th order moment of the distribution $F_A(x)$. We can conclude that the empirical distribution of the eigenvalues of $\mathbf{A}$ not only characterizes the spectrum of the matrix, but is also the distribution of the non-commutative random variable $\mathbf{A}$ belonging to the $*$-algebra of $N \times N$ random Hermitian matrices. This last interpretation will be at the basis of the techniques presented in the following sections.

It should be kept in mind, however, that the distribution function does not need to exist. For instance, a generic non-Hermitian square complex matrix may have non-real eigenvalues and, hence, a distribution function like the one in (2.1) cannot be defined. In a more abstract sense, the distribution of a non-commutative random variable simply is the collection of all its moments[8]. Equivalently, in a non-commutative probability space $(\mathcal{A}, \varphi)$, two random variables $a, b \in \mathcal{A}$ are equally distributed whenever $\varphi(a^k) = \varphi(b^k)$ for all k. As we have seen for the Hermitian matrices, in some occasions it is possible to express explicitly the distribution function: it can be shown that only normal variables (i.e., such that $a^*a = aa^*$) have this property. The support of the correspondent probability measure will be included in $\mathbb{C}$.

Following the same approach, the joint distribution of multiple non-commutative random variables is also defined in terms of all joint moments. For example, given two random variables a, b in the non-commutative probability space $(\mathcal{A}, \varphi)$, their joint distribution is the collection of moments

$$\varphi(a), \varphi(b), \varphi(a^2), \varphi(b^2), \varphi(ab), \varphi(a^2b), \\ \varphi(ab^2), \varphi(a^2b^2), \varphi(abab), \varphi(a^2bab), \varphi(a^3b^2), \ldots$$

Observe that, due to the non-commutative structure of $\mathcal{A}$, moments like, e.g., $\varphi(a^2b^2)$ and $\varphi(abab)$ are different and they all have to be considered when characterizing the distribution. Conversely, forms like, e.g., $\varphi(ab)$ and $\varphi(ba)$ coincide since the functional φ is assumed to be tracial unless otherwise specified.

When $\mathcal{A}$ is a $*$-algebra, the joint distribution of a and a^* is called $*$-distribution of the random variable a. The joint $*$-distribution of multiple non-commutative random variables $\{a_1, a_2, \ldots, a_n\}$ is, thus, the collection of moments of $\{a_1, a_1^*, a_2, a_2^*, \ldots, a_n, a_n^*\}$.

[8] This is not the case in a classical probability space: a classical distribution function is univocally determined by its moments only if it is supported by a finite interval; for infinite supports, supplementary conditions like the Carleman's one are needed (see, e.g., [79, Section 4.5]).

For classic probability spaces, it is well known that joint distribution functions of independent random variables factorize into the product of the individual distributions. Similarly, one can wonder whether complicated joint moments like $\varphi(a^2bab)$ may be expressed in terms of the simple moments of a and b. In the next section, the answer is shown to be positive for free non-commutative random variables. In non-commutative probability, *freeness*, or *free independence*, plays a role analogous to independence in classic probability: both of them refer to special but very fundamental situations where particular relations hold among joint moments of some variables. We will see, however, that the factorization rules are somehow different in the two contexts.

2.3.1 Free random variables

In this section, a first formal definition of freeness will be presented, together with some insights into this aspect of fundamental importance for the theory we are developing. First, to highlight analogies and differences, we formalize the idea of independent random variables in the new algebraic definition of probability space.

Definition 2.4. Let $(\mathcal{A}, \varphi)$ be a probability space. The m subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_m \subseteq \mathcal{A}$ are said independent if:

1. they commute, i.e. $ab = ba$ for any $a \in \mathcal{A}_i$ and $b \in \mathcal{A}_j$, with $i, j = 1, \ldots, m$ and $i \neq j$;

2. the expectation φ factorizes as follows:

$$\varphi(a_1 \cdots a_m) = \varphi(a_1) \cdots \varphi(a_m),$$

for all $a_1 \in \mathcal{A}_1, \ldots, a_m \in \mathcal{A}_m$.

By extension, two random variables $a, b \in \mathcal{A}$ are independent when the subalgebras they generate are independent, meaning that $ab = ba$ and $\varphi(a^s b^t) = \varphi(a^s)\varphi(b^t)$, for all $s, t \in \mathbb{N}$.

Note that the definition of independence does not apply to general algebras, but only to commutative ones. This is the case for the algebra generated by classical random variables, as mentioned in the previous section: classic independence fits with this new definition when choosing the expected value $\mathbb{E}[\cdot]$ as the functional $\varphi(\cdot)$. Conversely, two $N \times N$ square random matrices with independent entries are not necessarily independent since, in general, they do not commute. Strictly speaking, only diagonal random matrices are (with an opportune choice of φ).

The great advantage of dealing with independent (commutative) random variables is the possibility to express joint distributions and moments

in terms of the respective individual distributions and moments. For example, the moments of the sum and the product of two independent random variables a, b are given by

$$\varphi\Big((a+b)^k\Big) = \sum_{i=0}^{k} \binom{k}{i} \varphi(a^i)\varphi(b^{k-i}),$$
$$\varphi\Big((ab)^k\Big) = \varphi(a^k)\varphi(b^k).$$

Freeness identifies another class of random variables for which similar properties hold. Then, if $a, b \in (\mathcal{A}, \varphi)$ are two non-commutative free random variables, the collection of moments of $a+b$ and ab can be written explicitly from the single distributions of a and b. In particular, the distribution of the eigenvalues of sums and products of free matrices may be computed from the spectra of the individual matrices. The resulting equations, however, are different from those obtained for independent variables.

Definition 2.5. Let $(\mathcal{A}, \phi)$ be a non-commutative probability space. The unital subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_m \subseteq \mathcal{A}$ are said to be free if $\varphi(a_1 \cdots a_n) = 0$ for any $n \in \mathbb{N}$ and any choice of the random variables $a_1, \ldots, a_n$ such that:

1. $a_i \in \mathcal{A}_{j(i)}$, with $j(i) \in \{1, \ldots, m\}$;
2. $j(i) \neq j(i+1)$, for $i = 1, \ldots, n-1$ and
3. $\varphi(a_i) = 0$, for $i = 1, \ldots, n$.

The n random variables $a_1, \ldots, a_n \in \mathcal{A}$ (or, more generally, the n subsets $\chi_1, \ldots, \chi_n \subset \mathcal{A}$) are said to be free if the unital subalgebras they generate are free.

Sometimes, the name $*$-freeness is used when dealing with $*$-algebras.

Let us consider some examples to clarify this definition and its implications. Take three free random variables $a, b, c \in (\mathcal{A}, \varphi)$ with zero mean, i.e. $\varphi(a) = \varphi(b) = \varphi(c) = 0$. The freeness definition ensures that, e.g.,

$$\varphi(abcbacbababc) = 0,$$

since no two consecutive variables are the same (i.e. belong to the same subalgebra). Conversely, at first sight nothing can be said about, e.g., $\varphi(ababc^2ba)$: $c^2 = c \cdot c$ belongs to the same subalgebra as c but its moment does not need to be zero ($\varphi(c) = 0 \not\Rightarrow \varphi(c^2) = 0$). Nevertheless, the definition of freeness offers more insight, as shown by the following example.

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two free subalgebras in the non-commutative probability space $(\mathcal{A}, \varphi)$ and let $a \in \mathcal{A}_1$ and $b \in \mathcal{A}_2$ be two free random variables without any constraint on their first-order moment. Now, build the two

centered random variables $\bar{a} = a - \varphi(a)\mathbf{1} \in \mathcal{A}_1$ and $\bar{b} = b - \varphi(b)\mathbf{1} \in \mathcal{A}_2$. Then, $\bar{a}$ and $\bar{b}$ belong to two free subalgebras and have zero mean, yielding

$$\varphi\big(\bar{a}\bar{b}\big) = \varphi((a - \varphi(a)\mathbf{1})(b - \varphi(b)\mathbf{1})) = 0.$$

The linearity of φ implies $\varphi(ab) = \varphi(a)\varphi(b)$. More generally

$$\varphi\big(\big(a^m - \varphi\big(a^m\big)\mathbf{1}\big)\big(b^n - \varphi\big(b^n\big)\mathbf{1}\big)\big) = 0 \Rightarrow \varphi\big(a^m b^n\big) = \varphi\big(a^m\big)\varphi\big(b^n\big).$$

Even though this factorization is exactly the same as the one obtained for independent random variables, differences appear when considering more complicated moments. For instance, by developing as before the equation

$$\varphi\big(\bar{a}\bar{b}\bar{a}\bar{b}\big) = 0$$

one obtains

$$\varphi(abab) = \varphi(a^2)\varphi^2(b) + \varphi^2(a)\varphi(b^2) - \varphi^2(a)\varphi^2(b).$$

With a similar approach, all the joint moments of a and b can be computed in terms of individual moments of lower order. Thus, the joint distribution of the couple of random variables $\{a, b\}$ can be characterized in terms of the distributions of a and b. In Section 2.3.3 we will see a more schematic approach to compute joint moments which will also yield to a simpler and more intuitive definition of freeness.

To conclude the comparison between independent and free random variables, let us show that two independent random variables cannot be free, except in the trivial case where one of the two is deterministic. Indeed, if a and b are independent then they commute and we get:

$$\varphi(a^2)\varphi(b^2) = \varphi(a^2b^2) = \varphi(abab) = \varphi(a^2)\varphi^2(b) + \varphi^2(a)\varphi(b^2) - \varphi^2(a)\varphi^2(b),$$

which is equivalent to

$$\big[\varphi(a^2) - \varphi^2(a)\big]\big[\varphi(b^2) - \varphi^2(b)\big] = 0.$$

This is true only if at least one of the two terms is zero or, equivalently, if the variance of one of the two random variables is zero.

Before going a little further into the understanding of freeness, the next section brings some practicality to the topic, showing how abstract free random variables may represent special classes of random matrices.

2.3.2 Asymptotic freeness of random matrices

The notions described in the previous sections should be enough to sense that free probability may result very helpful in studying the spectra of sums and products of some random matrices. At that purpose, it has already

been mentioned that the $*$-algebra of $N \times N$ random matrices with complex entries forms a non-commutative probability space together with the natural functional $\varphi_N(\cdot) = \frac{1}{N}\mathbb{E}\mathrm{tr}[\cdot]$, where the expected value is computed with respect to the distributions of the matrix entries and is assumed to be finite. We still need to show which matrices can be assimilated to free random variables.

Definition 2.6. Let $\{\mathbf{A}_1, \dots, \mathbf{A}_m\}$ be a collection of $N \times N$ complex random matrices and let $\mu^N_{\mathbf{A}_1,\dots,\mathbf{A}_m}$ be their joint $*$-distribution as non-commutative random variables, i.e. the collection of all their joint moments. Formally, the set of random matrices $\{\mathbf{A}_1, \dots, \mathbf{A}_m\}$ is said to be asymptotically free as $N \to +\infty$ when $\mu^N_{\mathbf{A}_1,\dots,\mathbf{A}_m}$ tends to the joint $*$-distribution of m free random variables or, equivalently, when the joint moments tend to the joint moments of m free random variables. More generally, given the disjoint sets of indeces $\mathcal{I}_1, \dots, \mathcal{I}_m$, the sets of random matrices $\{\mathbf{A}_i, i \in \mathcal{I}_1\}, \dots, \{\mathbf{A}_i, i \in \mathcal{I}_m\}$ form an *asymptotically free family* if any collection of random matrices, each one chosen from a different set, is asymptotically free.

Asymptotic freeness of some particular classes of random matrices aroused the interest in Voiculescu's research on free probability and free operators. It was the same Voiculescu who first showed that any collection of Hermitian Gaussian random matrices together with the set of deterministic diagonal matrices form an asymptotically free family as the dimension grow indefinitely [80]. The following theorem appears, together with its proof, in [81] and expresses that first result in a more general form.

Theorem 2.1. *Let $\mathbf{A}(N, s)$ be a collection of $N \times N$ random matrices with complex entries denoted by $a(i, j; N, s), 1 \le i, j \le N$, indexed by $s \in \mathbb{N}$. Assume $a(i, j; N, s) = a^*(j, i; N, s)$ (the matrices are Hermitian) and that*

$$\{a(i, j; N, s) : 1 \le i \le j \le N, s \in \mathbb{N}\}$$

is a set of independent random variables such that

$$\begin{aligned} \mathbb{E}[a(i, j; N, s)] &= 0 && 1 \le i \le j \le N, s \in \mathbb{N} \\ \mathbb{E}\big[|a(i, j; N, s)|^2\big] &= \frac{1}{N} && 1 \le i \le j \le N, s \in \mathbb{N} \\ \sup_{1\le i\le j\le N} \mathbb{E}\big[|a(i, j; N, s)|^m\big] &= O\big(N^{-m/2}\big) && m \ge 1, s \in \mathbb{N}. \end{aligned}$$

Consider also the collection $(\mathbf{D}(N, t))_{t\in\mathbb{N}} \subseteq \Delta_N$, where Δ_N is the algebra of $N \times N$ deterministic matrices with uniformly bounded spectral radius[9] and a limit eigenvalue distribution as $N \to +\infty$. Then, the family of sets

$$\big(\{\mathbf{D}(N, t) : t \in \mathbb{N}\}, \{\mathbf{A}(N, 1)\}, \{\mathbf{A}(N, 2)\}, \dots\big)$$

[9]The spectral radius of a square matrix is maximum among the moduli of its eigenvalues. If the matrix has infinite dimensions, the spectral radius is taken as the supremum over all possible eigenvalues.

is asymptotically free as $N \to +\infty$ and, moreover, the limit distribution of each $\mathbf{A}(N,s)$ is a semicircle law.

This theorem implies that any Hermitian random matrix $\mathbf{A}(N,s_1)$ satisfying the conditions above is asymptotically free from any other similar matrix $\mathbf{A}(N,s_2)$ with independent entries. Furthermore, $\mathbf{A}(N,s_1)$ is also asymptotically free from any deterministic matrix in $\{\mathbf{D}(N,t) : t \in \mathbb{N}\}$. Observe that this does not mean in any way that two deterministic matrices $\mathbf{D}(N,t_1)$ and $\mathbf{D}(N,t_2)$ are asymptotically free.

Example 2.9. Let $\mathbf{A}$ and $\mathbf{B}$ be two $N \times N$ Hermitian random matrices such that the entries of $\sqrt{N}\mathbf{A}$ and $\sqrt{N}\mathbf{B}$ are independent, zero mean and have bounded moments of all orders. Let $\mathbf{D}$ be a deterministic $N \times N$ matrix satisfying the properties listed in the theorem. Then, as $N \to +\infty$, the three matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{D}$ become models for free non-commutative random variables with respect to the usual functional $\varphi_N(\cdot) = \frac{1}{N}\mathbb{E}\mathrm{tr}[\cdot]$ and, for instance,

$$\frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{A}^k\mathbf{D}^l\big] \to \frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{A}^k\big]\frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{D}^l\big], \tag{2.7}$$

$$\frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{A}^k\mathbf{D}^l\mathbf{B}^m\big] \to \frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{A}^k\big]\frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{B}^m\big]\frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{D}^l\big], \tag{2.8}$$

for any $k,l,m \in \mathbb{N}$ and $N \to +\infty$. These joint moments, and more complicated ones, can be derived by applying the freeness rules presented in the previous section.

Reference [77] defines the so-called *asymptotic freeness almost everywhere*, by replacing convergence in expectation of the moments by almost sure convergence. For example, if $\mathbf{A}$ and $\mathbf{B}$ are two matrices asymptotically free almost everywhere, then

$$\lim_{N\to+\infty}\frac{1}{N}\mathrm{tr}\big[\mathbf{A}^k\mathbf{B}^l\big] = \frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{A}^k\big]\frac{1}{N}\mathbb{E}\mathrm{tr}\big[\mathbf{B}^l\big],$$

almost surely and for all $k,l \in \mathbb{N}$. Similarly to Theorem 2.1, there exist classes of random matrices that can be shown to be asymptotically free almost everywhere. However, to the best of the author's knowledge, all published results require more stringent conditions: typically, the distribution of the matrices should be multi-variate Gaussian or, more generally, (bi-)unitarily invariant[10]. We include here the statement of [77, Theorems 4.3.5 and 4.3.11], which will be useful in Chapter 5:

[10] A $N \times N$ random matrix $\mathbf{U}$ is unitarily invariant if its distribution is equal to that of $\mathbf{V}_1\mathbf{U}\mathbf{V}_1$, while is bi-unitarily invariant if its distribution is equal to that of $\mathbf{V}_1\mathbf{U}\mathbf{V}_2$, for any $N \times N$ unitary matrices $\mathbf{V}_1, \mathbf{V}_2$.

Theorem 2.2. *Let $\mathbf{A}(N,s)$ be an independent family of $N \times N$ unitarily invariant Hermitian random matrices with complex entries, indexed by $s \in \mathbb{N}$. Consider also the collection $(\mathbf{D}(N,t))_{t\in\mathbb{N}}$ of $N \times N$ deterministic complex matrices such that $\sup_N \|\mathbf{D}(N,t)\| < +\infty$ for each $t \in \mathbb{N}$ ($\|\cdot\|$ denotes the operator norm) and $(\mathbf{D}(N,t)\mathbf{D}^H(N,t))_{t\in\mathbb{N}}$ has a limit distribution as $N \to +\infty$. If, for all s, $\mathbf{A}(N,s)$ converges in distribution (with respect to $\frac{1}{N}\mathrm{tr}\{\cdot\}$) almost surely to a measure μ_s with real compact support, then the family of sets*

$$\big(\{\mathbf{D}(N,t), \mathbf{D}^H(N,t) : t \in \mathbb{N}\}, \{\mathbf{A}(N,1)\}, \{\mathbf{A}(N,2)\}, \dots\big)$$

is asymptotically free almost everywhere as $N \to +\infty$.

Theorem 2.3. *Let $\mathbf{A}(N,s)$ be an independent family of $N \times N$ bi-unitarily invariant random matrices with complex entries, indexed by $s \in \mathbb{N}$. Consider also the collection $(\mathbf{D}(N,t))_{t\in\mathbb{N}}$ defined in Theorem 2.2. If, for all s, $\mathbf{A}(N,s)^H\mathbf{A}(N,s)$ converges in distribution almost surely to a measure μ_s with positive real compact support, then the family of sets*

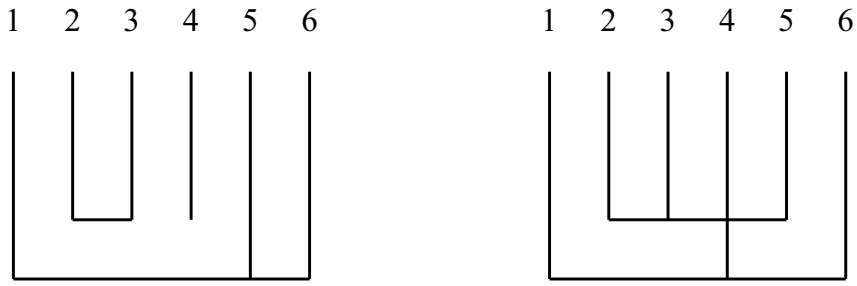
$$\big(\{\mathbf{D}(N,t), \mathbf{D}^H(N,t) : t \in \mathbb{N}\}, \{\mathbf{A}(N,1), \mathbf{A}^H(N,1)\}, \{\mathbf{A}(N,2), \mathbf{A}^H(N,2)\}, \dots\big)$$

is asymptotically free almost everywhere as $N \to +\infty$.

2.3.3 Combinatorics and free probability: non-crossing partitions and free cumulants

The original definition of freeness given in Section 2.3.1 brought us some interesting insights on free random variables. However, its requirements are not easy to verify. What follows is a different but equivalent approach based on combinatorics, due to R. Speicher [75,82]. *Free cumulants* are introduced: they are polynomial expressions of the moments of random variables, much more suitable to characterize freeness than the original moments.

Free cumulants are based on *non-crossing partitions*. Since the formal definition of non-crossing partitions is not very self-explanatory, we will now describe how they are built. For clarity's sake, we will focus on the set of the first n positive integers $\mathcal{I} = \{1, \dots, n\}$: it is straightforward to extend the definition to any finite ordered set. Recall that a partition π of $\mathcal{I}$ is a collection of disjoint subsets, called "blocks", $\mathcal{I}_1, \dots, \mathcal{I}_r \subseteq \mathcal{I}$ such that $\mathcal{I}_1 \cup \dots \cup \mathcal{I}_r = \mathcal{I}$. To verify whether π is a non-crossing partition, we may proceed as follows. First, write the elements of $\mathcal{I}$ in increasing order and trace a vertical line under each one of them. Then, connect with a horizontal line all the elements of the same block. If these arcs can be drawn without crossing one another, then π is a non-crossing partition. Figure 2.4 depicts an example of non-crossing partition as opposed to a generic one. Let $NC(n)$



(a) $\{(1,5,6),(2,3),(4)\}$ (b) $\{(1,4,6),(2,3,5)\}$.

Figure 2.4: Two examples of allowed (a) and forbidden (b) partitions in $NC(6)$.

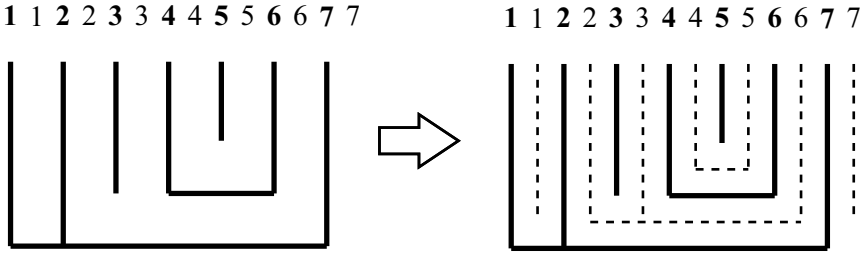


Figure 2.5: Example of non-crossing partition and its complementation map: $\pi = \{(1,2,7),(3),(4,6),(5)\}$ (solid line) and $K(\pi) = \{(1),(2,3,6),(4,5),(7)\}$ (dashed line), respectively.

denote the set of all non-crossing partitions of n elements. It can be shown that its cardinality is given by the n-th Catalan number

$$C_n = \frac{1}{n}\binom{2n}{n-1}.$$

Another concept that will be useful in the following is the complementation map $K(\pi)$ associated to the non-crossing partition π, which can be derived from the graphical representation just introduced, see Figure 2.5. Duplicate the elements of $\mathcal{I}$, alternating the old and the new sets (the new i-th element appears between the i-th and the $(i+1)$-th elements of the old set, while the new n-th element closes the sequence). Connect the new elements as described above, without intersecting lines of neither the new graph nor the old one. The resulting partition is the complementation map of π. The example of Figure 2.6 reports all the $C_4 = 14$ non-crossing partitions, together with their complementation maps, of $NC(4)$.

Finally, to simplify equations below, let us introduce a notation for the factorization of a moment according to a non-crossing partition. Let $a_1, \ldots, a_n$ be non-commutative random variables in the probability space $(\mathcal{A}, \varphi)$. Given a particular non-crossing partition $\pi = \{\mathcal{I}_1, \ldots, \mathcal{I}_r\}$, we will

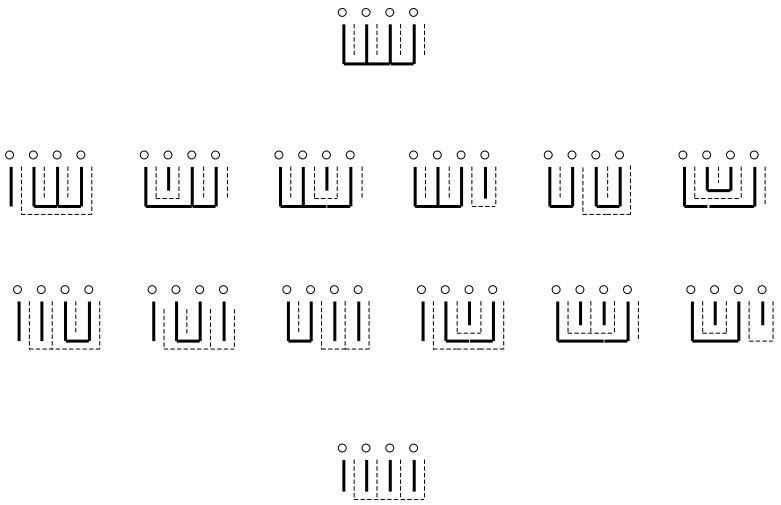


Figure 2.6: Non-crossing partitions (solid lines) and relative complementation maps (dashed line) of a four-element set.

write:

$$\varphi_\pi[a_1,\ldots,a_n] \stackrel{\text{def}}{=} \varphi\left(\prod_{i\in\mathcal{I}_1} a_i\right)\cdots\varphi\left(\prod_{i\in\mathcal{I}_r} a_i\right), \tag{2.9}$$

where the multiplications on the right-hand side are made respecting the increasing order of the indeces.

Free cumulants

As mentioned before, free cumulants (as classical cumulants) consist of polynomial expressions of moments of random variables. More specifically, they are built according to the following recursive definition.

Definition 2.7. Let $a_1,\ldots,a_n$ be non-commutative random variables in the probability space $(\mathcal{A},\phi)$. The *joint free cumulant* of $a_1,\ldots,a_n$, denoted by $k_n(a_1,\ldots,a_n)$, is the solution to the system of equations

$$\varphi(a_1\cdots a_n) = \sum_{\pi\in NC(n)} k_\pi[a_1,\ldots,a_n]. \tag{2.10}$$

In the last equation, we introduced the quantity $k_\pi[a_1,\ldots,a_n]$ which is a short notation for the multiplication of all the joint free cumulants corresponding to the blocks of the non-crossing partition π. Specifically, if $\pi=\{\mathcal{I}_1,\ldots,\mathcal{I}_r\}$ is the given partition and $\mathcal{I}_i=\{a_{i,1},\ldots,a_{i,|\mathcal{I}_i|}\}$ its i-th block, then

$$k_\pi[a_1,\ldots,a_n] \stackrel{\text{def}}{=} \prod_{i=1}^{r} k_{|\mathcal{I}_i|}(a_{i,1},\ldots,a_{i,|\mathcal{I}_i|}), \tag{2.11}$$

where, as before, $k_{|\mathcal{I}_i|}(a_{i,1},\ldots,a_{i,|\mathcal{I}_i|})$ is the joint free cumulant of the variables $a_{i,1},\ldots,a_{i,|\mathcal{I}_i|}$. This factorization is analogous to what was done in (2.9) for the joint moments.

Note that (2.10) can be applied recursively to compute the cumulants $k_n(\cdot)$ for all possible $n\in\mathbb{N}$, as shown below for $n=1,2$ and 3:

$n=1$**.** In this case one simply has $\varphi(a_1)=k_{\{(1)\}}(a_1)$, implying

$$k_1(a_1)=\varphi(a_1);$$

$n=2$**.** The $C_2=2$ possible non-crossing partitions with two elements are $\{(1,2)\}$ and $\{(1),(2)\}$. Then, from (2.10) and (2.11),

$$\varphi(a_1,a_2)=k_{\{(1,2)\}}[a_1,a_2]+k_{\{(1),(2)\}}[a_1,a_2]=k_2(a_1,a_2)+k_1(a_1)k_1(a_2).$$

Comparing to the previous result, $k_2(\cdot)$ is given by

$$k_2(a_1,a_2)=\varphi(a_1,a_2)-\varphi(a_1)\varphi(a_2);$$

$n=3$**.** The set $\mathcal{I}=\{1,2,3\}$ presents $C_3=5$ possible non-crossing partitions, namely $\{(1,2,3)\},\{(1,2),(3)\},\{(1),(2,3)\},\{(1,3),(2)\}$ and $\{(1),(2),(3)\}$. Hence, the recurrence relation (2.10) is written

$$\begin{aligned}\varphi(a_1a_2a_3)&=k_{\{(1,2,3)\}}[a_1,a_2,a_3]+k_{\{(1,2),(3)\}}[a_1,a_2,a_3]+\\&\quad+k_{\{(1),(2,3)\}}[a_1,a_2,a_3]+k_{\{(1,3),(2)\}}[a_1,a_2,a_3]+\\&\quad+k_{\{(1),(2),(3)\}}[a_1,a_2,a_3]\\&=k_3(a_1,a_2,a_3)+k_2(a_1,a_2)k_1(a_3)+k_1(a_1)k_2(a_2,a_3)+\\&\quad+k_2(a_1,a_3)k_1(a_2)+k_1(a_1)k_1(a_2)k_1(a_3).\end{aligned}$$

After some algebra and recalling the two previous results, we have

$$\begin{aligned}k_3(a_1,a_2,a_3)=\varphi(a_1a_2a_3)&-\varphi(a_1a_2)\varphi(a_3)-\varphi(a_1)\varphi(a_2a_3)-\\&-\varphi(a_1a_3)\varphi(a_2)+2\varphi(a_1)\varphi(a_2)\varphi(a_3).\end{aligned}$$

As mentioned above, the introduction of free cumulants is motivated by strong interconnection with freeness. Indeed, as classic independence results in a straightforward factorization of joint moments, free independence may be re-defined very simply in terms of free cumulants.

Theorem 2.4 ([82, Theorem 4.2.1])**.** *Let $(\mathcal{A},\varphi)$ be a non-commutative probability space and consider the unital subalgebras $\mathcal{A}_1,\ldots,\mathcal{A}_m\subset\mathcal{A}$. Then, the two following statements are equivalent:*

1. *$\mathcal{A}_1,\ldots,\mathcal{A}_m$ are free and*

2. the mixed joint free cumulants vanish, i.e.

$$k_n(a_1, \ldots, a_n) = 0$$

for any $n \geq 2$ and any collection of random variables $a_1 \in \mathcal{A}_{j(1)}, \ldots, a_n \in \mathcal{A}_{j(n)}$ chosen from at least two different subalgebras (i.e. $j(1) = j(2) = \cdots = j(n)$ is the only forbidden case).

The proof of this theorem is out of the scope of this dissertation, but it practically consists in showing, thanks to (2.10), that the conditions on the moments in Definition 2.5 translate into the vanishing condition on the cumulants given in the theorem. Observe that here we do not require the random variables to be zero mean, nor consecutive random variables to belong to two different subalgebras. These "weaker" assumptions make the characterization of freeness in terms of cumulants much more useful in practice.

The relation between free cumulants and free independence goes much further than Theorem 2.4. However, the purpose here was just to offer some better understanding of free probability. Once again, the interested reader is referred to [75, 82]. For completeness, we just report a result were the complementation map $K(\pi)$ of a non-crossing partition π comes into play. Let $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$ be two sets of free non-commutative random variables. The, we have

$$\varphi(a_1 b_1 a_2 b_2 \ldots a_n b_n) = \sum_{\pi \in NC(n)} k_\pi[a_1, a_2, \ldots, a_n] \varphi_{K(\pi)}[b_1, b_2, \ldots, b_n],$$

where $\varphi_\pi[\cdot]$ and $k_\pi[\cdot]$ are defined in (2.9) and (2.11), respectively.

2.3.4 Additive and multiplicative free convolutions of measures

In classical probability theory, independent random variables are interesting objects. Indeed, their joint distribution and moments can be factorized in terms of the individual ones. Furthermore, simple formulas exist to compute the distribution of their sums and products. Similarly, in a non-commutative probability space, the previous sections showed that the joint moments of free random variables can derived from the individual moments. To conclude the analogy, it still remains to give practical ways to compute the distributions of sums and products of free random variables. These tools represent the solution to our problem of characterizing the spectra of sums and products of large random matrices, whose asymptotic interpretation as free non-commutative random variables as been given in Section 2.3.2.

Additive free convolution

Take two free random variables a and b in the non-commutative space $(\mathcal{A}, \varphi)$ and denote their distribution with μ_a and μ_b. Recall that, in this context, a distribution does not need to be a function but just a collection of moments. The new random variable $a + b$ also belongs to $(\mathcal{A}, \varphi)$ and has a proper distribution. Indeed, since φ is linear and $\{a, b\}$ are free, previous results imply that the moments of $a+b$, i.e. $\varphi\Big((a+b)^n\Big)$, can be expressed in terms of the singular moments $\varphi(a^s)$ and $\varphi(b^t)$, for $n, s, t \in \mathbb{N}$. In short, we say that the distribution μ_{a+b} of $a+b$ results from the *additive free convolution* of μ_a and μ_b and we write

$$\mu_{a+b} = \mu_a \boxplus \mu_b. \tag{2.12}$$

The name "convolution" is imported from classic probability theory, where the density of the sum of two independent random variables can be expressed as the convolution of their individual densities. In free probability, however, this operation is highly non-linear, as it will be evident later on.

Observe that (2.12) is only the definition of the operator $\boxplus$ and lacks operative value. In practice, additive free convolution is carried out by means of the so-called "R-transform", which translates the additive convolution of two distribution into the sum of their transformations. In classic probability theory, a similar result is obtained by considering the logarithm of the Fourier transform of the probability density function.

R-transform. Consider the random variable a with distribution μ_a and k-th order moment $m_k = \varphi(a^k)$. The Stieltjes transform of μ_a is the formal power series

$$m_{\mu_a}(z) = z^{-1} + \sum_{k=1}^{+\infty} m_k z^{-1-k} \in \mathbb{C}[[z^{-1}]]. \tag{2.13}$$

Observe that this definition of the Stieltjes transform coincides with Definition 2.2 when the distribution μ_a can be expressed as a real function and has compact support. The R-transform of the distribution μ_a is defined as

$$R_{\mu_a}(z) = K_{\mu_a}(z) - z^{-1}, \tag{2.14}$$

where $K_{\mu_a}(z)$ is the formal inverse of $m_{\mu_a}(z)$, i.e.

$$m_{\mu_a}(K_{\mu_a}(z)) = K_{\mu_a}(m_{\mu_a}(z)) = z. \tag{2.15}$$

It certainly exists since the constant term of $m_{\mu_a}(z)$ is zero.

As mentioned above, we introduce the R-transform as a tool to characterize the distribution of the sum of free random variables. More specifically, if a

and b are two free non-commutative random variables, then the R-transform of μ_{a+b} is the sum of the R-transforms of μ_a and μ_b or, equivalently,

$$R_{\mu_{a+b}}(z) = R_{\mu_a \boxplus \mu_b}(z) = R_{\mu_a}(z) + R_{\mu_b}(z). \tag{2.16}$$

This result shows, on the one hand, that the additive convolution $\boxplus$ is a commutative operator. More importantly, it also represents an operative approach for computing the distribution of the sum of random variables. Indeed, knowing the R-transform of a distribution, its Stieltjes transform can be computed thanks to (2.15) and (2.14); then, the Stieltjes inversion formula can be applied.

The equivalence between additive convolution and sum of R-transforms can be easily proven after realizing that the R-transform can be written as the formal power series

$$R_{\mu_a}(z) = \sum_{n=0}^{+\infty} k_{n+1}^a z^n,$$

where $k_n^a = k_n(a, \ldots, a)$ is the n-th free cumulant of a as defined in the previous section. Now, if a and b are two different free random variables, Theorem 2.4 implies

$$k_n^{a+b} = k_n(a+b, \ldots, a+b) = k_n(a, \ldots, a) + k_n(b, \ldots, b) = k_n^a + k_n^b.$$

The result in (2.16) follows straightforwardly.

Example 2.10. Let ν be a distribution with density $f_\nu(x) = \frac{1}{2}[\delta_{-1}(x) + \delta_1(x)]$, where $\delta_t(x)$ is the Dirac delta centered in $x = t$. Then, it can be shown that its R-transform is $R_\nu(z) = (\sqrt{1+4z^2} - 1)/(2z)$. Furthermore, by inverting $2R_\nu(z)$, we can compute the density of the distribution $s = \nu \boxplus \nu$, namely

$$f_s(x) = \frac{1}{\pi\sqrt{4 - x^2}}, \text{ for } |x| \leq 2 \text{ and 0 otherwise.}$$

This is called the arcsin law.

This example is very representative of the non linearity of the additive free convolution: "summing" to itself a discrete distribution we end up with an absolutely continuous density with measurable support.

To conclude, it is worth introducing the following results, even though they are not strictly related to the contents of this thesis. Briefly, in free probability, the semicircular law (see Example 2.3) takes the role assumed by the Gaussian distribution in classic probability theory.

Example 2.11. Let denote with $\gamma_{a,r}$ the semicircle distribution of mean a and variance $r^2/4$, whose density is

$$f_{\gamma_{a,r}}(x) = \frac{2}{\pi r^2}\sqrt{r^2 - (x-a)^2}, \text{ for } |x - a| \leq r \text{ and 0 otherwise.}$$

Since its R-transform is

$$R_{\gamma_{a,r}}(z) = a + \frac{r^2}{4}z,$$

it is straightforward that

$$\gamma_{a,r} \boxplus \gamma_{b,s} = \gamma_{a+b,\sqrt{r^2+s^2}},$$

for any two free random variables with distribution $\gamma_{a,r}$ and $\gamma_{b,s}$, respectively.

This example shows that the sum of semicircular-distributed free random variables is also semicircular-distributed, similarly to what happens when adding Gaussian independent random variables. The following results brings this analogy further.

Theorem 2.5 (Free central limit theorem)**.** *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $a_1, a_2, \cdots \in \mathcal{A}$ a sequence of free and identically distributed random variables with zero mean and variance σ^2, i.e. $\varphi(a_i) = 0$ and $\varphi(a_i^2) = \sigma^2$ for all i. Then, the random variable S_N, defined as*

$$S_N = \frac{a_1 + \cdots + a_N}{\sqrt{N}},$$

tends as $N \to +\infty$ to a random variable with distribution a semicircle law with zero mean and variance σ^2.

Proof. A very simple proof can be given in terms of the R-transform. Recall that, in non-commutative probability, convergence in distribution is equivalent to convergence of the moments and, hence, of the free cumulants. Also recall that the free cumulants k_n are the coefficients of the R-transform expressed as a formal power series. Then, since $R_{\lambda a}(z) = \lambda R_a(\lambda z)$, one obtains

$$\begin{aligned} R_{S_N}(z) = R_{(a_1+\cdots+a_N)/\sqrt{N}}(z) = \\ = \frac{1}{\sqrt{N}} R_{a_1+\cdots+a_N}\Big(\frac{z}{\sqrt{N}}\Big) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} R_{a_i}\Big(\frac{z}{\sqrt{N}}\Big) = \\ = \sqrt{N} R_{a_i}\Big(\frac{z}{\sqrt{N}}\Big) = \sqrt{N}\Big(k_1 + \frac{z}{\sqrt{N}}k_2 + \frac{z^2}{N}k_3 + \dots\Big), \end{aligned}$$

where the third equality follows from freeness and the fourth one from the fact that the a_i's are identically distributed. Now, since $k_1 = \varphi(a_i) = 0$ and $k_2 = \varphi(a_1^2) - \varphi^2(a_i) = \sigma^2$, one has

$$R_{S_N}(z) \overset{N\to+\infty}{\longrightarrow} z\sigma^2,$$

which is the R-transform of a centered semicircle law with variance σ^2. □

This theorem is the equivalent in free probability of the classic central limit theorem (see, e.g., [59, Section 27]), which states the convergence to a Gaussian random variable of a sum of an infinite number of independent and identically distributed random variables. We can hence conclude that the semicircle law is the true analog of the Gaussian distribution in non-commutative probability theory.

Multiplicative free convolution

Similarly to what has just been done for the additive convolution, the *multiplicative free convolution* is defined as the binary operator that produces the distribution of the product of two free random variables a and b from their individual distributions μ_a and μ_b, respectively. Denoting by $\boxtimes$ such an operator, we write

$$\mu_{ab} = \mu_a \boxtimes \mu_b.$$

This definition is justified by the fact that the moments of ab are polynomial functions of the moments of a and b alone.

Next, the S-transform is introduced. This is the main tool to compute the result of a multiplicative convolution in practice: as with the Mellin transform in classic probability theory, the S-transform of μ_{ab} is simply the product between the S-transforms of μ_a and μ_b. Observe that this result implies the commutativity of the operator $\boxtimes$.

S-transform Let a be a non-commutative random variable with distribution μ_a and moments $m_k, k = 1, 2, \ldots$ Consider the formal power series

$$\psi_{\mu_a}(z) = \sum_{k=1}^{+\infty} m_k z^k = z^{-1} m_{\mu_a}(z^-1) - 1 \in \mathbb{C}[[z^{-1}]],$$

where $m_{\mu_a}(z)$ is the Stieltjes transform of μ_a as defined in (2.13). Since $\psi_{\mu_a}(z)$ has no constant term, its formal inverse $\chi_{\mu_a}(z)$ exists and is well-defined:

$$\chi_{\mu_a}(\psi_{\mu_a}(z)) = \psi_{\mu_a}(\chi_{\mu_a}(z)) = z.$$

The S-transform of the distribution μ_a is then defined as

$$S_{\mu_a}(z) = \frac{1+z}{z}\chi_{\mu_a}(z).$$

As mentioned before, the main property of the S-transform is to translate the multiplicative free convolution into a trivial multiplication, that is

$$S_{\mu_{ab}}(z) = S_{\mu_a \boxtimes \mu_b}(z) = S_{\mu_a}(z) S_{\mu_b}(z).$$

2.4 Conclusions

This chapter tried to give a reasonable introduction to three important mathematical topics, namely the convergence of sequences of random variables, random matrix theory and free probability. Particular attention was given to the last two subjects, whose results are not widely known. The main idea was to show that some classes of large random matrices are "nice" mathematical objects since (i) their spectra present a deterministic characterization and (ii) eigenvalues of sums and products of matrices can be easily computed.

Most of the results will be useful to solve technical problems in the following chapters. Few others, as the free central limit theorem, will not find direct application. However, they have been mentioned to give a more general and complete overview of the fascinating theories of random matrices and free probability.

Chapter 3

Randomized i.i.d. LD-STBC for the Relay Channel: the Proposed AF Scheme and its Spectral Efficiency

The ultimate aim of every communications system is to transfer information from one point (or multiple points) of the space to another location (or other locations), where that information is demanded. Consider a simple point-to-point system as the one depicted in Figure 3.1. It is desirable that a user having access to the output signal Y be able to reconstruct the information described by the input signal X. Then, efforts should be made to maximize the mutual information $I(X;Y)$, that is to minimize the uncertainty about X knowing Y. Obviously, communications costs like, for instance, power consumption and allocated resources must be taken into account. From a thorough analysis of communications systems, it can be shown that $I(X;Y)$ depends both on some system-related quantities (e.g., the channel quality or the amount of noise) and on the model used to generate the signal X. The channel capacity is hence defined as

$$C = \max_{p_X(x)} I(X;Y),$$

where $p_X(x)$ is the distribution of the random process that models the input signal X.

The channel capacity is hence the theoretical maximum transmission rate that can be achieved on the channel. According to this definition, the capacity of the general relay channel is still unknown. As mentioned in Chapter 1, there exist results for some special cases (e.g. [14, 35, 40]), but we still do not fully understand which rates may be achieved by introducing relays to aid point-to-point communications.

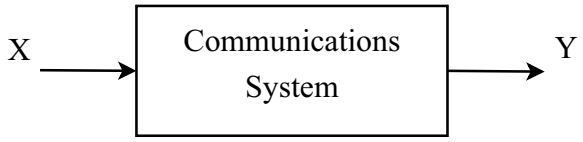


Figure 3.1: Block diagram of a basic point-to-point communications system.

This chapter does not pretend to solve that issue. Instead, we will introduce a specific relaying architecture and compute its spectral efficiency. Similarly to the channel capacity, the spectral efficiency is also defined as the maximum achievable mutual information (normalized with respect to the occupied resources) between input and output system signals. However, it takes into account implementation details. More specifically, we will consider different coding rates at the relays and different receiver schemes at the destination.

3.1 System description

The system analyzed in this dissertation is the classical relay channel as presented, for instance, in [24, 29, 39, 55, 83]: a source (or transmitter) communicates with a destination supported by a set of half-duplex relays. In particular, this chapter focuses on the amplify-and-forward (AF) relaying strategy, which is more general. Indeed, the decode-and-forward (DF) results presented in the following chapter can be easily derived from the AF ones by, for instance, considering unitary unnoisy source–relay channels.

As a result of the half-duplex assumption, communications occur in two phases or time-slots, depicted in Figure 3.2:

1. the source broadcasts a message to the relays and the destination;

2. the relays linearly transform the received noisy samples and forward them to the destination, while the source remains silent[1].

The destination will combine the signals received in both time-slots and estimate the original information. Two different receivers will be presented and compared in terms of spectral efficiency: the optimal maximum-likelihood (ML) receiver and the linear minimum-mean-square-error (LMMSE) filter. The system is also compared to the direct link in order to identify under which conditions relaying is superior to conventional transmission.

The main difference with respect to previous results is the signal processing at the relays. As explained in Section 1.3, the main idea is to introduce a distributed randomized linear-dispersion space-time block code (LD-STBC)

[1]This two-phase protocol without source transmitting in phase two is often referred to as orthogonal relaying.

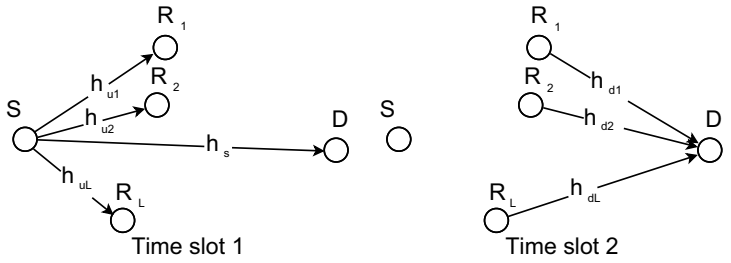


Figure 3.2: The two-phase transmission scheme.

which allows exploiting diversity while being flexible and suitable for a dynamic system where the number of active relays continuously varies in time. More specifically, the l-th relay ($l = 1, \ldots, L$, with L the total number of relays) is assigned its particular $N \times K$ linear-dispersion matrix $\mathbf{C}_l$ which linearly transforms the K-symbol source vector $\mathbf{s}$ into N new symbols to be forwarded to the destination. We will denote by $\alpha = K/N$ the coding rate, which also coincides with the ratio between the lengths of the two phases (the bandwidth of all links is kept constant). Intuitively, the system spectral efficiency depends on α, since this parameter tunes the weight of source and relay contributions at the destination.

Observe that each column of $\mathbf{C}_l$ spreads the relative element of $\mathbf{s}$ over the following N relay channel accesses, opening to a DS/CDMA[2] interpretation of the coding scheme. For this reason, we will often refer to the coding matrices as spreading matrices and to their columns as spreading sequences or signatures.

3.1.1 Signal model

In time-slot 1, the source transmits the vector of K complex i.i.d. symbols $\mathbf{s} = [s_1, s_2, \ldots, s_K]^T$, each one with zero mean and variance $\mathbb{E}[|s_k|^2] = P_s$. The l-th relay gets a noisy copy through the uplink channel h_{ul}: $\mathbf{r}_l = h_{ul}\mathbf{s} + \mathbf{n}_{ul}$, with $\mathbf{n}_{ul}$ the additive noise. Similarly, being h_s the direct source–destination channel and $\mathbf{n}_1$ the relative noise, the destination receives $\mathbf{d}_1 = h_s\mathbf{s} + \mathbf{n}_1$. In time-slot 2, each relay multiplies the received samples by the complex amplification factor g_l and by the $N \times K$ signature matrix $\mathbf{C}_l$; the N resulting symbols are then sent to the destination through the downlink channel h_{dl}. Thus, the signal received in the second phase can be expressed as: $\mathbf{d}_2 = \sum_{l=1}^{L} g_l h_{dl} \mathbf{C}_l \mathbf{r}_l + \mathbf{n}_d$, where $\mathbf{n}_d$ denotes the additive noise at the destination. Replacing $\mathbf{r}_l$ by its expression, the global signal model is

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} h_s \mathbf{I}_K \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{H}}_u \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{n}_1 \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{\Psi}}\mathbf{n}_u + \mathbf{n}_d \end{bmatrix}; \tag{3.1}$$

[2]Direct-sequence code division multiple access.

where:

- $\widetilde{\mathbf{C}} = [\mathbf{C}_1\, \mathbf{C}_2\, \cdots\, \mathbf{C}_L]$ collects all the signature matrices;
- $\widetilde{\mathbf{\Psi}} = \mathbf{\Psi} \otimes \mathbf{I}_K$. $\mathbf{\Psi} = \mathrm{diag}\{g_1 h_{d1}, g_2 h_{d2}, \ldots, g_L h_{dL}\}$ is a diagonal matrix where the l-th entry $g_l h_{dl}$ is the equivalent downlink channel coefficient for relay l;
- $\widetilde{\mathbf{H}}_u = \mathbf{h}_u \otimes \mathbf{I}_K$, with $\mathbf{h}_u = [h_{u1}\, h_{u2}\, \cdots\, h_{uL}]^T$;
- $\mathbf{n}_u = [\mathbf{n}_{u1}^T\, \mathbf{n}_{u2}^T\, \cdots\, \mathbf{n}_{uL}^T]^T$.

All the channels are assumed to be independent and affected by frequency-flat, quasi-static fading, i.e. channel coefficients remain constant over all the message transmission. All the noises are modeled as additive white Gaussian noise (AWGN), with $\mathbf{n}_1 \sim \mathcal{CN}(\mathbf{0}, \sigma_d^2 \mathbf{I}_K)$, $\mathbf{n}_u \sim \mathcal{CN}(\mathbf{0}, \sigma_u^2 \mathbf{I}_{KL})$ and $\mathbf{n}_d \sim \mathcal{CN}(\mathbf{0}, \sigma_d^2 \mathbf{I}_N)$, independent of one another. We will work under the hypothesis of global perfect channel state information at the receiver: the destination knows (in magnitude and phase) the channel coefficients of all the source–relay, relay–destination and source–destination links, besides the noise variances. On the contrary, relays and source do not have any channel state information (one can assume the existence of an error-free feedback channel to inform the source about the maximum supported rate).

The coding matrices

The N-chip-long sequences $\mathbf{c}_{l,k}$, columns of the matrices $\mathbf{C}_l$, and hence of the matrix $\widetilde{\mathbf{C}}$, are composed by i.i.d. random elements with zero mean, variance $1/N$ and bounded moments (the exact distribution is not important). Recall that randomness only models the sequence generation: after system initialization, the matrices $\mathbf{C}_l$ remain fixed and are known at the receiver. Moreover, observe that each matrix $\mathbf{C}_l$ is a block of an LD-STBC [84].

To justify further this choice for the spreading sequences, let us briefly analyze two special cases: (i), $\mathbf{C}_l = \mathbf{I}_K, \forall l$, i.e. relays do not apply any code and, (ii), $\widetilde{\mathbf{C}}$ is a $KL \times KL$ orthogonal matrix. For the signal model in (3.1), it is straightforward to show that the spectral efficiency (in nat/s/Hz) in the two cases will be, respectively:

$$I_{(i)} = \frac{1}{2} \ln \left(1 + \frac{P_s}{\sigma_d^2} |h_s|^2 + \frac{P_s \left| \sum_{l=1}^{L} g_l h_{dl} h_{ul} \right|^2}{\sigma_u^2 \sum_{l=1}^{L} |g_l h_{dl}|^2 + \sigma_d^2} \right), \tag{3.2}$$

$$I_{(ii)} = \frac{1}{1+L} \ln \left(1 + \frac{P_s}{\sigma_d^2} |h_s|^2 + \sum_{l=1}^{L} \frac{P_s |g_l h_{dl} h_{ul}|^2}{\sigma_u^2 |g_l h_{dl}|^2 + \sigma_d^2} \right). \tag{3.3}$$

In the first case, the channel is not properly accessed: the numerator inside the logarithm of equation (3.2) shows how cooperation can be potentially useless. Indeed, contributions can generally combine destructively,

unless the relays have perfect knowledge of both uplink and downlink channel phases. In the literature [24, 29, 39, 83], it is common to assume that relays simply fix the power gain $|g_l|^2$ to meet some power constraint.

On the contrary, with orthogonal codes (3.3), relay contributions always sum up constructively. However, as the number of relays increases, the orthogonality constraint drastically reduces the spectral efficiency due to the presence of the factor $1/(1+L)$, fraction of degrees of freedom effectively employed. We will see that our randomly generated LD-STBC results in a coherent sum of relay contributions while allowing us to limit the waste of resources.

The noise

For the sake of clarity, let us summarize the characteristics of the final noise affecting the signal at the receiver, which can be expressed as

$$\mathbf{n} = \begin{bmatrix} \mathbf{n}_1 \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{\Psi}}\mathbf{n}_u + \mathbf{n}_d \end{bmatrix}.$$

This noise, even if it is still a Gaussian process with zero mean, is not white anymore due to a colored component introduced by the correlation among relay signatures. Indeed, the covariance matrix can be expressed as

$$\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \begin{bmatrix} \sigma_d^2\mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix};$$
$$\text{with } \mathbf{R} = \sigma_u^2\widetilde{\mathbf{C}}\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{C}}^H + \sigma_d^2\mathbf{I}_N.$$

3.2 The maximum-likelihood receiver

This section deals with the assumption of maximum-likelihood reception at the destination, which maximizes the spectral efficiency. Generally, the ML receiver needs to be implemented as an exhaustive search among all the possible source messages. However, the linear-dispersion structure of the considered codes allows for more efficient decoding techniques (e.g. sphere decoding [85]) as explained in [84].

For the MIMO-like signal model in (3.1), it is well known that the maximum spectral efficiency is theoretically achieved by random Gaussian source coding and may be expressed in nat/s/Hz by the formula

$$I_{\mathrm{AF,ML}} = \frac{1}{N+K}\ln\det\left(\mathbf{I}_K + P_s\begin{bmatrix} h_s\mathbf{I}_K \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{H}}_u \end{bmatrix}^H \begin{bmatrix} \sigma_d^2\mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} h_s\mathbf{I}_K \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{H}}_u \end{bmatrix}\right), \tag{3.4}$$

(see, e.g., [10]). A direct analysis of this expression is not straightforward, due to its matrix formulation and to the dependence on the particular realization of the linear-dispersion matrices. Simplifications can be sought in

the asymptotic domain: letting the number of source symbols K and the spreading-sequence length N grow without bound but with constant coding rate $\alpha = K/N$ ($0 < \alpha < +\infty$), one can show that $|I_{\mathrm{ML}} - I^{\mathrm{iid}}_{\mathrm{AF,ML}}| \to 0$ almost surely (with probability 1), where

$$I^{\mathrm{iid}}_{\mathrm{AF,ML}} = \frac{1}{1+\alpha}\left[\alpha \ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) + \right. \\ \left. + \alpha \sum_{l=1}^{L} \ln \frac{1+\lambda_l\phi_1}{1+|g_l h_{dl}|^2\phi_2} + \ln\frac{\phi_2}{\phi_1} + \frac{\sigma_d^2}{\sigma_u^2}(\phi_1-\phi_2)\right] \quad (3.5)$$

is the deterministic asymptotic equivalent of the spectral efficiency (the proof is given in the appendix). The following quantities have been defined:

- $\{\lambda_l\}_{l=1}^{L}$ are the eigenvalues of $\mathbf{\Psi}\mathbf{\Psi}^H + \frac{P_s/\sigma_u^2}{1+P_s|h_s|^2/\sigma_d^2}\mathbf{\Psi}\mathbf{h}_u\mathbf{h}_u^H\mathbf{\Psi}^H$;
- ϕ_1 is the unique positive solution to

$$\phi_1 = \left(\frac{\sigma_d^2}{\sigma_u^2} + \alpha\sum_{l=1}^{L}\frac{\lambda_l}{1+\lambda_l\phi_1}\right)^{-1}; \quad (3.6)$$

- ϕ_2 is the unique positive solution to

$$\phi_2 = \left(\frac{\sigma_d^2}{\sigma_u^2} + \alpha\sum_{l=1}^{L}\frac{|g_l h_{dl}|^2}{1+|g_l h_{dl}|^2\phi_2}\right)^{-1}. \quad (3.7)$$

Asymptotic equivalents are very useful in practice, since they turn out to be good approximations of the finite reality. Observe that there is no direct reference to K, N or to the codes in the expression of the asymptotic spectral efficiency, which depends only on the coding rate α and on the channel coefficients.

The coefficients ϕ_1 and ϕ_2 do not have an evident physical meaning. However, they are related to the interference generated by the relays when they receive useful signal plus noise (ϕ_1, see also the comments on equation (3.12) in Section 3.3) or only noise (ϕ_2): indeed, ϕ_1 converges to ϕ_2 when $P_s \to 0$.

Unfortunately, equation (3.5) does not clarify whether relay and interference contributions sum up constructively or destructively with the direct one. Indeed, nothing can be said about the aggregate sign of the last three terms. On the one hand, it can be easily shown that $\phi_1 \leq \phi_2$: it is enough to notice that the unique positive solution to

$$\phi = \left(z + \alpha\sum_{l=1}^{L}\frac{a_l}{1+a_l\phi}\right)^{-1} \quad (3.8)$$

is decreasing in any of the parameters a_l (if, without loss of generality, we let $|g_1 h_{d1}|^2 \leq |g_2 h_{d2}|^2 \leq \cdots \leq |g_L h_{dL}|^2$ and $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L$, one has $|g_l h_{dl}|^2 \leq \lambda_l$, see Appendix 3.A.3). On the other hand, nothing can be said when comparing $\lambda_l \phi_1$ and $|g_l h_{dl}|^2 \phi_2$. The problem of whether relays are effectively helpful in increasing the link reliability is addressed in the next section.

3.2.1 Comparison with the direct link

The introduction of relays may potentially lead to some drawbacks. Indeed, the available degrees of freedom are not properly exploited since only K symbols are transmitted during $K + N$ channel accesses. For this reason, one should evaluate whether the diversity benefits introduced by the relays compensate for the losses just discussed. For this purpose, we will now compare the spectral efficiency (3.5) with the one of the direct link, namely $I_s = \ln(1 + P_s |h_s|^2 / \sigma_d^2)$.

Intuition suggests that the direct link channel gain $|h_s|^2$ would be the determining parameter: whenever this channel is good, the receiver can detect the information correctly and there is no need for relaying. The following results confirm this idea.

The horizontal asymptote

For an exhaustive study of the problem, one should identify which conditions must be satisfied for $\hat{I} = I^{\text{iid}}_{\text{AF,ML}} - I_s$ to be positive. Unfortunately, a mathematical analysis of $\hat{I}$ is not a simple task, mainly because of the involved expressions of ϕ_1 and ϕ_2 and their dependence on all the other parameters. Simpler expressions can be found for $\alpha \to \infty$. In this regime, indeed, it is straightforward to show that $I^{\text{iid}}_{\text{AF,ML}}$ tends to I_s: the larger α, the longer time-slot 1 will be with respect to time-slot 2, i.e. relays will transmit for a very short time compared to the source. Studying how $I^{\text{iid}}_{\text{AF,ML}}$ approaches the horizontal asymptote I_s (independent of α) leads to a sufficient condition for the superiority of relaying: if $\hat{I}$ tends to zero from above, continuity implies that there exists a finite α for which $I^{\text{iid}}_{\text{AF,ML}} > I_s$ and, thus, for which relaying is convenient.

From (3.8), one can show that $\phi = \left(\alpha \sum_{l=1}^{L} a_l\right)^{-1} + o\big(\alpha^{-1}\big)$ for $\alpha \to \infty$ and hence

$$\phi_1 = \frac{1}{\alpha \sum_{l=1}^{L} \lambda_l} + o\big(\alpha^{-1}\big), \qquad \phi_2 = \frac{1}{\alpha \sum_{l=1}^{L} |g_l h_{dl}|^2} + o\big(\alpha^{-1}\big),$$

which implies

$$\lim_{\alpha \to \infty} \alpha \hat{I}_1 = -\ln\left(1 + \frac{P_s}{\sigma_d^2} |h_s|^2\right) + \ln \frac{\sum_{l=1}^{L} \lambda_l}{\sum_{l=1}^{L} |g_l h_{dl}|^2}.$$

This means that relaying is convenient if

$$\frac{\sum_{l=1}^{L}\lambda_l}{\sum_{l=1}^{L}|g_l h_{dl}|^2} > 1+\frac{P_s}{\sigma_d^2}|h_s|^2$$

or, equivalently, if

$$\frac{P_s}{\sigma_u^2}\sum_{l=1}^{L}\frac{|h_{ul}|^2}{1+\left(\sum_{m\neq l}|g_m h_{dm}|^2\right)/|g_l h_{dl}|^2} > \left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right)\frac{P_s}{\sigma_d^2}|h_s|^2. \quad (3.9)$$

The last expression has been derived using the fact that (see Appendix 3.A.3)

$$\sum_{l=1}^{L}\lambda_l = \sum_{l=1}^{L}|g_l h_{dl}|^2\left(1+\frac{\frac{P_s}{\sigma_u^2}|h_{ul}|^2}{1+\frac{P_s}{\sigma_d^2}|h_s|^2}\right). \quad (3.10)$$

Observe that the left-hand side of (3.9) is a weighted sum of the relay uplink signal-to-noise ratios (SNR's), namely $P_s|h_{ul}|^2/\sigma_u^2$; the weight associated with the l-th relay increases as its downlink equivalent gain $|g_l h_{dl}|^2$ becomes larger with respect to the other relays. The right-hand side, instead, is a quadratic function of the direct-link SNR. As expected, whenever the direct channel is bad with respect to those of the relays, the proposed scheme performs better than a point-to-point transmission. Note that condition (3.9) is only sufficient and nothing can be said about its necessity ($I_{\mathrm{AF,ML}}^{\mathrm{iid}}$ can theoretically approach the asymptote from below even if there exist a finite α for which $I_{\mathrm{AF,ML}}^{\mathrm{iid}} > I_s$).

For the cases where relaying is convenient, it would be interesting to establish which is the best time-sharing strategy, i.e. the value of the coding rate α that maximizes the spectral efficiency. Once again, the complexity of (3.5) makes this issue particularly involved. In the next section we will look for some more insights in the one-relay case.

The one-relay case

When considering a single relay, the asymptotic expressions simplify considerably. First of all, neglecting all trivial subscripts, one can write

$$\lambda = \left(\frac{P_s|h_u|^2/\sigma_u^2}{1+P_s|h_s|^2/\sigma_d^2}+1\right)|gh_d|^2,$$

and ϕ_1, ϕ_2 are the positive solutions to second order equations, i.e. they take the form

$$\phi = \frac{1}{2az}\left[a-\alpha a-z+\sqrt{(a-\alpha a-z)^2+4az}\right],$$

cf. (3.8).

Hereafter we will show that, in this case, condition (3.9) is also necessary, besides sufficient, and we will be able to compute the optimum α, even though just numerically.

If only one relay is present, the sufficient condition for relaying superiority is $\lambda/|gh_d|^2 > 1 + P_s|h_s|^2/\sigma_d^2$ or, equivalently

$$\frac{P_s}{\sigma_u^2}|h_u|^2 > \left(1 + \frac{P_s}{\sigma_d^2}|h_s|^2\right)\frac{P_s}{\sigma_d^2}|h_s|^2. \tag{3.11}$$

It is interesting, and somehow surprising, to note that this condition does not depend on the relay downlink gain $|gh_d|^2$, meaning that the relay should be used each time its uplink SNR is higher than the threshold given by (3.11), whatever the quality of the relay–destination channel.

Lemma 3.1. *Consider a system with one single AF relay implementing the i.i.d. LD-STBC. Then, condition (3.11) is the necessary and sufficient condition for the superiority of relaying over the direct link.*

Proof. Sufficiency is already proven since (3.11) is simply the particularization of (3.9) to the one-relay case.

To show that condition (3.11) is also necessary, let us consider the derivative of $I_{\mathrm{AF,ML}}^{\mathrm{iid}}$ with respect to α, namely

$$\begin{aligned}\frac{\partial I_{\mathrm{AF,ML}}^{\mathrm{iid}}}{\partial\alpha} = {} & \frac{1}{(1+\alpha)^2}\left[\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) + \ln\frac{1+\lambda\phi_1}{1+|gh_d|^2\phi_2} - \ln\frac{\phi_2}{\phi_1} - \right.\\ & \left. - \frac{\sigma_d^2}{\sigma_u^2}(\phi_1-\phi_2)\right] + \frac{1}{1+\alpha}\left[\underbrace{\left(\alpha\frac{\lambda}{1+\phi_1\lambda} - \frac{1}{\phi_1} + \frac{\sigma_d^2}{\sigma_u^2}\right)}_{\text{(a)}}\frac{\partial\phi_1}{\partial\alpha} - \right.\\ & \left. - \underbrace{\left(\alpha\frac{|gh_d|^2}{1+\phi_2|gh_d|^2} - \frac{1}{\phi_2} + \frac{\sigma_d^2}{\sigma_u^2}\right)}_{\text{(b)}}\frac{\partial\phi_2}{\partial\alpha}\right]\\ = {} & \frac{1}{(1+\alpha)^2}\left[\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) + \ln\frac{1+\lambda\phi_1}{1+|gh_d|^2\phi_2} - \right.\\ & \left. - \ln\frac{\phi_2}{\phi_1} - \frac{\sigma_d^2}{\sigma_u^2}(\phi_1-\phi_2)\right]\end{aligned}$$

The second step follows since the terms (a) and (b) are null due to (3.6) and (3.7), respectively.

The sign of the derivative is obviously given by the sign of

$$\Gamma(\alpha) = \ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) + \ln\frac{1+\lambda\phi_1}{1+|gh_d|^2\phi_2} - \ln\frac{\phi_2}{\phi_1} - \frac{\sigma_d^2}{\sigma_u^2}(\phi_1-\phi_2).$$

This is still difficult to study, but one can reason as follows. First of all, note that both ϕ_1 and ϕ_2 tend to σ_u^2/σ_d^2 for $\alpha \to 0$, hence $\Gamma(\alpha)$ tends to $\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) + \ln\frac{\sigma_d^2/\sigma_u^2+\lambda}{\sigma_d^2/\sigma_u^2+|gh_d|^2}$, which is always greater than zero since $\lambda > |gh_d|^2$. The derivative of $\Gamma(\alpha)$ can be shown to be

$$\Gamma'(\alpha) = -(1+\alpha)\left[\frac{1}{\left(\frac{1+\lambda\phi_1}{\lambda\phi_1}\right)^2-\alpha} - \frac{1}{\left(\frac{1+|gh_d|^2\phi_2}{|gh_d|^2\phi_2}\right)^2-\alpha}\right],$$

which is always negative since $\lambda > |gh_d|^2$ implies $\lambda\phi_1 > |gh_d|^2\phi_2$. Then, $\Gamma(\alpha)$ is monotonically decreasing in α.

Summarizing, $\Gamma(\alpha)$ is positive in $\alpha = 0$ and monotonically decreasing. Besides, for $\alpha \to +\infty$, $\Gamma(\alpha)$ tends to $\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) - \ln(\lambda/|gh_d|^2)$, which is positive if and only if $\lambda < |gh_d|^2\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right)$. Then, if (3.11) does not hold, $\Gamma(\alpha)$ is always positive and $I_{\mathrm{AF,ML}}^{\mathrm{iid}}$ monotonically increases from 0 to I_s. Conversely, if (3.11) holds true, $\Gamma(\alpha)$ changes its sign only once from positive to negative and $I_{\mathrm{AF,ML}}^{\mathrm{iid}}$ presents a maximum larger than I_s. Note that the zero of $\Gamma(\alpha)$ and, hence, the coding rate that maximizes $I_{\mathrm{AF,ML}}^{\mathrm{iid}}$ can be found by binary search or other more efficient methods (e.g. Newton-Raphson). These considerations imply (3.11) is the necessary and sufficient condition for the superiority of relaying. □

3.3 The LMMSE receiver

The ML receiver analyzed so far is optimal, meaning that it maximizes the spectral efficiency at its output. However, it is non-linear and, thus, computationally expensive. This section studies a sub-optimal solution, namely the linear minimum-mean-square-error filter, which is known to be the linear receiver with the largest output signal-to-interference-plus-noise ratio (SINR). Roughly speaking, the LMMSE receiver detects the information on a symbol-by-symbol basis, instead of considering the entire message as a whole, and no information is extracted from inter-symbol interference (the interference generated by symbols other than the considered one because of the non-zero cross-correlation of the signatures).

As mentioned in Section 3.1, it is not difficult to find analogies between the presented coding scheme and DS/CDMA. Then, according to [86, 87], the inter-symbol interference can be considered as Gaussian when K and N are large enough. This fact implies that the output SINR of a LMMSE filter is a sufficient statistic to characterize the performance of the system; the spectral efficiency can be computed by means of the Shannon's "$\ln(1+SINR)$" formula and achieved by Gaussian source coding.

3.3.1 Spectral efficiency

We compute now the SINR for a generic symbol, e.g. s_1, without loss of generality. For the transmission scheme $\mathbf{d} = \mathbf{a}s_1 + \mathbf{n}_E$, the LMMSE filter and its output SINR are well known to be, respectively:

$$\mathbf{w} = \frac{P_s}{1 + P_s \mathbf{a}^H \mathbf{R}_E^{-1} \mathbf{a}} \mathbf{R}_E^{-1} \mathbf{a}, \qquad SINR = P_s \mathbf{a}^H \mathbf{R}_E^{-1} \mathbf{a};$$

where $\mathbf{R}_E = \mathbb{E}[\mathbf{n}_E \mathbf{n}_E^H]$ is the covariance matrix of the equivalent noise. To apply this result to our scheme, let us rewrite (3.1) as:

$$\mathbf{d} = \begin{bmatrix} h_s \\ \mathbf{0} \\ \widetilde{\mathbf{C}}_1 \mathbf{\Psi} \mathbf{h}_u \end{bmatrix} s_1 + \begin{bmatrix} [\mathbf{n}_1]_1 \\ h_s \mathbf{s}_1 + [\mathbf{n}_1]_{2:K} \\ \bar{\mathbf{C}}_1 \bar{\mathbf{\Psi}} \bar{\mathbf{H}}_u \mathbf{s}_1 + \widetilde{\mathbf{C}} \widetilde{\mathbf{\Psi}} \mathbf{n}_u + \mathbf{n}_d \end{bmatrix},$$

where:

- $\mathbf{s}_1 = [s_2\, s_3\, \cdots\, s_K]^T$
- the columns of the $N \times L$ matrix $\widetilde{\mathbf{C}}_1$ are the first columns of all the relay signature matrices $\mathbf{C}_l$;
- the matrix $\bar{\mathbf{C}}_1$ is made up by the columns of $\widetilde{\mathbf{C}}$ that remain after extracting $\widetilde{\mathbf{C}}_1$;
- $\bar{\mathbf{\Psi}} = \mathbf{\Psi} \otimes \mathbf{I}_{K-1}$ and $\bar{\mathbf{H}}_u = \mathbf{h}_u \otimes \mathbf{I}_{K-1}$;
- $\mathbf{n}_1 = \left[[\mathbf{n}_1]_1 \, [\mathbf{n}_1]_{2:K}^T \right]^T$.

The resulting SINR is, thus,

$$SINR_{\mathrm{AF}} = \frac{P_s}{\sigma_d^2} |h_s|^2 + \frac{P_s}{\sigma_u^2} \mathbf{h}_u^H \mathbf{h}_u - \frac{P_s}{\sigma_u^2} \mathbf{h}_u^H \left(\mathbf{I}_L + \mathbf{\Psi}^H \widetilde{\mathbf{C}}_1^H \mathbf{Q}^{-1} \widetilde{\mathbf{C}}_1 \mathbf{\Psi} \right)^{-1} \mathbf{h}_u,$$

with

$$\mathbf{Q} = \bar{\mathbf{C}}_1 \bar{\mathbf{\Psi}} \left(\frac{\frac{P_s}{\sigma_u^2}}{1 + \frac{P_s}{\sigma_d^2} |h_s|^2} \bar{\mathbf{H}}_u \bar{\mathbf{H}}_u^H + \mathbf{I}_{L(K-1)} \right) \bar{\mathbf{\Psi}}^H \bar{\mathbf{C}}_1^H + \frac{\sigma_d^2}{\sigma_u^2} \mathbf{I}_N.$$

Once again, the SINR is a random quantity because function of the random linear-dispersion matrices. However, its limit for $K = \alpha N$ tending to infinity results in a deterministic quantity. Recalling that the number L of relays is finite, [71, Theorem 7][3] can be applied to the L^2 entries of $\widetilde{\mathbf{C}}_1^H \mathbf{Q}^{-1} \widetilde{\mathbf{C}}_1$ to show that this matrix converges to $\phi_1 \mathbf{I}_L$ with probability 1, where ϕ_1

[3] See also Lemma 2.1 and Example 2.5 with $\mathbf{A} = \mathbf{0}$.

and $\{\lambda_l\}_{l=1}^{L}$ are exactly the same as in (3.6). The almost sure asymptotic equivalent of the SINR is hence:

$$SINR_{\mathrm{AF}}^{\mathrm{iid}} = \frac{P_s}{\sigma_d^2}|h_s|^2 + \frac{P_s}{\sigma_u^2}\sum_{l=1}^{L}\frac{|g_l h_{dl} h_{ul}|^2}{|g_l h_{dl}|^2 + \frac{1}{\phi_1}}, \tag{3.12}$$

where the dependence on the codes is concentrated in the coding rate α. The contributions of the source and of each one of the L relays can be clearly identified. Comparing the latter with the SNR of a single AF relay, i.e. $\frac{P_s|gh_d h_u|^2}{\sigma_u^2|gh_d|^2+\sigma_d^2}$, we can state that the effect of the random signatures in the LMMSE receiver is that each relay behaves as if it were the only AF relay in the system, except for the downlink noise which increases from σ_d^2 to σ_u^2/ϕ_1 (from (3.6), $\phi_1 < \sigma_u^2/\sigma_d^2$).

Observe that all the source symbols will be received with the same asymptotic SINR (3.12), which does not depend on the considered symbol s_1. Thanks to logarithm continuity, the asymptotic equivalent of the spectral efficiency can be simply expressed as:

$$I_{\mathrm{AF,LMMSE}}^{\mathrm{iid}} = \frac{\alpha}{1+\alpha}\ln(1 + SINR_{\mathrm{AF}}^{\mathrm{iid}}) \tag{3.13}$$

in nats per degree of freedom. A comparison with the two special cases in (3.2) and (3.3) reinforces the considerations exposed in Section 3.1.1: random codes, as orthogonal ones, obtain a coherent combination of relay contributions and, at the same time, allow controlling the fraction of exploited degrees of freedom. As it happened for the optimal receiver, however, the coding rate α should be tuned to get the best tradeoff between $\alpha/(1+\alpha)$ and $SINR_{\mathrm{AF}}^{\mathrm{iid}}$, respectively an increasing and a decreasing function of α.

3.3.2 A sufficient condition for relaying superiority

Similarly to the ML receiver, a sufficient condition for the superiority of relaying when employing the LMMSE receiver can be derived analyzing how $I_{\mathrm{LMMSE}}^{\mathrm{iid}}$ converges to I_s. Following the same steps as above, the following system of conditions can be written:

$$\begin{cases} \dfrac{P_s}{\sigma_u^2}\displaystyle\sum_{l=1}^{L}\dfrac{|h_{ul}|^2}{1+\left(\sum_{m\neq l}|g_m h_{dm}|^2\right)/|g_l h_{dl}|^2} > \\ \qquad > \left(1+\dfrac{P_s}{\sigma_d^2}|h_s|^2\right)\dfrac{\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right)}{1-\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right)}, \\ \dfrac{P_s}{\sigma_d^2}|h_s|^2 < e-1. \end{cases} \tag{3.14}$$

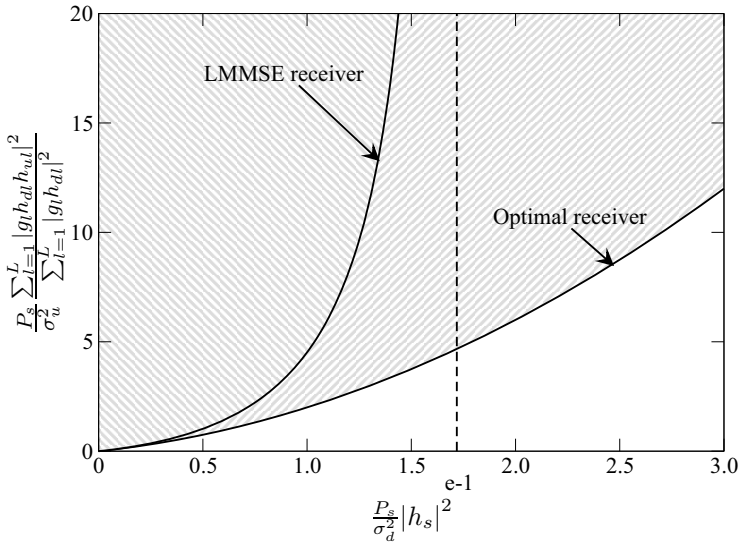


Figure 3.3: Relaying is certainly convenient whenever the weighted sum of the relay uplink SNR's falls in the shadowed regions above the curves.

The first inequality simplifies to

$$\frac{P_s}{\sigma_u^2}|h_u|^2 > \left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right)\frac{\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right)}{1-\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right)}$$

when only one relay is present.

In the LMMSE-receiver case, as in the ML one considered before, the sufficient condition (3.14) compares the same weighted sum of the relay uplink SNR's with an increasing function of $P_s|h_s|^2/\sigma_d^2$, the direct link SNR. Here, however, the condition is weaker, in the sense that it leaves out a large number of cases where relaying is convenient: simulation results show that the condition is not necessary, not even in the single relay case. Intuitions for this behavior can be found in Figure 3.3, which compares the regions where conditions (3.9) and (3.14) are satisfied.

3.4 Numerical results

This section gives a numerical assessment of the theoretical results presented in this chapter.

First of all, let us stress again the fact that the asymptotic equivalents derived in the paper are excellent approximations of the finite reality. To illustrate this fact, Figure 3.4 compares the asymptotic and the finite spectral efficiencies (averaged out over 1000 instances of the spreading matrices), for the two considered receivers and for $\alpha = 2.5$. In both cases, it can be noticed that there is no significant difference between the two curves, in spite

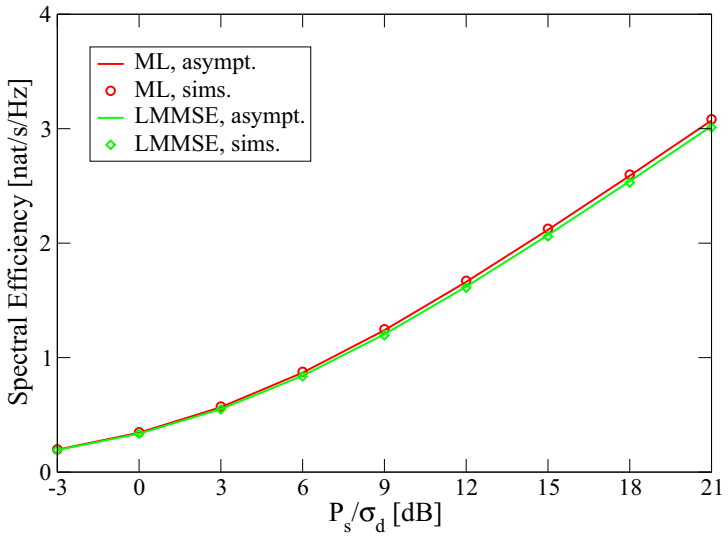


Figure 3.4: Spectral efficiency as a function of P_s/σ_d^2 for $\sigma_d^2/\sigma_u^2 = 1$, $|h_s|^2 = 0.4$, $L = 1$, $|h_u|^2 = |gh_d|^2 = 1$ and $\alpha = 2.5$. The two curves marked as "sims." represent, for $K = 5$ and $N = 2$, the mean value over 1000 realization of the spreading matrices.

of having fixed K and N to very low values (5 and 2, respectively). The following comments are hence based on the asymptotic results encountered in the paper.

Figure 3.5 shows the asymptotic spectral efficiencies (3.5) and (3.13) as functions of α, for different values of $|h_s|^2$. These curves confirm intuition and the sufficient condition derived in Section 3.2.1: relays are useful only when the direct link is weak and the receiver cannot reconstruct the information only from the source contribution. Moreover, when relaying is superior to the direct link, one can maximize the spectral efficiency by properly choosing α. In Section 3.2.1, for systems with one relay, it has been shown that $\alpha_{\max}$ can be found by binary search. Particular cases exist where $\alpha_{\max}$ can be expressed in closed form, as the low-power regime discussed in [88].

In Section 3.3, the spectral efficiency is computed for the best possible linear receiver, i.e. the LMMSE filter. The sub-optimality of this filter is evident when looking at Figures 3.4 and 3.5. The performance loss with respect to the optimal ML receiver is due to the fact that, in time-slot 2, the LMMSE filter does not exploit the correlation among the symbols introduced by the relay signatures. From Figure 3.4, however, it can be noted that the two receivers behave very similarly for low transmitted power: indeed, it is straightforward to show that

$$\lim_{P_s \to 0} \frac{1}{P_s} I^{\text{iid}}_{\text{AF,ML}} = \lim_{P_s \to 0} \frac{1}{P_s} I^{\text{iid}}_{\text{AF,LMMSE}} = \dot{I}(0)$$

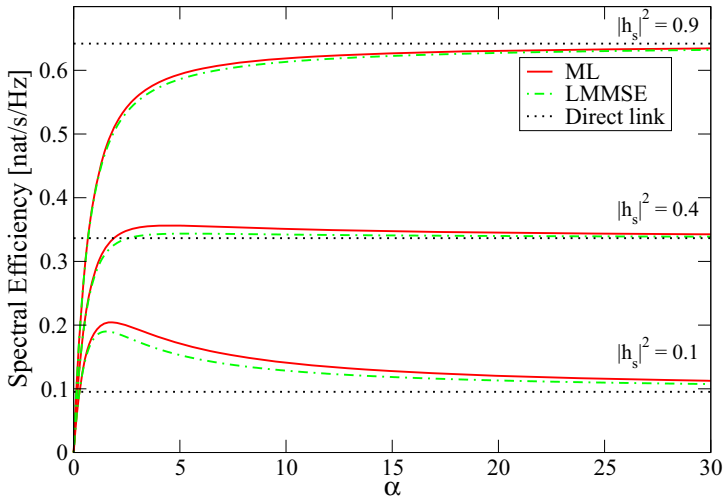


Figure 3.5: Asymptotic spectral efficiency as a function of α for one relay, $P_s/\sigma_d^2 = 1$, $\sigma_d^2/\sigma_u^2 = 1$ and $|h_u|^2 = |gh_d|^2 = 1$. With these assumptions, the threshold value of $|h_s|^2$ for relaying to be convenient is approximately 0.62.

with $0 < \dot{I}(0) < +\infty$. This result implies that the two receivers sustain reliable communications while the normalized energy per bit takes values as low as $(E_b/N_0)_{\min} = (\ln 2)/\dot{I}(0)$, where N_0 is the noise power spectral density [88, 89].

3.5 Concluding comments

This chapter has evaluated the instantaneous spectral efficiency of the proposed coding scheme, for both the ML and the LMMSE receiver. We have assumed AF relaying and full channel state information at the destination. Furthermore, it has been shown that the system behaves as a deterministic one when the dimensions K and N of the coding matrices grow indefinitely but keeping constant the coding rate $\alpha = K/N$.

Based on the asymptotic spectral efficiency, the relaying protocol has been compared to the simple point-to-point transmission over the direct link. A sufficient condition (which is also necessary in the single-relay case) has been found for the superiority of relaying. As suggested by intuition, relays should be employed only if the direct link is too weak to sustain a reliable communication.

The main drawback of orthogonal relaying protocols as the one considered here is probably the fact the source remains silent in the second phase of transmission, thus limiting the spectral efficiency. In [55], the same randomized LD-STBC is implemented also in non-orthogonal relaying , i.e. when the source transmits N new symbols during phase two. Even though the paper

shows that the non-orthogonal protocol is always superior to the orthogonal one, this dissertation only considers orthogonal relaying. Indeed, the resulting expressions of the spectral efficiency are simpler and it is easier to follow with the analysis of the proposed distributed random linear-dispersion coding scheme.

In spite of the encouraging results, time-invariant channels are not the best assumption when dealing with relays. Indeed, typical wireless communications systems are usually affected by some kind of fading: as we will see in the next chapter, it is in this situation that diversity plays a fundamental role, since it increases the probability to find a good path to convey information from the source to the destination.

Appendix 3.A Proof of asymptotic results

In this appendix we will prove that (3.5) is the almost sure asymptotic equivalent of the finite spectral efficiency (3.4).

To start with, equation (3.4) can be rewritten as:

$$I_{\mathrm{ML}} = \frac{\alpha}{1+\alpha}\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) + \frac{1}{N+K}\ln\det\left[\mathbf{I}_N + \right.$$
$$\left. + \frac{\sigma_u^2}{\sigma_d^2}\widetilde{\mathbf{C}}\left(\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{\Psi}}^H + \frac{\frac{P_s}{\sigma_u^2}}{1+\frac{P_s}{\sigma_d^2}|h_s|^2}\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{H}}_u\widetilde{\mathbf{H}}_u^H\widetilde{\mathbf{\Psi}}^H\right)\widetilde{\mathbf{C}}^H\right] -$$
$$- \frac{1}{N+K}\ln\det\left[\mathbf{I}_N + \frac{\sigma_u^2}{\sigma_d^2}\widetilde{\mathbf{C}}\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{C}}^H\right] \quad (3.15)$$

after some algebra. Our problem reduces to studying the asymptotic behavior of the last two terms, which have the common structure $\ln\det(\mathbf{I}_N + z\widetilde{\mathbf{C}}\widetilde{\mathbf{T}}\widetilde{\mathbf{C}}^H)$. The proof is based on the concepts introduced in Proposition 2.1 and Example 2.8.

3.A.1 General preliminaries

We will now study the function

$$\theta(z) = \frac{1}{N}\ln\det\left(\mathbf{I}_N + z\widetilde{\mathbf{C}}\widetilde{\mathbf{T}}\widetilde{\mathbf{C}}^H\right), \quad (3.16)$$

where $\widetilde{\mathbf{C}}$ is the coding matrix introduced in Section 3.1.1 and $\widetilde{\mathbf{T}} = \mathbf{T}\otimes\mathbf{I}_K$, being $\mathbf{T}$ a $L\times L$ deterministic, nonnegative-definite matrix with eigenvalues $\{\tau_l\}_{l=1}^L$. In particular, the limit for $K, N\to\infty$ with $K/N\to\alpha$ ($0<\alpha<+\infty$) will be computed.

Since

$$\theta(z) = \frac{1}{N}\ln\det\left(\mathbf{I}_N + z\widetilde{\mathbf{C}}\widetilde{\mathbf{T}}\widetilde{\mathbf{C}}^H\right) = \frac{1}{N}\ln\det\left(\mathbf{I}_{KL} + z\widetilde{\mathbf{T}}^{1/2}\widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^H\widetilde{\mathbf{T}}^{1/2}\right),$$

it was shown in Example 2.8 that we can write

$$\theta(z) = \frac{KL}{N} \int_0^z \left(\frac{1}{t} + \frac{1}{t^2} m\left(-\frac{1}{t}\right) \right) \mathrm{d}t, \tag{3.17}$$

where $m(x)$ is the Stieltjes transform associated to the KL eigenvalues of $\widetilde{\mathbf{T}}^{1/2}\widetilde{\mathbf{C}}^H\widetilde{\mathbf{C}}\widetilde{\mathbf{T}}^{1/2}$, namely

$$m(x) = \frac{1}{KL} \mathrm{tr}\left(\left[x\mathbf{I}_{KL} - \widetilde{\mathbf{T}}^{1/2}\widetilde{\mathbf{C}}^H\widetilde{\mathbf{C}}\widetilde{\mathbf{T}}^{1/2} \right]^{-1} \right).$$

Let us define

$$f(x) = \frac{x}{1 - L\alpha + L\alpha x m(x)}.$$

Then, with the change of variable $x = -\frac{1}{t}$, (3.17) can be further rewritten as

$$\theta(x) = \int_{-\infty}^{-1/z} \left(\frac{1}{f(x)} - \frac{1}{x} \right) \mathrm{d}x. \tag{3.18}$$

Asymptotically, i.e. for $K = \alpha N \to \infty$, and almost surely, $m(x)$ is the unique positive solution to

$$m(x) = \frac{1}{L} \sum_{l=1}^{L} \frac{1}{x - \tau_l(1 - L\alpha + L\alpha x m(x))}$$

(see Example 2.6 and, for instance, [69]). Thus, $f(x)$ satisfies:

$$x = f(x) \left(1 - \alpha \sum_{l=1}^{L} \frac{\tau_l}{\tau_l - f(x)} \right). \tag{3.19}$$

The derivative of $f(x)$ is given by

$$f'(x) = \left(1 - \alpha \sum_{l=1}^{L} \left(\frac{\tau_l}{\tau_l - f(x)} \right)^2 \right)^{-1}.$$

Injecting (3.19) into (3.18) and writing the unity as $1 = f'(x)/f'(x)$, we obtain

$$\begin{aligned} \theta(z) &= -\int_{-\infty}^{-1/z} \frac{1}{f(x)} \frac{\alpha \sum_{l=1}^{L} \frac{\tau_l}{\tau_l - f(x)}}{1 - \alpha \sum_{l=1}^{L} \frac{\tau_l}{\tau_l - f(x)}} \left(1 - \alpha \sum_{l=1}^{L} \left(\frac{\tau_l}{\tau_l - f(x)} \right)^2 \right) f'(x)\, \mathrm{d}x \\ &= -\int_{-\infty}^{f(-1/z)} \frac{1}{y} \frac{\alpha \sum_{l=1}^{L} \frac{\tau_l}{\tau_l - y}}{1 - \alpha \sum_{l=1}^{L} \frac{\tau_l}{\tau_l - y}} \left(1 - \alpha \sum_{l=1}^{L} \left(\frac{\tau_l}{\tau_l - y} \right)^2 \right) \mathrm{d}y, \end{aligned} \tag{3.20}$$

where we made the change of variable $y = f(x)$.

With the new change of variable $s = -1/y$, we have

$$\theta(z) = \int_0^{\phi(z)} \frac{\alpha \sum_{l=1}^{L} \frac{\tau_l}{1+s\tau_l}}{1-\alpha \sum_{l=1}^{L} \frac{s\tau_l}{1+s\tau_l}} \left(1 - \alpha \sum_{l=1}^{L} \left(\frac{s\tau_l}{1+s\tau_l}\right)^2\right) \mathrm{d}s,$$

where $\phi(z) = -1/f(-\frac{1}{z})$ is the unique positive solution to

$$1 = \frac{\phi(z)}{z} + \alpha \sum_{l=1}^{L} \frac{\phi(z)\tau_l}{1+\phi(z)\tau_l}. \tag{3.21}$$

Writing

$$\begin{aligned}
1 - \alpha \sum_{l=1}^{L} \left(\frac{s\tau_l}{1+s\tau_l}\right)^2 =& 1 - \alpha \sum_{l=1}^{L} \frac{s\tau_l}{1+s\tau_l} + \alpha \sum_{l=1}^{L} \frac{s\tau_l}{1+s\tau_l} - \alpha \sum_{l=1}^{L} \left(\frac{s\tau_l}{1+s\tau_l}\right)^2 \\
=& 1 - \alpha \sum_{l=1}^{L} \frac{s\tau_l}{1+s\tau_l} + s\alpha \sum_{l=1}^{L} \frac{\tau_l}{(1+s\tau_l)^2},
\end{aligned}$$

$\theta(z)$ can be reformulated as:

$$\begin{aligned}
\theta(z) =& \alpha \sum_{l=1}^{L} \int_0^{\phi(z)} \frac{\tau_l}{1+s\tau_l} \mathrm{d}s + \int_0^{\phi(z)} \frac{s\alpha \sum_{l=1}^{L} \frac{\tau_l}{1+s\tau_l} \alpha \sum_{l=1}^{L} \frac{\tau_l}{(1+s\tau_l)^2}}{1-\alpha \sum_{l=1}^{L} \frac{s\tau_l}{1+s\tau_l}} \mathrm{d}s \\
=& \alpha \sum_{l=1}^{L} \ln[1+\phi(z)\tau_l] - \alpha \sum_{l=1}^{L} \int_0^{\phi(z)} \frac{\tau_l}{(1+s\tau_l)^2} \mathrm{d}s - \\
& - \int_0^{\phi(z)} \frac{-\alpha \sum_{l=1}^{L} \frac{\tau_l}{(1+s\tau_l)^2}}{1-\alpha \sum_{l=1}^{L} \frac{s\tau_l}{1+s\tau_l}} \mathrm{d}s \\
=& \alpha \sum_{l=1}^{L} \ln[1+\phi(z)\tau_l] + \alpha \sum_{l=1}^{L} \left(\frac{1}{1+\phi(z)\tau_l} - 1\right) - \\
& - \ln\left[1 - \alpha \sum_{l=1}^{L} \frac{\phi(z)\tau_l}{1+\phi(z)\tau_l}\right].
\end{aligned}$$

Comparing the last two terms of the last equation with (3.21), one obtains the asymptotic almost sure expression of $\theta(z)$, namely

$$\theta(z) = \alpha \sum_{l=1}^{L} \ln\left[1+\phi(z)\tau_l\right] + \ln\left[\frac{z}{\phi(z)}\right] + \frac{\phi(z)}{z} - 1.$$

3.A.2 Asymptotic spectral efficiency

The previous result can be used to show that

$$\left| \frac{1}{N} \ln \det \left[\mathbf{I}_N + \frac{\sigma_u^2}{\sigma_d^2} \widetilde{\mathbf{C}} \left(\widetilde{\mathbf{\Psi}} \widetilde{\mathbf{\Psi}}^H + \frac{\frac{P_s}{\sigma_u^2}}{1 + \frac{P_s}{\sigma_d^2} |h_s|^2} \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{H}}_u \widetilde{\mathbf{H}}_u^H \widetilde{\mathbf{\Psi}}^H \right) \widetilde{\mathbf{C}}^H \right] - \right.$$
$$\left. - \left[\alpha \sum_{l=1}^{L} \ln(1 + \phi_1 \lambda_l) + \ln \left(\frac{\sigma_u^2}{\sigma_d^2 \phi_1} \right) + \frac{\sigma_d^2}{\sigma_u^2} \phi_1 - 1 \right] \right| \xrightarrow{\text{a.s.}} 0,$$

and

$$\left| \frac{1}{N} \ln \det \left[\mathbf{I}_N + \frac{\sigma_u^2}{\sigma_d^2} \widetilde{\mathbf{C}} \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{C}}^H \right] - \right.$$
$$\left. - \left[\alpha \sum_{l=1}^{L} \ln(1 + \phi_2 |g_l h_{dl}|^2) + \ln \left(\frac{\sigma_u^2}{\sigma_d^2 \phi_2} \right) + \frac{\sigma_d^2}{\sigma_u^2} \phi_2 - 1 \right] \right| \xrightarrow{\text{a.s.}} 0.$$

Both results hold almost surely for $K = \alpha N \to \infty$. The quantities ϕ_1 and ϕ_2 are defined as in (3.6) and (3.7), respectively. It is now straightforward to express the almost sure, deterministic asymptotic equivalent of the spectral efficiency as in (3.5).

3.A.3 The eigenvalues of $\mathbf{\Psi}\mathbf{\Psi}^H + \chi \mathbf{\Psi} \mathbf{h}_u \mathbf{h}_u^H \mathbf{\Psi}^H$

The coefficients $\{\lambda_l\}_{l=1}^{L}$, which appear in the definition of (3.6), are the eigenvalues of a matrix of the form $\mathbf{\Psi}\mathbf{\Psi}^H + \chi \mathbf{\Psi} \mathbf{h}_u \mathbf{h}_u^H \mathbf{\Psi}^H$, with $\mathbf{\Psi}$ and $\mathbf{h}_u$ as in the system model. We will show now how to solve the characteristic equation

$$f(\lambda) = \det \left(\mathbf{\Psi}\mathbf{\Psi}^H + \chi \mathbf{\Psi} \mathbf{h}_u \mathbf{h}_u^H \mathbf{\Psi}^H - \lambda \mathbf{I}_L \right) = 0.$$

First of all, note that if the diagonal matrix $\mathbf{\Psi}\mathbf{\Psi}^H$ has n entries equal to $\hat{\psi}$, with $2 \leq n \leq L$, then $\lambda = \hat{\psi}$ is a zero of order $n-1$ of $f(\lambda)$. It is enough to write all the entries of the matrix $\mathbf{\Psi}\mathbf{\Psi}^H + \chi \mathbf{\Psi} \mathbf{h}_u \mathbf{h}_u^H \mathbf{\Psi}^H$ and note that the n rows corresponding to $\hat{\psi}$ only differ for a constant factor.

Knowing this, we can now focus on the case where $\mathbf{\Psi}\mathbf{\Psi}^H - \lambda \mathbf{I}_L$ is full rank to seek for the other eigenvalues. One has:

$$\begin{aligned}
&\det \left(\mathbf{\Psi}\mathbf{\Psi}^H + \chi \mathbf{\Psi} \mathbf{h}_u \mathbf{h}_u^H \mathbf{\Psi}^H - \lambda \mathbf{I}_L \right) \\
&\quad = \det \left(\mathbf{\Psi}\mathbf{\Psi}^H - \lambda \mathbf{I}_L \right) \cdot \\
&\qquad \cdot \det \left[\mathbf{I}_L + \chi (\mathbf{\Psi}\mathbf{\Psi}^H - \lambda \mathbf{I}_L)^{-1} \mathbf{\Psi} \mathbf{h}_u \mathbf{h}_u^H \mathbf{\Psi}^H \right] \\
&\quad = \det \left(\mathbf{\Psi}\mathbf{\Psi}^H - \lambda \mathbf{I}_L \right) \cdot \\
&\qquad \cdot \left[1 + \chi \mathbf{h}_u^H \mathbf{\Psi}^H (\mathbf{\Psi}\mathbf{\Psi}^H - \lambda \mathbf{I}_L)^{-1} \mathbf{\Psi} \mathbf{h}_u \right].
\end{aligned}$$

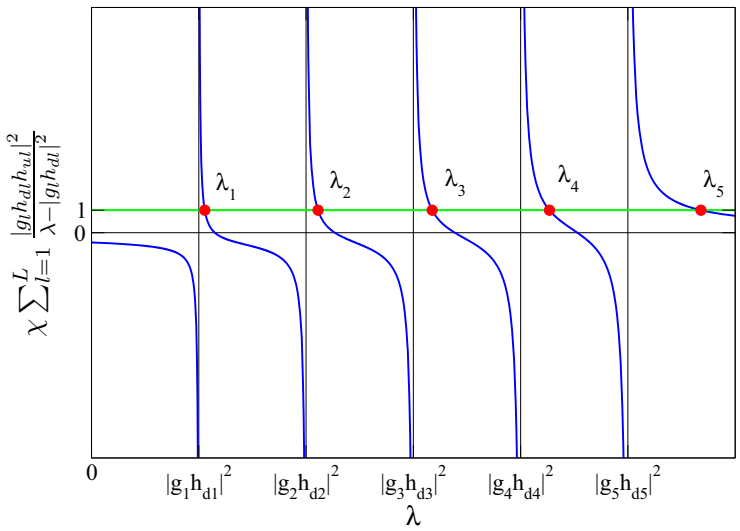


Figure 3.6: Graphical solution of the equation $\chi\sum_{l=1}^{L}\frac{|g_l h_{dl} h_{ul}|^2}{\lambda-|g_l h_{dl}|^2}=1$.

The only possible zeros are those of the second term, namely the solutions to

$$\begin{aligned} 0 &= 1+\chi\mathbf{h}_u^H\mathbf{\Psi}^H\left(\mathbf{\Psi}\mathbf{\Psi}^H-\lambda\mathbf{I}_L\right)^{-1}\mathbf{\Psi}\mathbf{h}_u \\ &= 1+\chi\sum_{l=1}^{L}\frac{|g_l h_{dl} h_{ul}|^2}{|g_l h_{dl}|^2-\lambda}. \end{aligned} \tag{3.22}$$

These solutions can be encountered graphically as in Figure 3.6, which also shows that, taking $|g_1 h_{d1}|^2 \le |g_2 h_{d2}|^2 \le \cdots \le |g_L h_{dL}|^2$ and $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_L$, one has

$$|g_l h_{dl}|^2 \le \lambda_l < |g_{l+1} h_{d(l+1)}|^2 \qquad \text{for } l=1,\ldots,L-1$$

and

$$\lambda_L \ge |g_L h_{dL}|^2$$

Rewriting (3.22) in polynomial form, one can further show that

$$\sum_{l=1}^{L}\lambda_l = \sum_{l=1}^{L}|g_l h_{dl}|^2\left(1+\chi|h_{ul}|^2\right),$$

where all possible λ's (including those derived in the first step) have been considered.

Chapter 4

Large-SNR Outage Analysis of Relay Channels Employing i.i.d. Randomized LD-STBC

In the previous chapter, we computed the spectral efficiency achieved by two different receivers in an AF relay channel implementing the randomized i.i.d. LD-STBC. For specific values of the channel coefficients, the spectral efficiency represents the maximum information rate that can be sustained free of decoding errors with indefinitely long source codes. Nevertheless, relay networks and, more generally, most of wireless communications systems are characterized by time-varying fading channels (we will only consider frequency-flat fading, as opposed to frequency-selective one). Under this assumption, the idea of maximum sustainable rate loses part of its meaning. Indeed, for most of the statistical channel models, there exists a non-zero probability of deep fading where no communications can be established at any rate.

One possible way to evaluate the system performance under the fading assumption is to consider the ergodic spectral efficiency, i.e. the expected value of the spectral efficiency with respect to the channel distribution. This quantity represents the mean transmission rate that results from transmitting at a rate equal to the instantaneous spectral efficiency at every channel realization (the channel coefficients are modeled as stationary ergodic random variables[1]).

Observe that the ergodic spectral efficiency hides the real channel fluctuations. For most applications, it is more interesting to know how often a given rate cannot be achieved[2].

[1] A stationary process is a stochastic process whose joint distribution does not change when shifted in time. A stochastic process is said to be ergodic if its statistical properties can be deduced from a single, sufficiently long sample.

[2] For more information about ergodic and outage capacity, please refer to, e.g., [1].

Outage probability

In what follows, we are hence going to consider the spectral efficiency I as a random quantity, since function of the random channel coefficients. As mentioned in Section 3.1.1, all channels are assumed to undergo independent, frequency-flat, quasi-static fading. This means that they can be modeled by random variables that take on independent values at any new source message. We will not make any special assumption on the probability distribution of the channel gains, except that they should be right-continuous at zero[3]. The exponential distribution, which characterizes the typical Rayleigh fading model, satisfies this hypothesis.

Now, the outage probability $P_{out}(R)$ is defined as the probability that the system cannot support a given target rate R, namely

$$P_{out}(R) = \Pr[I < R].$$

Unfortunately, except for some very specific cases, it is difficult to derive the distribution of the spectral efficiency and, thus, the outage probability. For this reason, we will focus on a high SNR approximation and derive the diversity order d ($0 < d < +\infty$) and the outage gain κ ($0 < \kappa < +\infty$) that satisfy the following equation:

$$\kappa = \lim_{SNR \to +\infty} SNR^d P_{out}(R). \tag{4.1}$$

The diversity order gives an idea of how fast the outage probability decays when increasing the transmitted power.

On the other hand, when more power is available, one can also decide to increase the information rate R (equivalently, to work at a higher κ). Assuming that the target rate grows with the SNR as $R = r \ln SNR$ (r being the multiplexing gain), the diversity–multiplexing tradeoff (DMT, see [46]) is defined as

$$d(r) = -\lim_{SNR \to +\infty} \frac{\ln P_{out}(r \ln SNR)}{\ln SNR}$$

and provides an intuitive idea about how the system shares its degrees of freedom between diversity and multiplexing capabilities.

In what follows, we will study the large-SNR outage probability of the asymptotic spectral efficiency I^{iid} defined in Chapter 3. Note that this is not necessarily the same as considering the asymptotic (in K and N) behavior of the outage probability. Formally, interchanging the two limits for $K = \alpha N \to +\infty$ and for $SNR \to +\infty$ does not need to yield the same result. However, based on simulation results, we conjecture that this equality holds true.

[3] In other words, their density does not escape to infinity at zero

As discussed in Section 3.1.1, one of the main issues of the AF relaying strategy is the colored structure of the forwarded noise. For this reason, we start the outage analysis from the decode-and-forward (DF) version of the LD-STBC relaying protocol, where relays forward a perfect non-noisy copy of the source message.

4.1 LD-STBC in DF relaying

The communications system considered here is totally analogous to the one presented in Section 3.1.1, except that we replace AF relays by DF ones. Then, at the end of the first transmission phase, relays will attempt to decode the source message $\mathbf{s}$: only the relays that successfully decode $\mathbf{s}$ participate in the second transmission phase[4]. Let $\mathcal{L} = \{1, \ldots, L\}$ be the set of all relays and $\mathcal{L}' \subseteq \mathcal{L}$ the decoding subset, that is the set of relays that are allowed to join phase two. The signal model for DF relaying (c.f. (3.1)) is given by

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} h_s \mathbf{I}_k \\ \frac{1}{\sqrt{\alpha}} \sum_{l \in \mathcal{L}'} h_{dl} \mathbf{C}_l \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix}, \tag{4.2}$$

where we fixed $|g_l|^2 = \frac{1}{\alpha}$, $l = 1, \ldots, L$, to force the mean relay symbol power to be equal to P_s.

4.1.1 The ML receiver

With very similar reasoning as in Section 3.2, the spectral efficiency obtained by the ML receiver from the signal model in (4.2) may be expressed as

$$I_{\mathrm{DF,ML}}^{(\mathcal{L}')} = \frac{1}{K+N} \ln \det \Bigg(\mathbf{I}_{K+N} + \\ + \frac{P_s}{\sigma_d^2} \begin{bmatrix} h_s \mathbf{I}_K \\ \frac{1}{\sqrt{\alpha}} \sum_{l \in \mathcal{L}'} h_{dl} \mathbf{C}_l \end{bmatrix} \begin{bmatrix} h_s^* \mathbf{I}_K & \frac{1}{\sqrt{\alpha}} \sum_{l \in \mathcal{L}'} h_{dl}^* \mathbf{C}_l^H \end{bmatrix} \Bigg),$$

and is obviously dependent on the decoding set $\mathcal{L}'$. Using the results in Appendix 3.A.1, it is not difficult to show that the asymptotic spectral

[4]For DF relaying, the analysis of the spectral efficiency and its comparison to the direct link is not as meaningful as in the AF case (see Chapter 3), since we do not know how many relays decode source information and participate in the second transmission phase.

efficiency[5] conditioned to the decoding set $\mathcal{L}'$ is given by

$$I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}} = \frac{1}{1+\alpha}\Bigg[\alpha \ln\left(1+\rho|h_s|^2\right) + \\ + \alpha \ln\left(1+\frac{\rho\sum_{l\in\mathcal{L}'}|h_{dl}|^2}{\alpha\beta_{\mathcal{L}'}(1+\rho|h_s|^2)}\right) + \ln\beta_{\mathcal{L}'} + \frac{1}{\beta_{\mathcal{L}'}} - 1\Bigg], \quad (4.3)$$

where $\rho = P_s/\sigma_d^2$ is the reference SNR and $\beta_{\mathcal{L}'}$ is the unique positive solution to

$$\beta_{\mathcal{L}'} = 1 + \alpha\beta_{\mathcal{L}'}\frac{\frac{\rho}{1+\rho|h_s|^2}\sum_{l\in\mathcal{L}'}|h_{dl}|^2}{\alpha\beta_{\mathcal{L}'} + \frac{\rho}{1+\rho|h_s|^2}\sum_{l\in\mathcal{L}'}|h_{dl}|^2}, \quad (4.4)$$

namely

$$\beta_{\mathcal{L}'} = \frac{\alpha+\alpha\chi-\chi+\sqrt{(\alpha+\alpha\chi-\chi)^2+4\alpha\chi}}{2\alpha}.$$

To avoid cumbersome notation, we denoted by χ the term $\frac{\rho}{1+\rho|h_s|^2}\sum_{l\in\mathcal{L}'}|h_{dl}|^2$.

Outage analysis

Since the spectral efficiency depends on the decoding set, the outage probability has to be computed by means of the total probability theorem as follows:

$$P_{out}(R) = \sum_{\mathcal{L}'\subseteq\mathcal{L}} \Pr[\mathcal{L}' \text{ is the decoding set}]\,\Pr\big[I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}} < R\big],$$

where the sum is over all possible subsets of $\mathcal{L}$. Without going into implementation details, we will simply assume that the relay r_l is part of the decoding set whenever the source–relay link is not in outage, i.e. when

$$I_{ul} = \frac{\alpha}{1+\alpha}\ln\big(1+z\rho|h_{ul}|^2\big) \geq R,$$

which implies

$$\Pr[\mathcal{L}' \text{ is the decoding set}] = \prod_{l\in\mathcal{L}'}\Pr[I_{ul}\geq R]\prod_{l\in\overline{\mathcal{L}'}}\Pr[I_{ul}<R],$$

where $\overline{\mathcal{L}'} = \mathcal{L}\backslash\mathcal{L}'$ is the complement of the decoding set (see also [26,29,32]).

[5] By this, we mean that $I_{\mathrm{DF,ML}}^{(\mathcal{L}')} \to I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}}$ almost surely when $K, N \to +\infty$, $K/N \to \alpha$ with $0 < \alpha < +\infty$.

Theorem 4.1. *Consider the system described in Section 4.1 with L DF relays employing the randomized LD-STBC. Assume maximum-likelihood reception with full channel state information, resulting in the asymptotic spectral efficiency (4.3). Then, the diversity order is independent of the coding rate α and equal to the total number of transmitters (relays plus source) $L+1$. The outage gain is given by the integral*

$$\kappa = \lim_{\rho\to+\infty} \rho^{L+1} P_{out}(R) = \zeta_s \sum_{l=0}^{L} \binom{L}{l} \left(\frac{\zeta_u Q}{z}\right)^{L-l} \zeta_d^l \cdot \\ \cdot \int_{\mathbb{R}_+^{l+1}} \mathbb{1}\big\{f(a, \mathbf{c}_1^l) < (1+\alpha)R\big\}\, \mathrm{d}a\, \mathrm{d}\mathbf{c}_1^l, \quad (4.5)$$

where we have introduced the function

$$f(a, \mathbf{c}_1^l) = \alpha \ln(1+a) + \alpha \ln\left(1 + \frac{\sum_{l=1}^{L} c_l}{\alpha \beta_{\mathcal{L}'}(1+a)}\right) + \ln \beta_{\mathcal{L}'} + \frac{1}{\beta_{\mathcal{L}'}} - 1$$

and the quantity

$$Q = \exp\Big(\frac{1+\alpha}{\alpha} R\Big) - 1.$$

Proof. See Appendix 4.A.1. □

As mentioned in the introduction, the results of this chapter are derived without considering any special probability distribution for the channel gains. The only required assumption is right-continuity in zero of the density functions. Even though not necessary, but in order to simplify the exposition, we assume all the uplink channels to be identically distributed. Same thing for the downlink channels. In Theorem 4.1, and the following ones, we refer to the quantities:

$$\zeta_s = \lim_{x\to 0^+} f_{|h_s|^2}(x), \quad \zeta_u = \lim_{x\to 0^+} f_{|h_u|^2}(x), \quad \zeta_d = \lim_{x\to 0^+} f_{|h_d|^2}(x), \quad (4.6)$$

where $f_{|h|^2}(x)$ is the probability density function (pdf) of the channel gain $|h|^2$. This general model includes, for instance, the classical Rayleigh fading case, where $|h|^2$ is exponentially distributed with variance $1/\zeta^2$. We also employ the following notations for $i \le j$:

$$\mathrm{d}\mathbf{x}_i^j = \mathrm{d}x_i\, \mathrm{d}x_{i+1} \dots \mathrm{d}x_j;$$
$$f_{|h|^2}(\mathbf{x}_i^j) = f_{|h|^2}(x_i) f_{|h|^2}(x_{i+1}) \dots f_{|h|^2}(x_j),$$

and, similarly, for any subset $\mathcal{L}'$

$$\mathrm{d}\mathbf{x}_{\mathcal{L}'} = \prod_{l\in\mathcal{L}'} \mathrm{d}x_l;$$
$$f_{|h|^2}(\mathbf{x}_{\mathcal{L}'}) = \prod_{l\in\mathcal{L}'} f_{|h|^2}(x_l).$$

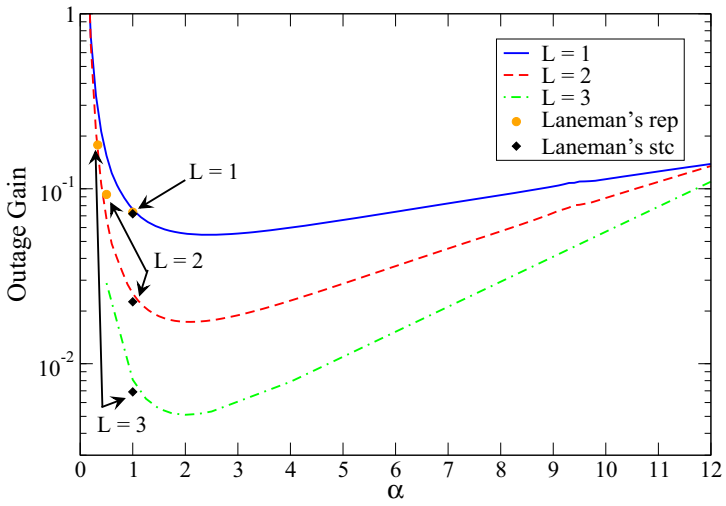


Figure 4.1: Outage gain as a function of α for DF relays, ML receiver, unitary channel variances, $z = 1$ and $R = 0.1$ nat/s/Hz. This scheme achieves full diversity order $L + 1$.

The interpretation of Theorem 4.1 is very simple: the necessary condition for the system to suffer an outage event is that each one of the relays cannot communicate with either the source or the destination. Note that full diversity is achieved even for large values of the coding rate ($\alpha \gg 1$ or, equivalently, $K \gg N$), which correspond to short source silences. However, fixing α equal to a large value is not, in general, a good relaying strategy since it does not minimize the outage gain. Unfortunately, this parameter can be computed only by solving numerically the integral in (4.5) and the task of minimizing such expression as a function of α is not an easy one. However, numerical results as the three examples depicted in Figure 4.1 suggest that the outage gain is a convex function of α with a minimum located at a relative low value of α (around 2 in the reported examples).

4.1.2 The LMMSE receiver

Following the same guidelines as in Section 3.3, the LMMSE SINR for the signal model in (4.2) can be expressed as:

$$SINR_{\mathrm{DF}}^{(\mathcal{L}')} = \frac{P_s}{\sigma_d^2}|h_s|^2 + \frac{P_s}{\alpha\sigma_d^2}\left(\sum_{l\in\mathcal{L}'} h_{dl}^* \mathbf{c}_{1,l}^H\right) \cdot$$
$$\cdot \left(\frac{P_s/\sigma_d^2}{\alpha(1+P_s|h_s|^2/\sigma_d^2)}\left(\sum_{l\in\mathcal{L}'} h_{dl}^* \mathbf{c}_{1,l}^H\right)\left(\sum_{l\in\mathcal{L}'} h_{dl}\mathbf{c}_{1,l}\right) + \mathbf{I}_N\right)^{-1}\left(\sum_{l\in\mathcal{L}'} h_{dl}\mathbf{c}_{1,l}\right),$$

where the focus is on the first source symbol s_1, without loss of generality.

When we let $K, N \to +\infty$, $K/N \to \alpha$ with $0 < \alpha < +\infty$, the SINR converges almost surely to a deterministic quantity. Given the decoding set $\mathcal{L}'$, the asymptotic LMMSE SINR is

$$SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}} = \rho|h_s|^2 + \frac{\rho}{\alpha\beta_{\mathcal{L}'}} \sum_{l\in\mathcal{L}'} |h_{dl}|^2, \tag{4.7}$$

with $\beta_{\mathcal{L}'}$ as in (4.4). As in the AF case, the proof is a direct application of [71, Theorem 7].

Prior to considering the outage probability, let us recall the parameter

$$Q = \exp\Big(\frac{1+\alpha}{\alpha}R\Big) - 1, \tag{4.8}$$

which is a function of α. Note that

$$I_{\mathrm{DF,LMMSE}}^{(\mathcal{L}'),\mathrm{iid}} = \frac{\alpha}{1+\alpha}\ln(1 + SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}}) < R$$

if and only if

$$SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}} < Q.$$

Now, let α_{th} denote the unique positive solution to

$$\alpha_{th} = 1 + \frac{1}{Q(\alpha_{th})}. \tag{4.9}$$

The large-SNR analysis of the outage probability leads to the following result:

Theorem 4.2. *Consider the system described in Section 4.1 with L DF relays employing the randomized LD-STBC. Assume LMMSE reception with full channel state information, resulting in the asymptotic SINR (4.7). Then, the diversity order is a function of the coding rate α, namely*

$$d = \begin{cases} L+1 & \text{for } \alpha \le \alpha_{th}; \\ 1 & \text{for } \alpha > \alpha_{th}. \end{cases}$$

With the definitions in (4.6), the closed-form expressions of the relative outage gains are given by

$$\kappa = \lim_{\rho\to+\infty} \rho^{L+1} P_{out}(R) = \zeta_s \sum_{l=0}^{L} \binom{L}{l} \left(\frac{\zeta_u Q}{z}\right)^{L-l} \frac{(\alpha\zeta_d)^l}{l!} \mathcal{P}_l, \tag{4.10}$$

when $\alpha \le \alpha_{th}$ (coefficients $\{\mathcal{P}_l : l = 1, \ldots, L\}$ are defined by (4.32) in Appendix 4.A.2) and

$$\kappa = \lim_{\rho\to+\infty} \rho P_{out}(R) = \zeta_s \left(Q - \frac{Q+1}{\alpha}\right), \tag{4.11}$$

when $\alpha > \alpha_{th}$.

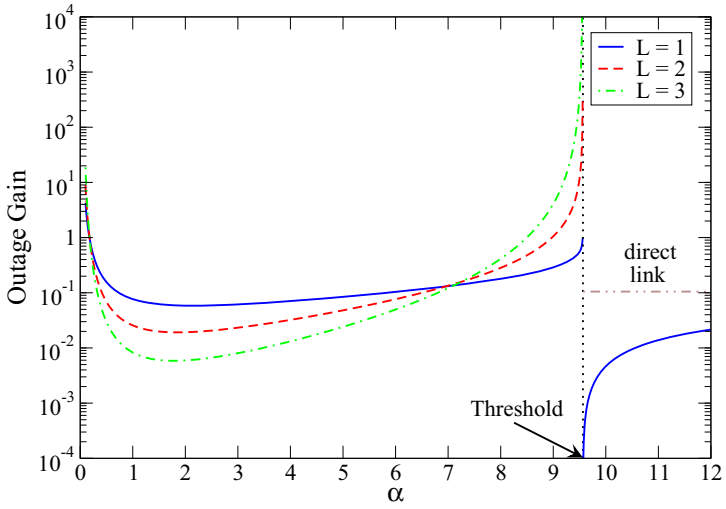


Figure 4.2: Outage gain as a function of α for DF relays, LMMSE receiver, unitary channel variances, $z = 1$ and $R = 0.1$ nat/s/Hz. The diversity order is $L + 1$ on the left of the threshold and 1 on its right.

Proof. See Appendix 4.A.2. □

As opposed to the ML receiver, the LMMSE filter may exploit diversity only for low values of α. Nevertheless, since $\alpha_{th} > 1$, we can achieve full diversity even if the source transmission phase is longer than the relaying one. In other words, it is theoretically possible to transmit over more than half of the total degrees of freedom and still achieve full diversity order $L+1$ with a linear receiver.

As before, the best relaying strategy could be identified by minimizing the outage gain as a function of α. Unfortunately, even if (4.10) gives a closed-form expression to compute κ, the analytical study of this function is not an easy task and one should locate the optimal α numerically, by means of curves as those in Figure 4.2.

On the other hand, the sub-optimality of linear receivers is evident for $\alpha > \alpha_{th}$. In this case, the source signal is excessively compressed at the relays and the information recovered by the LMMSE filter is not enough to improve diversity with respect to the direct link.

4.1.3 Diversity–Multiplexing Tradeoff

This section is devoted to the derivation of the DMT for the presented scheme, for both ML and LMMSE receivers. For this purpose, it results simpler to start from the LMMSE case. For the ML receiver, indeed, we are only able to give bounds on the real DMT, due to the lack of a closed-form expression for the outage gain.

DMT for the LMMSE receiver

According to Theorem 4.2, the diversity behavior of the LMMSE receiver depends on whether α is lower or higher than the threshold value α_{th}. Obviously, this fact also has influence on the DMT. When computing the DMT, we might as well consider $\alpha_{th} = 1$, since Q tends to infinity when $R = r \ln \rho$ and $\rho \to +\infty$.

Now, noting that[6]

$$Q \doteq \rho^{\frac{1+\alpha}{\alpha} r}$$

and applying some algebra to the right-hand sides of (4.10) and (4.11), the DMT for DF relays with LMMSE reception can be written as

$$d_{\text{LMMSE}}(r) = \begin{cases} (L+1)\left(1 - \dfrac{1+\alpha}{\alpha} r\right) & \text{for } \alpha \leq 1, \\ 1 - \dfrac{1+\alpha}{\alpha} r & \text{for } \alpha > 1. \end{cases}$$

It is straightforward to notice that the first line is maximized by $\alpha = 1$ (i.e. $K = N$, the two transmission phases have the same duration), while the second one is only maximized by letting $\alpha \to +\infty$ (i.e. $N = 0$, the relays never transmit and communications are supported by the direct link alone). Equivalently, the best DMT is

$$d^*_{\text{LMMSE}}(r) = \begin{cases} (L+1)(1-2r) & \text{for } r \leq \dfrac{L}{2L+1}, \\ 1 - r & \text{otherwise}, \end{cases} \tag{4.12}$$

which is represented by the line connecting the three points $(0, L+1), A = (\frac{L}{2L+1}, \frac{L+1}{2L+1})$ and $(1, 0)$ in Figure 4.3. In practice, according to the DMT, the best transmission strategy is to relay with a coding rate $\alpha = 1$ when the multiplexing gain is lower that $L/(2L+1)$. Conversely, it is preferable to skip the relaying phase for higher values of r.

[6] We recall the definition of exponential equality from [46]: we write $f(z) \doteq z^d$ if $\lim_{z \to +\infty} \frac{\ln f(z)}{\ln z} = d$.

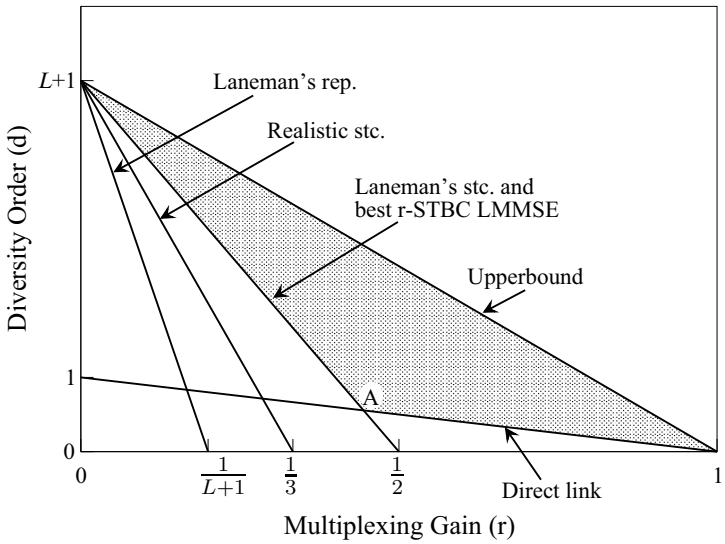


Figure 4.3: DMT for DF relays, different relaying protocols. The best DMT for random STBC's with ML receiver is included in the shadowed region.

DMT for the ML receiver

As mentioned before, the exact DMT for DF relays with ML reception cannot be easily derived in closed form because the outage gain is only known as the solution to the integral in (4.5). For this reason, we will confine it between two bounds.

Proposition 4.1. *The DMT for DF relays with ML reception is upper-bounded as follows:*

$$d_{\mathrm{ML}}(r) \leq (L+1)\left(1 - \frac{1+\alpha}{\alpha} r\right).$$

Proof. By studying the first derivative of (4.3) as a function of $\beta_{\mathcal{L}'}$ (considered this time as a stand-alone variable, independent of α and the channels), we readily see that

$$I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}} < I_{\mathrm{UP}}^{(\mathcal{L}')} = \frac{\alpha}{1+\alpha} \ln\left(1 + \rho|h_s|^2 + \frac{\rho}{\alpha} \sum_{l=1}^{|\mathcal{L}'|} |h_{dl}|^2\right).$$

Observe that the right-hand side is the spectral efficiency that we would obtain if it were always possible to build N orthogonal sequences of length K, whatever values of K and N. Using similar arguments as in Appendices 4.A.1 and 4.A.2, it is easy to prove that

$$\lim_{\rho \to +\infty} \rho^{|\mathcal{L}'|+1} \Pr\left[I_{\mathrm{UP}}^{(\mathcal{L}')} < R\right] = \zeta_s \zeta_d^{|\mathcal{L}'|} \frac{\alpha^{|\mathcal{L}'|} Q^{|\mathcal{L}'|+1}}{(|\mathcal{L}'|+1)!}.$$

Using the exponential relation $Q \doteq \rho^{\frac{1+\alpha}{\alpha}r}$, one can show

$$\begin{aligned} d_{\mathrm{ML}}(r) &= -\lim_{\rho\to+\infty} \frac{\ln P_{out}(r\ln\rho)}{\ln\rho} \\ &\leq -\lim_{\rho\to+\infty} \frac{\ln \sum_{\mathcal{L}'} \Pr[\mathcal{L}' \text{is the decoding set}] \Pr\left[I_{\mathrm{UP}}^{(\mathcal{L}')} < R\right]}{\ln\rho} \\ &= (L+1)\left(1-\frac{1+\alpha}{\alpha}r\right), \end{aligned}$$

which is the desired upper-bound. □

The simplest lower-bound (and the best one among those we could figure out) is the DMT of the LMMSE receiver, namely $d_{\mathrm{ML}}(r) \geq d_{\mathrm{LMMSE}}(r)$. Note that the two bounds coincide for $\alpha \leq 1$. Hence, in that situation,

$$d_{\mathrm{ML}}(r) = (L+1)\left(1-\frac{1+\alpha}{\alpha}r\right) \qquad \text{when } \alpha \leq 1.$$

With a similar reasoning as above, the best DMT for the ML receiver can be shown to satisfy

$$d^*_{\mathrm{LMMSE}}(r) \leq d^*_{\mathrm{ML}}(r) \leq (L+1)(1-r),$$

and falls within the shadowed area of Figure 4.3.

4.1.4 Numerical results

In this section, the analytical derivations made above are complemented with some numerical results. First, let us justify the choice of using the asymptotic deterministic spectral efficiency as an approximation of the finite reality. We have already seen with Figure 3.4 that the asymptotic spectral efficiency is an excellent approximation of the mean behavior of the system. As a further step, the points in Figure 4.4 represent the average spectral efficiency (normalized with respect to the limit for $K = \alpha N \to +\infty$), and its standard deviation, obtained over one thousand different realization of the coding matrices. The coding rate $\alpha = 4/3$ is kept constant over the whole experiment while the matrix dimensions grow proportionally to M, namely $K = 4M$ and $N = 3M$. Observe that, for codes of realistic dimensions, the error is already acceptably small. For instance, at $M = 10$, the standard deviation is already lower than 2%. Furthermore, a channel coherence time of $K+N = 70$ channel accesses is not unrealistic: for example, a normal data burst in the TDMA mode of GSM is 148 bits long[7], which corresponds to 74

[7]The data burst is made of 3 tail bits, 57 data bits, 1 toggle bit, 26 training bits, 1 toggle bit, 57 data bits and 3 tail bits, which sums a total of 148 bits. Between two consecutive bursts there is a guard time of 8.25 bits, leading to a total slot length of 156.25 bits or, equivalently, a slot time of 0.577 ms. The resulting raw bit rate is around 270 kilobits per second.

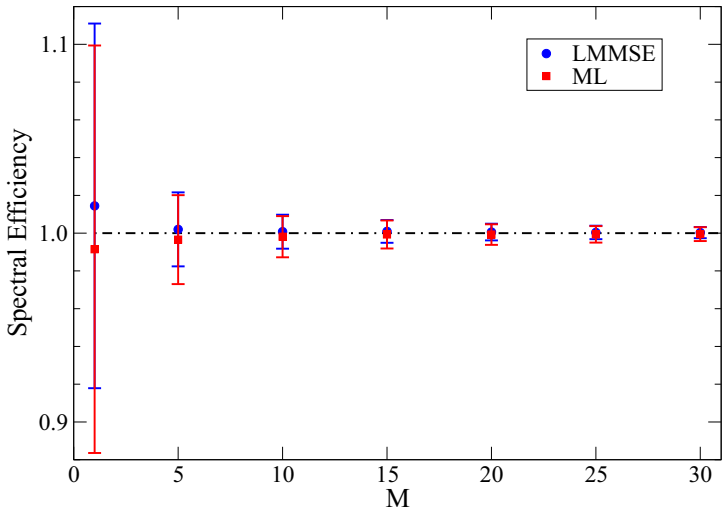


Figure 4.4: Average spectral efficiency and standard deviation (normalized with respect to the asymptotic value) over 1000 realizations of the coding matrices. The coding rate $\alpha = 4/3$ is kept constant while $K = 4M$ and $N = 3M$.

Gaussian-minimum-shift-keying symbols (see, e.g., [90, Chapter 3]). Quick convergence to the limit motivates our decision to carry out the large-SNR outage analysis based on the asymptotic spectral efficiencies (4.3) and (4.7). Figure 4.5 compares simulation-based outage probability curves with the respective high-SNR approximations.

Comparison to previous results

It is also interesting to compare the performance of the presented randomized LD-STBC with those of two existing protocols, namely those presented by J. N. Laneman *et al.* in [32]:

1. a TDMA-based repetition protocol where the relays in the decoding set simply retransmit the K-symbol source vector $\mathbf{s}$. Since relays transmit one at a time (TDMA, time division multiple access), the total duration of the relaying phase is LK channel accesses. According to our notation and power constraint, the spectral efficiency is shown to be
$$I_{\text{rep}}^{(\mathcal{L}')} = \frac{1}{L+1} \ln\Big(1 + \rho|h_s|^2 + \rho \sum_{l \in \mathcal{L}'} |h_{dl}|^2\Big),$$
for a given decoding set $\mathcal{L}'$;

2. an orthogonal-STC-based protocol where relays re-encode the source message according to a rate-1 (K information symbols over K channel

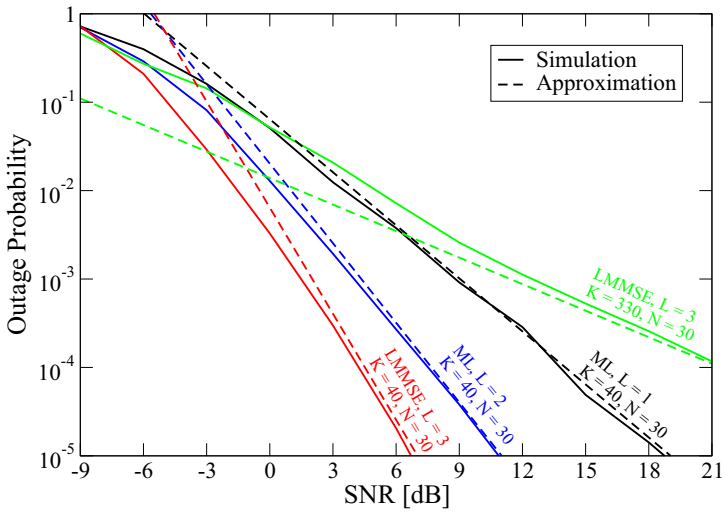


Figure 4.5: Comparison between simulation outage and large-SNR approximation for unitary channel variances, $z = 1$ and $R = 0.1$ nat/s/Hz.

accesses), full diversity-exploiting STC. Relays can then transmit at the same time and, hence, the relaying phase lasts K channel accesses as the source one. For the decoding set $\mathcal{L}'$, the resulting spectral efficiency is

$$I_{\text{stc}}^{(\mathcal{L}')} = \frac{1}{2}\ln\Big(1+\rho|h_s|^2\Big) + \frac{1}{2}\ln\Big(1+\rho\sum_{l\in\mathcal{L}'}|h_{dl}|^2\Big).$$

Both protocols achieve full diversity and the outage gains are, respectively:

$$\kappa_{\text{rep}} = \Big(e^{(L+1)R}-1\Big)^{L+1}\zeta_s\sum_{l=0}^{L}\binom{L}{l}\frac{\zeta_d^l(\zeta_u z^{-1})^{L-l}}{(l+1)!}$$

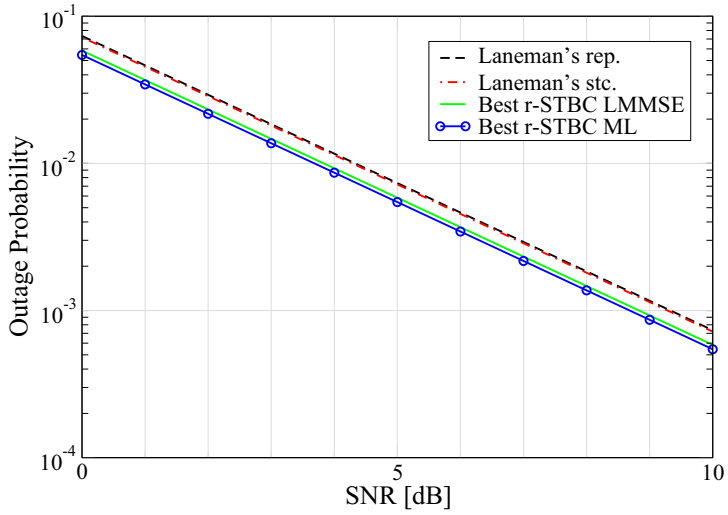
and

$$\kappa_{\text{stc}} = \Big(e^{2R}-1\Big)^{L+1}\zeta_s\sum_{l=0}^{L}\binom{L}{l}\zeta_d^l(\zeta_u z^{-1})^{L-l}A_l\Big(e^{2R}-1\Big),$$

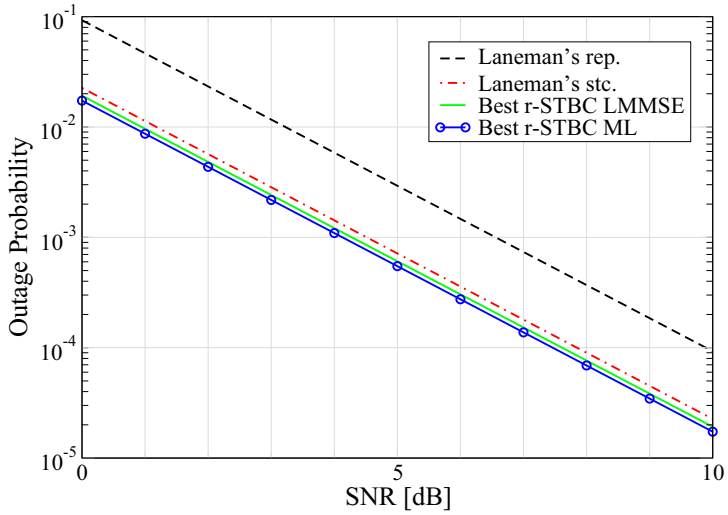
where

$$A_l(t) = \frac{1}{(l-1)!}\int_0^1 \frac{w^{l-1}(1-w)}{1+tw}\,\mathrm{d}w \tag{4.13}$$

and $A_0(t) = 1$. The coefficients ζ_s, ζ_u and ζ_d are defined by (4.6). The outage gains for one, two and three relays are reported in Figure 4.1 (all the system assumptions are the same as those of the curves presented above). Note that Laneman's deterministic scheme offers some benefit at its specific



(a) 1 relay



(b) 2 relays

Figure 4.6: Outage-probability comparison (large-SNR approximations) between Laneman's protocols and random STBC's, for 1 (a) and 2 (b) relays. The system parameters are set as in Figures 4.1 and 4.2.

coding rate $\alpha = 1$. However, randomized LD-STBC can achieve lower outage gains at higher values of α, due to a better exploitation of the degrees of freedom. Figure 4.6 compares the high-SNR approximations of the outage probability obtained by Laneman with those derived above for the randomized LD-STBC (the coding rate α is chosen to minimize the outage gain). Classic space-time codes lose something more than 0.5 dB with respect to the proposed randomized coding scheme.

Reference [32] also gives the DMT for both schemes, namely:

$$d_{\text{rep}}(r) = (L+1)[1-(L+1)r]$$

and

$$d_{\text{stc}}(r) = (L+1)(1-2r), \tag{4.14}$$

which are depicted in Figure 4.3 for comparison with the DMT's of the LD-STBC schemes computed above.

The main disadvantage of the repetition scheme is, thus, that the maximum achievable multiplexing gain, i.e. $r = 1/(L+1)$ is inversely proportional to the total number of relays. TDMA-based relaying protocols achieve full diversity at the cost of a huge waste of degrees of freedom. The suboptimality is evident in Figure 4.3, where the the DMT of the TDMA-based relaying protocol is shown to be always lower than the DMT of all the other relaying schemes.

Note that result in (4.14) is given by Laneman as a lower-bound, i.e. $d_{\text{stc}}(r) \geq (L+1)(1-2r)$. However, the equality can be shown straightforwardly after solving the integral in (4.13):

$$\begin{aligned}
\int_0^1 \frac{w^{l-1}(1-w)}{1+tw}\,\mathrm{d}w &= \\
&= -\frac{1}{t}\int_0^1 w^{l-1}\,\mathrm{d}w + (t+1)\sum_{j=2}^{l}\left(-\frac{1}{t}\right)^j \int_0^1 w^{l-j}\,\mathrm{d}w - \\
&\qquad - (t+1)\left(-\frac{1}{t}\right)^l \int_0^1 \frac{1}{1+tw}\,\mathrm{d}w \\
&= -\frac{1}{tl} + (t+1)\sum_{j=2}^{l}\left(-\frac{1}{t}\right)^j \frac{1}{l-j+1} + (t+1)\left(-\frac{1}{t}\right)^{(l+1)} \ln(t+1).
\end{aligned}$$

The DMT of Laneman's STC-based protocol is hence equal to the best DMT obtained with LD-STBC and LMMSE receiver (see (4.12) and Figure 4.3).

As a last comment, let us remark that Laneman's results about STC relays are optimistic: indeed, they are obtained by assuming an ideal code with rate 1. To the best of the author's knowledge, when considering complex

constellations, the only orthogonal STBC with unitary rate is the Alamouti code for two transmit antennas. For instance, STC's from orthogonal designs [8], cited by Laneman in his paper, provide codes with rate 3/4 for either 3 or 4 transmitters and codes with rate 1/2 for higher numbers of transmitters. This means that, for a high number of relays ($L \geq 5$), a more realistic DMT for Laneman's orthogonal-STC protocol would be

$$d_{stc}(r) = (L+1)(1-3r).$$

This DMT is also reported in Figure 4.3 (the "Realistic stc." line) which clearly shows that, in terms of DMT, the relaying schemes based on randomized LD-STBC performs better than those employing deterministic space-time codes based on orthogonal designs.

4.2 LD-STBC in AF relaying

The techniques introduced in the previous section to analyze the outage probability of DF relaying are now applied to the AF case. Recall that all AF relays participate in the second phase and, hence, there is no need to condition the outage probability on the decoding set. However, things become quite more complicated.

For DF relaying, indeed, the presented results mainly agree with the intuition that reliable communications can be established whenever at least one relay has good channels both from the source (the relay belongs to the decoding set) and to the destination. The only exception arises when LMMSE detection is combined to high coding rates (Theorem 4.2). Conversely, in AF relaying, the intuition is lost due to the correlation introduced by the randomized LD-STBC and by the forwarded colored noise. As expressed by the following results, the coding rate α plays a fundamental role.

4.2.1 The ML receiver

As usual, we first consider the ML receiver, which was thoroughly described in Section 3.2. We report here the resulting expression for the asymptotic spectral efficiency, namely

$$I_{\mathrm{AF,ML}}^{\mathrm{iid}} = \frac{1}{1+\alpha}\left[\alpha \ln\left(1+\rho|h_s|^2\right) + \right.$$
$$\left. + \alpha \sum_{l=1}^{L} \ln \frac{1+\lambda_l \phi_1}{1+\frac{z\rho|h_{dl}|^2}{\alpha(1+z\rho|h_{ul}|^2)}\phi_2} + \ln\frac{\phi_2}{\phi_1} + z(\phi_1-\phi_2)\right]. \quad (4.15)$$

Recall that the coefficients $\{\lambda_l\}_{l=1}^{L}$ are the L eigenvalues of the matrix $\mathbf{\Psi}\mathbf{\Psi}^H + \frac{z\rho}{1+\rho|h_s|^2}\mathbf{\Psi}\mathbf{h}_u\mathbf{h}_u^H\mathbf{\Psi}^H$, while ϕ_1 and ϕ_2 are the unique positive so-

lutions to

$$\phi_1 = \left(z + \alpha \sum_{l=1}^{L} \frac{\lambda_l}{1 + \lambda_l \phi_1} \right)^{-1} \tag{4.16}$$

and

$$\phi_2 = \left(z + \alpha \sum_{l=1}^{L} \frac{z\rho |h_{dl}|^2}{\alpha(1 + z\rho |h_{ul}|^2) + z\rho |h_{dl}|^2 \phi_2} \right)^{-1}, \tag{4.17}$$

respectively. To simplify notation, we introduced $z = \sigma_d^2/\sigma_u^2$. Furthermore, as in the DF case, the relay gains $\left\{ |g_l|^2 : l = 1, \ldots, L \right\}$, have been set to fix the mean relay symbol power to P_s, namely

$$|g_l|^2 = \frac{z\rho}{\alpha(1 + z\rho |h_{ul}|^2)}. \tag{4.18}$$

The large-SNR outage analysis of the asymptotic spectral efficiency above leads to the following result:

Theorem 4.3. *Consider the system described in Section 3.1.1, with L AF relays employing the randomized LD-STBC. Assume maximum-likelihood reception with full channel state information, resulting in the asymptotic spectral efficiency (4.15). Then, the diversity order depends on the coding rate α as*

$$d = \begin{cases} L + 1 & \text{for } \alpha < \dfrac{1}{L-1}, \\ M + 1 & \text{for } \dfrac{1}{M} \leq \alpha < \dfrac{1}{M-1}, \end{cases}$$

with $M = 1, \ldots, L-1$. The outage gain κ can be hence computed numerically according to its definition (4.1).

Proof. Unfortunately, the formal proof is not completed yet. More specifically, some convergence results are proven pointwise only and not uniformly. See Appendix 4.B for more details. □

Even if the mathematical derivations are very tedious, the idea behind the proof is very simple. Following the example of [33], a state is assigned to each relay according to the quality of both its channels. Then, for all the possible resulting partitions, the spectral efficiency is bounded to justify the applicability of the Lebesgue's dominated convergence theorem (LDCT), as in Appendix 4.A.1, and the probability of the outage event is computed for each possible partition. According to the value assumed by α, it turns out that only some of the partitions bring meaningful contributions to the outage probability, while the others can be neglected. These observations yield some interesting insights on the results of Theorem 4.3:

- When $\alpha = K/N < 1/(L-1)$, the contributions to the outage probability only come from the partitions where each relay has either one of the channels (uplink or downlink, but not both) in outage. The interpretation is the following. When using orthogonal spreading sequences, whose length would be at least $N = LK$, the system is equivalent to a transmission over $L+1$ parallel channels (same as time-division, see, e.g., [32]), also counting the direct link. With random, non-orthogonal signatures with $N > K(L-1)$, the ML receiver still sees $L+1$ parallel channels. Thus, the system undergoes an outage event only when all the relays cannot support the required rate, because either of their links cannot. The probability that deep fading affects both links for some relays and only one link for all the others is negligible. The resulting outage gain can be computed according to its definition. Indeed, applying the LDCT to (4.1), we can write

$$\kappa = \lim_{\rho\to+\infty} \rho^{L+1} P_{out}(R) = \sum_{k=0}^{L} \binom{L}{k} \zeta_s \zeta_u^k \zeta_d^{L-k} z^{-L} \cdot \\ \cdot \int_{\mathbb{R}_+^{L+1}} \mathbb{1}\left\{ f_k\left(a, \mathbf{b}_1^k, \mathbf{c}_{k+1}^L\right) < (1+\alpha)R \right\} \mathrm{d}a\, \mathrm{d}\mathbf{b}_1^k\, \mathrm{d}\mathbf{c}_{k+1}^L, \quad (4.19)$$

where we used the definition

$$f_k\left(a, \mathbf{b}_1^k, \mathbf{c}_{k+1}^L\right) = \\ = \alpha \ln\left[1 + a + \sum_{l=1}^{k} b_l + \frac{\bar{\phi}_1}{\alpha} \sum_{l=k+1}^{L} c_l\right] + \\ - (1-\alpha k)\left[1 - \ln(1-\alpha k)\right] - (1-\alpha k)\ln\left(z\bar{\phi}_1\right) + z\bar{\phi}_1.$$

Furthermore, writing

$$\bar{\lambda} = \frac{1}{\alpha} \frac{\sum_{l=k+1}^{L} c_l}{1 + a + \sum_{l=1}^{k} b_l},$$

$\bar{\phi}_1$ is the unique positive solution to

$$\bar{\phi}_1 = \begin{cases} \frac{\left[z-(1-\alpha(k+1))\bar{\lambda}\right]}{2z\bar{\lambda}} \left[-1 \pm \sqrt{1 + \frac{4z\bar{\lambda}(1-\alpha k)}{\left[z-(1-\alpha(k+1))\bar{\lambda}\right]^2}}\right] & k < L \\ \frac{1-\alpha L}{z} & k = L. \end{cases} \quad (4.20)$$

Generally, the integral in (4.19) has to be computed numerically;

- If $1/M \le \alpha < 1/(M-1)$, with $1 \le M < L$, the contributions to the outage probability are brought about by the cases with exactly

M relays with deep fading in the uplink. Their downlinks should not experience outage. Same thing must hold for both channels of the other $L-M$ relays. This means that the source–relay link is the dominant one in terms of outage:

1. If the source symbols are received by all relays with a high SNR, then it is always possible to convey information to the destination (except, of course, the case where all the downlinks and the direct link are corrupted. This case, however, has very low probability and should be accounted for only when $\alpha < 1/(L-1)$, see above);
2. Conversely, as α increases, there is a reduction in the minimum number of badly faded uplink channels which is sufficient to generate the outage event. For instance, when the direct link is corrupted and $\alpha \geq 1$, it is enough that one single relay, out of L, receives a low-SNR signal for data transmission to fail.

The intuition for the second point is that the non-orthogonal random coding employed in the presented scheme correlates the contributions of the relays, as can be noticed in (4.15): the quantities ϕ_1, ϕ_2 and all λ_l's depend on the channels of the totality of the links (direct, source–relay, relay–destination) of the system. Analogously to the previous case, the outage gain is given by the computation of the following integral (which needs to be computed numerically):

$$\kappa = \lim_{\rho\to+\infty} \rho^{M+1} P_{out}(R) = \binom{L}{M} \zeta_s \zeta_u^M z^{-M} \cdot \\ \cdot \int_{\mathbb{R}_+^{L+M+1}} \mathbb{1}\left\{ g(a, \mathbf{b}_1^M, \mathbf{y}_1^L) < (1+\alpha)R \right\} \mathrm{d}a \, \mathrm{d}\mathbf{b}_1^M f_{|h_d|^2}(\mathbf{y}_1^L) \, \mathrm{d}\mathbf{y}_1^L,$$

where we defined

$$g(a, \mathbf{b}_1^M, \mathbf{y}_1^L) = \alpha \ln(1+a) + \ln\frac{\theta_2}{\theta_1} + \alpha \sum_{l=L-M}^{L} \ln(1+\nu_l\theta_1) \\ - \alpha \sum_{k=1}^{M} \ln\left[1 + z\frac{y_k}{\alpha(1+b_k)}\theta_2\right].$$

The quantities θ_1 and θ_2 are the positive solutions to the equations

$$\alpha(M+1) - 1 = \alpha \sum_{l=L-M}^{L} \frac{1}{1+\nu_l\theta_1}, \qquad (4.21)$$

$$\alpha M - 1 = \alpha \sum_{k=1}^{M} \frac{1}{1+\frac{zy_k}{\alpha(1+b_k)}\theta_2},$$

with $\{\nu_l\}_{l=L-M}^{L}$ being the $M+1$ solutions to the following equation in ν

$$\sum_{k=1}^{M} \frac{b_k y_k}{(1+b_k)\nu - \frac{z y_k}{\alpha}} + \frac{1}{\nu} \sum_{k=M+1}^{L} y_k = \frac{\alpha(1+a)}{z}.$$

4.2.2 The LMMSE receiver

According to the power constraint considered in this chapter, which implies the relay gain in (4.18), the asymptotic LMMSE output SINR is given by

$$SINR_{\mathrm{AF}}^{\mathrm{iid}} = \rho|h_s|^2 + z^2\rho^2 \sum_{l=1}^{L} \frac{|h_{dl}h_{ul}|^2}{z\rho|h_{dl}|^2 + \alpha\big(1+z\rho|h_{ul}|^2\big)\frac{1}{\phi_1}}, \tag{4.22}$$

with ϕ_1 as defined in (4.16). Recalling the definition of Q as in (4.8), the large-SNR analysis of the outage probability $\Pr\big[I_{\mathrm{AF,LMMSE}}^{\mathrm{iid}} < R\big] = \Pr\big[SINR_{\mathrm{AF}}^{\mathrm{iid}} < Q\big]$ leads to the following result:

Theorem 4.4. *Consider the system described in Section 3.1.1, with L AF relays employing the randomized LD-STBC. Assume LMMSE reception with full channel state information, resulting in the asymptotic SINR (4.22). Then, the diversity order depends on the coding rate α as*

$$d = \begin{cases} L+1 & \text{for } \alpha \le \dfrac{1}{L}, \\ M+1 & \text{for } \dfrac{1}{M+1} < \alpha \le \dfrac{1}{M}, \\ 2 & \text{for } 1 < \alpha \le \alpha_{th}, \\ 1 & \text{for } \alpha > \alpha_{th}, \end{cases}$$

with $M = 1, \ldots, L-1$ and where α_{th} is the unique positive solution to (4.9). The outage gain κ can be hence computed numerically according to its definition (4.1).

Proof. As for the ML case, some convergence results are proven pointwise only and not uniformly. See Appendix 4.C for more details. □

This result is very similar to the one obtained in the previous section for the ML receiver. Nevertheless, some aspects are worth remarking:

- As before, the diversity order decreases as the coding rate α becomes larger. Since the LMMSE filter is suboptimal with respect to the ML receiver, this fact is intuitive enough. Note, however, that the conditions are stricter in this case. For instance, full diversity is achieved only if $\alpha < \frac{1}{L}$, as opposed to $\alpha < \frac{1}{L-1}$. In other words, the suboptimality of the linear receiver does not affect only the outage gain, but also the capacity of exploiting spatial diversity. This is true at any value of the coding rate;

- For $\alpha \leq \frac{1}{L}$ the transmission is in outage if and only if all relays suffer a deep fade in either of its channels. The cases where neither channels of a same relay are good are negligible. The diversity order is given by the following integral:

$$\kappa = \lim_{\rho \to +\infty} \rho^{L+1} P_{out}(R) = \sum_{k=0}^{L} \binom{L}{k} \zeta_s \zeta_u^k \zeta_d^{L-k} z^{-L} \cdot \\ \cdot \int_{\mathbb{R}_+^{L+1}} \mathbb{1}\left\{ a + \sum_{l=1}^{k} b_l + \frac{\bar{\phi}_1}{\alpha} \sum_{l=k+1}^{L} c_l < Q \right\} \mathrm{d}a \, \mathrm{d}\mathbf{b}_1^k \, \mathrm{d}\mathbf{c}_{k+1}^L,$$

  with $\bar{\phi}_1$ as defined in (4.20);

- For $\frac{1}{M+1} < \alpha \leq \frac{1}{M}$, with $M = 1, \ldots, L-1$, the only meaningful contributions are those associated to the cases where exactly M uplink channels experience a deep fade, while all the others have good quality. As for the ML receiver, then, the critical link is the source–relay one. Due to the signature correlation, a small number of relays receiving a low-quality signal can slow down the transmission, even when all the relay–destinations links are good. By means of the LDCT, the outage gain is given by the solution to the integral:

$$\kappa = \lim_{\rho \to +\infty} \rho^{M+1} P_{out}(R) = \binom{L}{M} \zeta_s \zeta_u^M z^{-M} \cdot \\ \cdot \int_{\mathbb{R}_+^{L+M+1}} \mathbb{1}\left\{ g(a, \mathbf{b}_1^M, \mathbf{y}_1^L) < Q \right\} \mathrm{d}a \, \mathrm{d}\mathbf{b}_1^M f_{|h_d|^2}(\mathbf{y}_1^L) \, \mathrm{d}\mathbf{y}_1^L,$$

  where we have defined

$$g(a, \mathbf{b}_1^M, \mathbf{y}_1^L) = a + z\theta_1 \sum_{l=1}^{M} \frac{b_l y_l}{z\theta_1 y_l + \alpha(1+b_l)} + z\frac{\theta_1}{\alpha} \sum_{l=M+1}^{L} y_l$$

  and θ_1 is defined in (4.21);

- For $1 < \alpha \leq \alpha_{th}$, the diversity order is two, meaning that it is enough that one single link is corrupted to interrupt the communication, besides the direct one. Discarding the situations whose probability decays too fast as $\rho \to +\infty$, the outage gain is given by the cases where one single relay has a bad uplink channel and $L-1$ relays have bad downlink channels, namely

$$\begin{aligned} \kappa &= \lim_{\rho \to +\infty} \rho^2 P_{out}(R) \\ &= L\zeta_s\zeta_u z^{-1} \int_{\mathbb{R}_+^2} \mathbb{1}\left\{ a + \frac{(1+a)b}{\alpha(1+a) + (\alpha-1)b} < Q \right\} \mathrm{d}a \, \mathrm{d}b \\ &= L\frac{\zeta_s\zeta_u}{z}\left\{ \frac{Q^2}{2} - \frac{\alpha-1}{\alpha}(Q+1)\left[Q + \frac{Q+1}{\alpha} \ln\left(1 - \alpha\frac{Q}{Q+1} \right) \right] \right\}. \end{aligned}$$

Only in the case where $L = 1$, another term has to be added, namely

$$\frac{\zeta_s\zeta_d}{z}\int_{\mathbb{R}_+^2}\mathbb{1}\{a<Q\}\mathbb{1}\left\{c<\frac{\alpha z(Q+1)(Q-a)}{1+a-(\alpha-1)(Q-a)}\right\}\mathrm{d}a\,\mathrm{d}c = \\ = \zeta_s\zeta_d(Q+1)\left[-Q-\frac{Q+1}{\alpha}\ln\left(1-\alpha\frac{Q}{Q+1}\right)\right],$$

which corresponds to the situation where the relay sees a good channel from the source but undergoes a deep fade on the downlink channel;

- For $\alpha > \alpha_{th}$, the information is carried only by the direct source–destination link. Relays reduce the outage gain, though. Indeed, one has

$$\begin{aligned}\kappa &= \lim_{\rho\to+\infty}\rho P_{out}(R)\\ &= \zeta_s\int_{\mathbb{R}_+}\mathbb{1}\left\{a<Q-\frac{Q+1}{\alpha}\right\}\mathrm{d}a\\ &= \zeta_s\left(Q-\frac{Q+1}{\alpha}\right),\end{aligned}$$

which vanishes as $\alpha \to \alpha_{th}$. Moreover, comparing the last outage-gain expression with (4.11), it is evident that there is no difference between AF and DF relays in this situation.

Figure 4.7 compares the diversity order d achieved by the two receivers as a function of the coding rate α when $L = 5$.

4.2.3 The single-relay case

To give a general idea of the structure of the proofs to the previous results (given in the Appendices 4.B and 4.C), we sketch here the case with one single relay. This case is much simpler than the general one, mainly because ϕ_1 and ϕ_2 can be computed explicitly as the positive root of a second-order polynomial and the close-form expression for λ is (suppressing indices)

$$\lambda = \left(\frac{z\rho|h_u|^2}{1+\rho|h_s|^2}+1\right)\frac{z\rho|h_d|^2}{\alpha(1+z\rho|h_u|^2)}.$$

In what follows, we prove (without being completely rigorous) that

$$\lim_{\rho\to+\infty}\rho^d P_{out}(R) = \kappa,$$

where the outage gain κ is finite and strictly positive and the diversity order d is equal to either one or two, according to the results for the general case, given in Theorems 4.3 and 4.4.

First, we define the four events

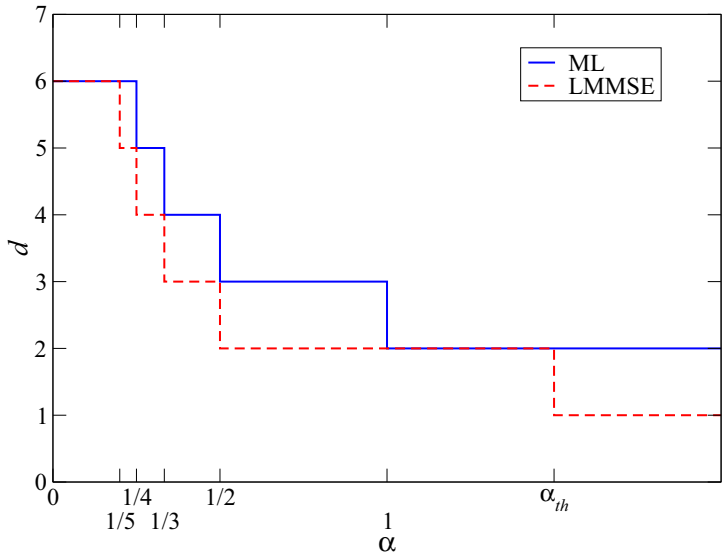


Figure 4.7: Diversity order as a function of the coding rate α for $L = 5$ AF relays.

$\mathcal{E}_1$ the system is in outage (i.e. $I^{\text{iid}}_{\text{AF,ML}} < R$ or $SINR^{\text{iid}}_{\text{AF}} < Q$, according to the receiver) and $|h_u|^2 = O(\rho^{-1})$, $|h_d|^2 > O(\rho^{-1})$ for $\rho \to +\infty$;

$\mathcal{E}_2$ the system is in outage and $|h_u|^2 > O(\rho^{-1})$, $|h_d|^2 = O(\rho^{-1})$ for $\rho \to +\infty$;

$\mathcal{E}_3$ the system is in outage and $|h_u|^2 > O(\rho^{-1})$, $|h_d|^2 > O(\rho^{-1})$ for $\rho \to +\infty$;

$\mathcal{E}_4$ the system is in outage and $|h_u|^2 = O(\rho^{-1})$, $|h_d|^2 = O(\rho^{-1})$ for $\rho \to +\infty$.

Note that the four events are disjoint and, thus, $P_{out}(R) = \sum_{i=1}^{4} \Pr[\mathcal{E}_i]$. Then, we study the probability of each event, namely

$$\lim_{\rho\to+\infty} \rho^d \Pr[\mathcal{E}_i] = \\ = \lim_{\rho\to+\infty} \rho^d \int_{\mathbb{R}^3_+} \mathbb{1}\{\mathcal{E}_i\} f_{|h_s|^2}(x) f_{|h_u|^2}(w) f_{|h_d|^2}(y)\, \mathrm{d}x\, \mathrm{d}w\, \mathrm{d}y. \quad (4.23)$$

Event $\mathcal{E}_1$

With the change of variables $a = \rho x$ and $b = z\rho w$, one has

$$\lambda = \left(\frac{b}{1+a} + 1\right) \frac{z\rho y}{\alpha(1+b)} \to +\infty$$

when $\rho \to +\infty$. As a result, it is straightforward to show that

$$\phi_1, \phi_2 \to \begin{cases} 0 & \text{for } \alpha > 1; \\ \dfrac{1-\alpha}{z} & \text{for } \alpha \leq 1 \end{cases}$$

and that, for $\alpha > 1$,

$$\lambda\phi_1, \frac{z\rho y}{\alpha(1+b)}\phi_2 \to \frac{1}{\alpha-1}.$$

Knowing these facts, the outage event $\mathcal{E}_1$ can be shown to be equivalent to

$$\mathcal{E}'_{1,\mathrm{ML}} : \alpha\ln(1+a) + \min\{1,\alpha\}\ln\left(\frac{b}{1+a}+1\right) < (1+\alpha)R,$$
$$\mathcal{E}'_{1,\mathrm{LMMSE}} : a + \frac{b(1+a)}{1+a+\max\{0,\alpha-1\}(1+a+b)} < Q,$$

in the limit for $\rho \to +\infty$, for the ML receiver and the LMMSE receiver, respectively. It is evident that $\mathbb{1}\{\mathcal{E}'_{1,\mathrm{ML}}\}$ represents a finite volume and (4.23) can be computed by means of the LDCT:

$$\begin{aligned}\lim_{\rho\to+\infty}\rho^2\Pr[\mathcal{E}_{1,\mathrm{ML}}] &= \\ = \frac{1}{z}\lim_{\rho\to+\infty}\int_{\mathbb{R}_+^3}\mathbb{1}\{\mathcal{E}_{1,\mathrm{ML}}\}f_{|h_s|^2}\left(\frac{a}{\rho}\right)f_{|h_u|^2}\left(\frac{b}{z\rho}\right)f_{|h_d|^2}(y)\,\mathrm{d}a\,\mathrm{d}b\,\mathrm{d}y &= \\ = \frac{\zeta_s\zeta_u}{z}\int_{\mathbb{R}_+^2}\mathbb{1}\{\mathcal{E}'_{1,\mathrm{ML}}\}\,\mathrm{d}a\,\mathrm{d}b.\end{aligned}$$

On the other hand, after some algebra, $\mathcal{E}'_{1,\mathrm{LMMSE}}$ can be rewritten in three different ways according to the value of the coding rate α:

- when $\alpha \le 1$, one has $\mathcal{E}'_{1,\mathrm{LMMSE}} : a + b < Q$ and

$$\begin{aligned}\lim_{\rho\to+\infty}\rho^2\Pr[\mathcal{E}_{1,\mathrm{LMMSE}}] &= \frac{\zeta_s\zeta_u}{z}\int_{\mathbb{R}_+^2}\mathbb{1}\{\mathcal{E}'_{1,\mathrm{LMMSE}}\}\,\mathrm{d}a\,\mathrm{d}b \\ &= \frac{\zeta_s\zeta_u}{z}\frac{Q^2}{2};\end{aligned}$$

- when $1 < \alpha \le \alpha_{th}$, the outage event is

$$\mathcal{E}'_{1,\mathrm{LMMSE}} : a < Q \text{ and } b < \frac{\alpha(Q-a)(1+a)}{1+Q-\alpha(Q-a)}.$$

Thus, the contribution to the outage gain is given by

$$\begin{aligned}\lim_{\rho\to+\infty}\rho^2\Pr[\mathcal{E}_{1,\mathrm{LMMSE}}] &= \frac{\zeta_s\zeta_u}{z}\int_{\mathbb{R}_+^2}\mathbb{1}\{\mathcal{E}'_{1,\mathrm{LMMSE}}\}\,\mathrm{d}a\,\mathrm{d}b = \\ = \frac{\zeta_s\zeta_u}{z}\left\{\frac{Q^2}{2} + \frac{\alpha-1}{\alpha}(Q+1)\left[-Q - \frac{Q+1}{\alpha}\ln\left(1-\alpha\frac{Q}{Q+1}\right)\right]\right\};\end{aligned}$$

- if $\alpha > \alpha_{th}$, the outage event does not depend on b, namely

$$\mathcal{E}'_{1,\mathrm{LMMSE}} : a < Q - \frac{Q+1}{\alpha}.$$

Hence,

$$\lim_{\rho\to+\infty} \rho \Pr[\mathcal{E}_{1,\mathrm{ML}}] = \\ = \frac{1}{z} \lim_{\rho\to+\infty} \left\{ \int_{\mathbb{R}_+} \mathbb{1}\{\mathcal{E}_{1,\mathrm{ML}}\} f_{|h_s|^2}\left(\frac{a}{\rho}\right) \mathrm{d}a \cdot \right. \\ \left. \cdot \underline{\int_{\mathbb{R}_+} f_{|h_u|^2}(w)\,\mathrm{d}w} \int_{\mathbb{R}_+} f_{|h_d|^2}(y)\,\mathrm{d}y \right\} = 0.$$

The underlined term vanishes because of the definition of $\mathcal{E}_1$.

Event $\mathcal{E}_2$

With the change of variables $a = \rho x$ and $c = z\rho y$ one has, for $\rho \to +\infty$:

$$\lambda \to \frac{c}{\alpha(1+a)}, \qquad \phi_1 \to \bar{\phi}_1 < +\infty,$$
$$\phi_2 \to \frac{1}{z}, \qquad \lambda\bar{\phi}_1 = \frac{1 - z\bar{\phi}_1}{\alpha - 1 + z\bar{\phi}_1}.$$

Note that the last equation implies $z\bar{\phi}_1 > 1 - \alpha$. The limit of the spectral efficiency is hence

$$\lim_{\rho\to+\infty} (1+\alpha) I^{\mathrm{iid}}_{\mathrm{AF,ML}} = \alpha \ln(1+a) + \alpha \ln \alpha + \\ + z\bar{\phi}_1 - \ln(z\bar{\phi}_1) - 1 - \alpha \ln(z\bar{\phi}_1 + \alpha - 1).$$

From the definition of ϕ_1, one obtains

$$\frac{1}{z\bar{\phi}_1} = 1 + \frac{\alpha c}{\alpha z(1+a) + z\bar{\phi}_1 c}.$$

Solving the last equation with respect to c, the outage event $\mathcal{E}_{2,\mathrm{ML}}$ is equivalent to

$$\mathcal{E}'_{2,\mathrm{ML}} : c < \frac{\alpha z(1+a)}{\gamma\left(\frac{\alpha}{1-\gamma} - 1\right)}.$$

The quantity γ, with $\max\{0, 1-\alpha\} < \gamma < 1$, can be computed as

$$\gamma = f_0^{-1}((1+\alpha)R - \alpha \ln(1+a)),$$

where $f_0^{-1}(\cdot)$ is the inverse of

$$f_0(t) = t - \ln t - \alpha \ln(t + \alpha - 1) + \alpha \ln \alpha - 1,$$

monotonically decreasing in t for $\max\{0, 1-\alpha\} < t < 1$.

Again, $\mathbb{1}\{\mathcal{E}'_{2,\mathrm{ML}}\}$ represents a finite volume and the limit in (4.23) is finite and can be computed with the LDCT:

$$\lim_{\rho\to+\infty} \rho^2 \Pr[\mathcal{E}_{2,\mathrm{ML}}] = \frac{\zeta_s\zeta_d}{z} \int_{\mathbb{R}^2_+} \mathbb{1}\{\mathcal{E}'_{2,\mathrm{ML}}\}\,\mathrm{d}a\,\mathrm{d}c.$$

When considering the LMMSE receiver, the results depend on the coding rate as in the previous event:

- if $\alpha \le \alpha_{th}$, the outage event is equivalent to

$$\mathcal{E}'_{2,\mathrm{LMMSE}} : a < Q \text{ and } c < \frac{z(Q+1)(Q-a)}{1+Q-\alpha(Q-a)}.$$

  Thus, the contribution to the outage gain is given by

$$\lim_{\rho\to+\infty} \rho^2 \Pr[\mathcal{E}_{2,\mathrm{LMMSE}}] = \frac{\zeta_s\zeta_d}{z} \int_{\mathbb{R}^2_+} \mathbb{1}\{\mathcal{E}'_{2,\mathrm{LMMSE}}\}\,\mathrm{d}a\,\mathrm{d}c = \\ = \zeta_s\zeta_d(Q+1)\left[-Q - \frac{Q+1}{\alpha}\ln\left(1-\alpha\frac{Q}{Q+1}\right)\right];$$

- when $\alpha > \alpha_{th}$ the outage event depends on a only and, analogously to $\mathcal{E}_{1,\mathrm{LMMSE}}$, one has

$$\lim_{\rho\to+\infty} \rho \Pr[\mathcal{E}_{2,\mathrm{LMMSE}}] = 0$$

Event $\mathcal{E}_3$

This case does not bring any contribution when employing the ML receiver since, with both the relay channels "not small", the spectral efficiency grows without bound, i.e. $\mathbb{1}\{\mathcal{E}_{3,\mathrm{ML}}\} = 0$, for ρ large enough.

On the contrary, the LMMSE receiver must be analyzed with more attention:

- for $\alpha \le \alpha_{th}$, the SINR is always larger than Q for ρ large enough and $\mathbb{1}\{\mathcal{E}_{3,\mathrm{LMMSE}}\} = 0$, which implies no contribution to the outage gain;
- if $\alpha > \alpha_{th}$, then the outage event is equivalent to

$$\mathcal{E}'_{3,\mathrm{LMMSE}} : a < Q - \frac{Q+1}{\alpha},$$

  which implies

$$\lim_{\rho\to+\infty} \rho \Pr[\mathcal{E}_{3,\mathrm{LMMSE}}] = \\ = \lim_{\rho\to+\infty} \int_{\mathbb{R}^3_+} \mathbb{1}\{\mathcal{E}_{3,\mathrm{LMMSE}}\} f_{|h_s|^2}\left(\frac{a}{\rho}\right) f_{|h_u|^2}(w) f_{|h_d|^2}(y)\,\mathrm{d}a\,\mathrm{d}w\,\mathrm{d}y = \\ = \zeta_s\left(Q - \frac{1+Q}{\alpha}\right).$$

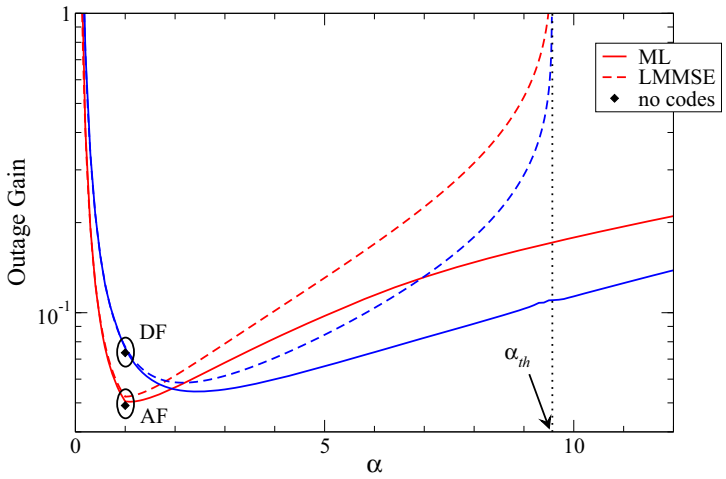


Figure 4.8: Single relay: outage gain as a function of the coding rate α for unitary channel variances, $\sigma_d^2/\sigma_u^2 = 1$ and target rate $R = 0.1$ nat/s/Hz.

Event $\mathcal{E}_4$

The contribution of this event is negligible. Indeed, with the change of variables $a = \rho x$, $b = z\rho w$ and $c = z\rho y$, the event $\mathcal{E}_4$ has a finite volume and

$$\lim_{\rho\to+\infty} \rho^2 \Pr[\mathcal{E}_4] = \frac{\zeta_s\zeta_u\zeta_d}{z^2} \lim_{\rho\to+\infty} \frac{1}{\rho} \int_{\mathbb{R}_+^3} \mathbb{1}\{\mathcal{E}_4\}\, \mathrm{d}a\, \mathrm{d}b\, \mathrm{d}c = 0$$

for both the receivers.

Figure 4.8 depicts the outage gain obtained by adding the contributions of the four cases (only the region with diversity order two is represented for the LMMSE receiver).

Comparison with the DF strategy

In Figure 4.8, the outage gain obtained by one DF relay according to the results of the previous section is reported as a term of comparison. We can notice that, for both receivers, the AF strategy works better than the DF one only for low values of the coding rate, namely for α lower than two, approximately. Intuition suggests the following explanation. When the signatures are long enough (i.e. α is small), the interference generated by the forwarded noise is not so heavy and the receiver may always extract information from the relay message. Conversely, when α is large, the interference is so strong that the message from an AF relay barely contributes at the receiver. In this case, it is hence better to use the DF strategy: relays transmit only when they could decode the source message but do not forward any noise.

Figure 4.8 also reports the outage gain given by repetition coding, i.e.

$\mathbf{C} = \mathbf{I}_K$ and $\alpha = 1$, namely

$$\kappa_{AF} = \frac{1}{2}\zeta_s\left(\zeta_d + \frac{\zeta_u}{z}\right)\left(e^{2R} - 1\right)^2, \qquad \kappa_{DF} = \zeta_s\left(\frac{\zeta_d}{2} + \frac{\zeta_u}{z}\right)\left(e^{2R} - 1\right)^2,$$

for AF and DF strategy, respectively. We can observe that the LD-STBC performs slightly worse at $\alpha = 1$. Indeed, when there is only one relay in the system, space-time coding is not necessary since there is no multiple-channel access. Furthermore, randomized LD-STBC introduces some inter-symbol interference due to its lack of orthogonality. However, for the DF strategy, the Figure 4.8 shows that the outage gain can be improved by adopting randomized LD-STBC at the relay. Indeed, since the relay contribution is complementary to the direct-link one, the relay may send a compressed version of the message (i.e. $\alpha > 1$) and still achieve diversity. In this way, we can better exploit the system degrees of freedom and minimize the outage gain.

4.3 Conclusions

In this chapter, we have followed the study of distributed randomized LD-STBC. More specifically, for the asymptotic (large linear-dispersion matrices) deterministic spectral efficiencies derived in the previous chapter, we have characterized the outage probability in the large-SNR regime by computing the diversity order and the outage gain.

According to these results, L DF relays always achieve full diversity order $L + 1$, except when the suboptimal LMMSE receiver is used together with a large coding rate. Furthermore, a comparison with previous works has shown interesting benefits. Indeed, even though the randomness of the codes introduces some interference with respect to orthogonal STC, the LD-STBC can achieve lower outage gains. This is due to its flexibility, that is the possibility of choosing coding rates higher than one, which better exploits the available degrees of freedom.

The AF relaying strategy, instead, cannot get all the benefits from randomized LD-STBC. Indeed, full diversity order is obtained only for very long signatures and decays as the coding rate increases (see Figure 4.7). This is mainly due to the colored noise forwarded by the relays.

Once again, we believe that these results are very representative of practical systems with not-so-large coding matrices. For instance, Figure 4.5 shows how the approximations derived in this chapter match the simulated outage probability of a system with signature length $N = 30$.

Appendix 4.A Proofs for the DF relaying strategy

4.A.1 Proof for the ML receiver

As explained in Section 4.1.1, we will resort to the total probability theorem to study the outage probability:

$$P_{out}(R) = \sum_{\mathcal{L}' \subseteq \mathcal{L}} \Pr[\mathcal{L}' \text{ is the decoding set}] \Pr\big[I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}} < R\big]. \tag{4.24}$$

Using the second definition in (4.6), it is straightforward to show that

$$\Pr[I_{ul} < R] = \Pr\left[|h_{ul}|^2 < \frac{Q}{z\rho}\right] = \zeta_u \frac{Q}{z\rho} + o\left(\frac{1}{\rho}\right)$$

and $\Pr[I_{ul} \geq R] = 1 + o(1)$, for $\rho \to +\infty$. Then, for very large ρ, the probability of the decoding subset $\mathcal{L}'$ can be approximated as

$$\Pr[\mathcal{L}' \text{ is the decoding subset}] = \left(\zeta_u \frac{Q}{z\rho}\right)^{L-|\mathcal{L}'|} + o\left(\frac{1}{\rho^{L-|\mathcal{L}'|}}\right), \tag{4.25}$$

where $|\cdot|$ is the set cardinality.

The second factor can be written as follows:

$$\begin{aligned}
&\Pr\big[I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}} < R\big] = \\
&\quad = \mathbb{E}\big[\mathbb{1}\big\{I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}} < R\big\}\big] \\
&\quad = \int_{\mathbb{R}_+^{|\mathcal{L}'|+1}} \mathbb{1}\big\{I'(x, \mathbf{y}_{\mathcal{L}'}; \rho) < (1+\alpha)R\big\} f_{|h_s|^2}(x) f_{|h_d|^2}(\mathbf{y}'_{\mathcal{L}}) \,\mathrm{d}x \,\mathrm{d}\mathbf{y}_{\mathcal{L}'} \\
&\quad = \rho^{-|\mathcal{L}'|-1} \int_{\mathbb{R}_+^{|\mathcal{L}'|+1}} \mathbb{1}\big\{I'(a, \mathbf{c}_{\mathcal{L}'}) < (1+\alpha)R\big\} f_{|h_s|^2}\left(\frac{a}{\rho}\right) f_{|h_d|^2}\left(\frac{\mathbf{c}'_{\mathcal{L}}}{\rho}\right) \mathrm{d}a \,\mathrm{d}\mathbf{c}_{\mathcal{L}'},
\end{aligned}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. We have also defined

$$I'(x, \mathbf{y}_{\mathcal{L}'}; \rho) = \alpha \ln(1+\rho x) + \alpha \ln\left(1 + \frac{\rho \sum_{l \in \mathcal{L}'} y_l}{\alpha \beta_{\mathcal{L}'}(1+\rho x)}\right) + \ln \beta_{\mathcal{L}'} + \frac{1}{\beta_{\mathcal{L}'}} - 1$$

and we made the change of variables $a = \rho x$ and $c_l = \rho y_l$, $\forall l \in \mathcal{L}'$.

Now, let us introduce the quantities

$$\begin{aligned}
I'_1 &= \alpha \ln(1+a) \geq 0; \\
I'_2 &= \alpha \ln\left(1 + \frac{\sum_{l \in \mathcal{L}'} c_l}{\alpha \beta_{\mathcal{L}'}(1+a)}\right) \geq 0; \\
I'_3 &= \ln \beta_{\mathcal{L}'} + \frac{1}{\beta_{\mathcal{L}'}} - 1 > 0,
\end{aligned}$$

such that $I'(a, \mathbf{c}_{\mathcal{L}'}) = I'_1 + I'_2 + I'_3$. The fact that $\mathbb{1}\{\sum_{i=1}^3 I'_i < (1+\alpha)R\} \leq \prod_{i=1}^3 \mathbb{1}\{I'_i < (1+\alpha)R\}$ will be used now to find an upper-bound on the outage probability.

The first term just needs to be rewritten as $\mathbb{1}\{\alpha \ln(1+a) < (1+\alpha)R\} = \mathbb{1}\{a < Q\}$, where Q is as defined in (4.8).

As for I'_2, one can note that

$$\begin{aligned}
\mathbb{1}\{I'_2 < (1+\alpha)R\} &= \mathbb{1}\left\{\alpha \ln\left(1 + \frac{\sum_{l\in\mathcal{L}'} c_l}{\alpha\beta_{\mathcal{L}'}(1+a)}\right) < (1+\alpha)R\right\} \\
&= \mathbb{1}\left\{\beta'_{\mathcal{L}} > \frac{\sum_{l\in\mathcal{L}'} c_l}{\alpha Q(1+a)}\right\} \\
&\overset{\text{(i)}}{=} \mathbb{1}\left\{\frac{\sum_{l\in\mathcal{L}'} c_l}{\alpha Q(1+a)}\left(1 - \alpha\frac{Q\sum_{l\in\mathcal{L}'} c_l}{\sum_{l\in\mathcal{L}'} c_l + Q\sum_{l\in\mathcal{L}'} c_l}\right) < 1\right\} \\
&= \mathbb{1}\left\{\frac{\sum_{l\in\mathcal{L}'} c_l(1+Q-\alpha Q)}{\alpha Q(1+a)(1+Q)} < 1\right\} \\
&= \begin{cases} \mathbb{1}\left\{\sum_{l\in\mathcal{L}'} c_l < \dfrac{\alpha Q(1+a)(1+Q)}{1+Q-\alpha Q}\right\} & \text{if } \alpha < 1 + \dfrac{1}{Q} \\ 1 & \text{if } \alpha \geq 1 + \dfrac{1}{Q}. \end{cases}
\end{aligned}$$

Equality (i) can be easily obtained by comparing the two sides of (4.4), see also Figure 4.9.

Similarly, the last term can be upper-bounded as follows:

$$\begin{aligned}
\mathbb{1}\{I'_3 < (1+\alpha)R\} &= \mathbb{1}\left\{\ln\beta_{\mathcal{L}'} + \frac{1}{\beta_{\mathcal{L}'}} - 1 < (1+\alpha)R\right\} \\
&\leq \mathbb{1}\{\ln\beta_{\mathcal{L}'} < (1+\alpha)R + 1\} = \mathbb{1}\{\beta_{\mathcal{L}'} < Q'\} \\
&\overset{(i)}{=} \mathbb{1}\left\{\frac{\sum_{l\in\mathcal{L}'} c_l}{\alpha Q'(1+a) + \sum_{l\in\mathcal{L}'} c_l} < \frac{Q'-1}{\alpha Q'}\right\} \\
&= \begin{cases} 1 & \text{if } \alpha \leq 1 - \dfrac{1}{Q'}, \\ \mathbb{1}\left\{\sum_{l\in\mathcal{L}'} c_l < \dfrac{\alpha Q'(Q'-1)(1+a)}{Q'(\alpha-1)+1}\right\} & \text{if } \alpha > 1 - \dfrac{1}{Q'}, \end{cases}
\end{aligned}$$

where we introduced $Q' = \exp[(1+\alpha)R + 1] > 1$. Once again, equality (i) comes directly from (4.4) and Figure 4.9.

Summarizing:

$$\mathbb{1}\{I'(a, \mathbf{c}_{\mathcal{L}'}) < (1+\alpha)R\} \leq \prod_{i=1}^3 \mathbb{1}\{I'_i < (1+\alpha)R\}$$

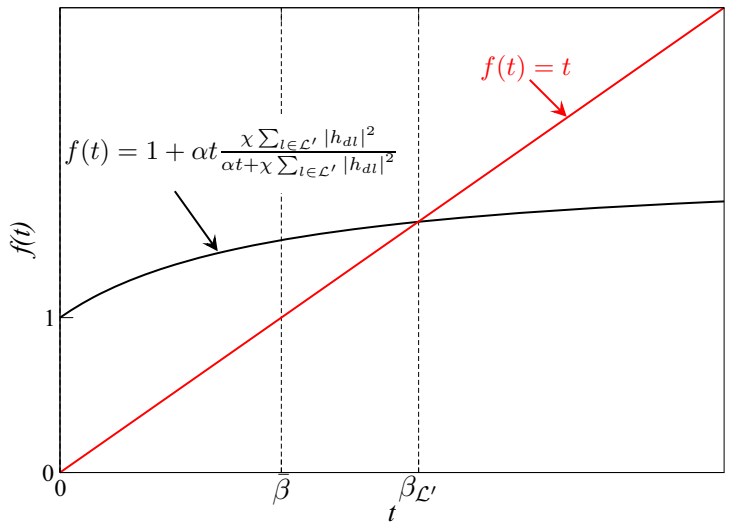


Figure 4.9: Graphical solution of equation (4.4). Note that for $\bar{\beta} < \beta_{\mathcal{L}'}$ it is $f(\bar{\beta}) = 1 + \alpha\bar{\beta}\frac{\chi \sum_{l\in\mathcal{L}'} |h_{dl}|^2}{\alpha\bar{\beta} + \chi \sum_{l\in\mathcal{L}'} |h_{dl}|^2} > \bar{\beta}$.

$$\leq \begin{cases} \mathbb{1}\{a < Q\}\mathbb{1}\left\{\sum_{l\in\mathcal{L}'} c_l < \dfrac{\alpha Q(1+a)(1+Q)}{1+Q-\alpha Q}\right\} \\ \qquad\qquad \text{if } \alpha < 1 + \dfrac{1}{Q}, \\ \mathbb{1}\{a < Q\}\mathbb{1}\left\{\sum_{l\in\mathcal{L}'} c_l < \dfrac{\alpha Q'(Q'-1)(1+a)}{Q'(\alpha-1)+1}\right\} \\ \qquad\qquad \text{if } \alpha \geq 1 + \dfrac{1}{Q}. \end{cases} \tag{4.26}$$

Now, it is straightforward to compute

$$\begin{aligned} &\lim_{\rho\to+\infty} \rho^{|\mathcal{L}'|+1} \Pr\left[I_{\mathrm{DF,ML}}^{(\mathcal{L}'),\mathrm{iid}} < R\right] = \\ &= \lim_{\rho\to+\infty} \int_{\mathbb{R}_+^{|\mathcal{L}'|+1}} \mathbb{1}\left\{I'(a, \mathbf{c}_{\mathcal{L}'}) < (1+\alpha)R\right\} f_{|h_s|^2}\left(\frac{a}{\rho}\right) f_{|h_d|^2}\left(\frac{\mathbf{c}_{\mathcal{L}'}}{\rho}\right) \mathrm{d}a\, \mathrm{d}\mathbf{c}_{\mathcal{L}'} \\ &\qquad = \zeta_s \zeta_d^{|\mathcal{L}'|} \int_{\mathbb{R}_+^{|\mathcal{L}'|+1}} \mathbb{1}\left\{I'(a, \mathbf{c}_{\mathcal{L}'}) < (1+\alpha)R\right\} \mathrm{d}a\, \mathrm{d}\mathbf{c}_{\mathcal{L}'}. \end{aligned} \tag{4.27}$$

Indeed, the limit and integral operations can be exchanged according to the Lebesgue's dominated convergence theorem (LDCT): from (4.26), the term $\mathbb{1}\left\{I'(a, \mathbf{c}_{\mathcal{L}'})\right\}$ represents a compact subset of $\mathbb{R}_+^{|\mathcal{L}'|+1}$ and, from (4.6), $f_{|h_s|^2}\left(\frac{a}{\rho}\right) f_{|h_d|^2}\left(\frac{\mathbf{c}_{\mathcal{L}'}}{\rho}\right)$ tends to the constant $\zeta_s \zeta_d^{|\mathcal{L}'|}$.

Finally, results (4.25) and (4.27) yield the following expression for the outage probability:

$$\lim_{\rho\to+\infty}\rho^{L+1}P_{out}(R)=\zeta_s\sum_{l=0}^{L}\binom{L}{l}\left(\frac{\zeta_u Q}{z}\right)^{L-l}\zeta_d^l\cdot \\ \cdot\int_{\mathbb{R}_+^{l+1}}\mathbb{1}\big\{I'(a,\mathbf{c}_1^l)<(1+\alpha)R\big\}\,\mathrm{d}a\,\mathrm{d}\mathbf{c}_1^l,$$

meaning that the diversity order is always maximum and equal to $L+1$. Generally, the last integral has to be computed numerically.

4.A.2 Proof for the LMMSE receiver

As before, the outage probability with the LMMSE receiver can be expressed as:

$$P_{out}(R)=\sum_{\mathcal{L}'\subseteq\mathcal{L}}\Pr[\mathcal{L}'\text{ is the decoding subset}]\Pr[SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}}<Q]. \quad (4.28)$$

The first term was already studied in Appendix 4.A.1. Here, we focus on the second one. First, note that

$$\begin{aligned}&\Pr[SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}}<Q]=\\ &\quad=\mathbb{E}[\mathbb{1}\{SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}}<Q\}]\\ &\quad=\int_{\mathbb{R}_+^{|\mathcal{L}'|+1}}\mathbb{1}\left\{\rho x+\frac{\rho}{\alpha\beta_{\mathcal{L}'}}\sum_{l\in\mathcal{L}'}y_l<Q\right\}f_{|h_s|^2}(x)f_{|h_d|^2}(\mathbf{y}'_{\mathcal{L}})\,\mathrm{d}x\,\mathrm{d}\mathbf{y}_{\mathcal{L}'}\\ &\quad=\rho^{-|\mathcal{L}'|-1}\int_{\mathbb{R}_+^{|\mathcal{L}'|+1}}\mathbb{1}\left\{a+\frac{1}{\alpha\beta_{\mathcal{L}'}}\sum_{l\in\mathcal{L}'}c_l<Q\right\}f_{|h_s|^2}\left(\frac{a}{\rho}\right)f_{|h_d|^2}\left(\frac{\mathbf{c}_{\mathcal{L}'}}{\rho}\right)\mathrm{d}a\,\mathrm{d}\mathbf{c}_{\mathcal{L}'}\end{aligned} \quad (4.29)$$

with the change of variables $a=\rho x$ and $c_l=\rho y_l$, $\forall l\in\mathcal{L}'$. The first term of the integrand can be rewritten as follows:

$$\begin{aligned}\mathbb{1}\left\{a+\frac{1}{\alpha\beta_{\mathcal{L}'}}\sum_{l\in\mathcal{L}'}c_l<Q\right\}&=\\ &=\mathbb{1}\{a<Q\}\mathbb{1}\left\{\frac{1}{\alpha\beta_{\mathcal{L}'}}\sum_{l\in\mathcal{L}'}c_l<Q-a\right\}\\ &=\mathbb{1}\{a<Q\}\mathbb{1}\left\{\beta_{\mathcal{L}'}>\frac{\sum_{l\in\mathcal{L}'}c_l}{\alpha(Q-a)}\right\}\\ &\overset{\text{(i)}}{=}\mathbb{1}\{a<Q\}\mathbb{1}\left\{\frac{\sum_{l\in\mathcal{L}'}c_l}{\alpha(Q-a)}\left(1-\alpha\frac{Q-a}{Q+1}\right)<1\right\}\\ &=\mathbb{1}\left\{a\le Q-\frac{Q+1}{\alpha}\right\}+\mathbb{1}\left\{Q-\frac{Q+1}{\alpha}<a<Q\right\}.\end{aligned}$$

$$\cdot \mathbb{1}\left\{\sum_{l\in\mathcal{L}'} c_l < \frac{\alpha(Q-a)(Q+1)}{Q+1-\alpha(Q-a)}\right\}.$$

The equality (i) can be easily obtained by comparing the two sides of (4.4), see also Figure 4.9. Substituting into (4.29), it gives:

$$\Pr[SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}} < Q] = \rho^{-1}\int_{\mathbb{R}_+} \mathbb{1}\left\{a \le Q - \frac{Q+1}{\alpha}\right\} f_{|h_s|^2}\left(\frac{a}{\rho}\right) \mathrm{d}a +$$
$$+ \rho^{-|\mathcal{L}'|-1}\int_{\mathbb{R}_+^{\mathcal{L}'+1}} \mathbb{1}\left\{Q - \frac{Q+1}{\alpha} < a < Q\right\}\mathbb{1}\left\{\sum_{l\in\mathcal{L}'} c_l < \frac{\alpha(Q-a)(Q+1)}{Q+1-\alpha(Q-a)}\right\}\cdot$$
$$\cdot f_{|h_s|^2}\left(\frac{a}{\rho}\right) f_{|h_d|^2}\left(\frac{\mathbf{c}_{\mathcal{L}'}}{\rho}\right)\mathrm{d}a\,\mathrm{d}\mathbf{c}_{\mathcal{L}'}. \quad (4.30)$$

The limit for $\rho \to +\infty$ of the first integral can be easily computed by means of the LDCT. Indeed, the indicator function represents a compact subset of $\mathbb{R}_+$ and $f_{|h_s|^2}\left(\frac{a}{\rho}\right)$ tends to the constant ζ_s. Similar reasoning holds for the second integral.

Recalling the definition of α_{th} in (4.9), we can distinguish two cases:

- if $\alpha \le \alpha_{th}$, the first term in (4.30) never exists and

$$\lim_{\rho\to+\infty} \rho^{|\mathcal{L}'|+1} \Pr[SINR_{\mathrm{DF}}^{(\mathcal{L}'),\mathrm{iid}} < Q] = \zeta_s \frac{(\alpha\zeta_d)^{|\mathcal{L}'|}}{|\mathcal{L}'|!}\mathcal{P}_{|\mathcal{L}'|}, \quad (4.31)$$

  where

$$\mathcal{P}_l = \int_0^Q \left[\frac{(Q+1)(Q-a)}{Q+1-\alpha(Q-a)}\right]^l \mathrm{d}a$$
$$= \left(-\frac{Q+1}{\alpha}\right)^l \left\{Q - l\cdot\frac{Q+1}{\alpha}\ln\left(\frac{Q+1}{Q+1-\alpha Q}\right) + \right.$$
$$\left. - \frac{Q+1}{\alpha}\sum_{j=2}^{l}\binom{l}{j}\frac{(-1)^{j-1}}{j-1}\left[\left(\frac{Q+1}{Q+1-\alpha Q}\right)^{j-1} - 1\right]\right\}. \quad (4.32)$$

  Inserting (4.25) and (4.31) into (4.28), one obtains

$$\lim_{\rho\to+\infty} \rho^{L+1} P_{out}(R) = \zeta_s \sum_{l=0}^{L}\binom{L}{l}\left(\frac{\zeta_u Q}{z}\right)^{L-l}\frac{(\alpha\zeta_d)^l}{l!}\mathcal{P}_l.$$

  By definition, this is the outage gain for a diversity order $d = L+1$;

- for $\alpha > \alpha_{th}$, the first term in (4.30) is well-defined and

$$\lim_{\rho\to+\infty} \rho P_{out}(R) = \zeta_s\left(Q - \frac{Q+1}{\alpha}\right),$$

meaning that the diversity order is $d = 1$. For $\alpha \to +\infty$, the right-hand side tends to the outage gain of the direct link, namely $\kappa_s = \zeta_s e^R$. Note that the only meaningful contribution comes from the case $\mathcal{L}' \equiv \mathcal{L}$, while the others tend to zero faster than $1/\rho$.

Appendix 4.B Proof for the AF relaying strategy with ML receiver

This appendix deals with the proof of Theorem 4.3. Briefly, we need to study the outage probability of the spectral efficiency

$$I(x, \mathbf{w}, \mathbf{y}; \rho) = \frac{1}{1+\alpha}\Bigg[\alpha \ln(1+\rho x) + (z\phi_1 - \ln \phi_1) - \\ - (z\phi_2 - \ln \phi_2) + \alpha \ln\Bigg(\prod_{l=1}^{L} \frac{1+\lambda_l \phi_1}{1+\gamma_l \phi_2}\Bigg)\Bigg], \quad (4.33)$$

where we have written $x = |h_s|^2, y_l = |h_{dl}|^2, w_l = |h_{ul}|^2$ and defined

$$\gamma_l = \frac{z\rho y_l}{\alpha(1+z\rho w_l)} = \frac{c_l}{\alpha(1+b_l)}.$$

In the last expression, the notation $a = \rho x, b_l = z\rho w_l$ and $c_l = z\rho y_l$ has been introduced. Recall from (4.16) and (4.17) that

$$\frac{1}{\phi_1} = z + \alpha \sum_{l=1}^{L} \frac{\lambda_l}{1+\lambda_l \phi_1} \qquad \frac{1}{\phi_2} = z + \alpha \sum_{l=1}^{L} \frac{\gamma_l}{1+\gamma_l \phi_2}.$$

The coefficients $\{\lambda_l\}_{l=1}^L$ are the solutions to the equation

$$\sum_{l=1}^{L} b_l \frac{\gamma_l}{\lambda - \gamma_l} = 1 + a,$$

c.f. Appendix 3.A.3 (for clarity's sake, we neglect the unlikely case where two or more relays have both channel coefficients equal to one another).

It is now possible to eliminate the direct dependence on the values $\{\lambda_l\}_{l=1}^L$ from (4.33). Indeed, it is enough to notice that

$$\prod_{l=1}^{L}(1+\lambda_l \phi_1) = (-\phi_1)^L P\Big(-\frac{1}{\phi_1}\Big),$$

where $P(\lambda)$ is the monic polynomial with roots $\{\lambda_l\}$, namely

$$P(\lambda) = \prod_{l=1}^{L}(\lambda - \lambda_l) = \prod_{l=1}^{L}(\lambda - \gamma_l) - \sum_{l=1}^{L} \frac{b_l}{1+a}\gamma_l \prod_{k \neq l}(\lambda - \gamma_k). \quad (4.34)$$

The last result implies

$$\prod_{l=1}^{L}(1+\lambda_l\phi_1) = \prod_{l=1}^{L}(1+\gamma_l\phi_1) + \sum_{l=1}^{L}\frac{b_l}{1+a}\phi_1\gamma_l\prod_{k\neq l}(1+\phi_1\gamma_k)$$

and, consequently,

$$\begin{aligned}(1+\alpha)I(x,\mathbf{w},\mathbf{y};\rho) &= \alpha\ln(1+\rho x) + f(\phi_1) - f(\phi_2) + \\ &\quad + \alpha\ln\left(1+\sum_{l=1}^{L}\frac{b_l}{1+a}\frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}\right) \\ &= f(\phi_1) - f(\phi_2) + \alpha\ln\left(1+a+\sum_{l=1}^{L}b_l\frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}\right),\end{aligned} \tag{4.35}$$

with

$$f(\phi) = z\phi - \ln\phi + \alpha\sum_{l=1}^{L}\ln(1+\gamma_l\phi), \quad \phi\in\left(0, z^{-1}\right).$$

Since $f(\phi)$ takes on its minimum value at $\phi=\phi_2$ (it is enough to compute the derivative and set it equal to zero), then

$$f(\phi_1) - f(\phi_2) > 0.$$

Hence, we can ensure

$$(1+\alpha)I(x,\mathbf{w},\mathbf{y};\rho) > \alpha\ln\left(1+a+\sum_{l=1}^{L}b_l\frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}\right)$$

and also

$$(1+\alpha)I(x,\mathbf{w},\mathbf{y};\rho) > f(\phi_1) - f(\phi_2),$$

two results that will be extensively used in what follows.

4.B.1 Preliminary considerations

We are now ready to investigate the large-SNR outage probability of the spectral efficiency (4.35), namely the asymptotic behavior of

$$\Pr[I(x,\mathbf{w},\mathbf{y};\rho) < R] = \mathbb{E}[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}]$$

as $\rho\to+\infty$. Recall that $\mathbb{1}\{\cdot\}$ is the indicator function and $\mathbb{E}[\cdot]$ the expected value with respect to the channel distributions. As mentioned in Section 4.2.1 and following the approach described in [33], one of four possible states is assigned to each relay according to the quality of its channels.

Consider the two quantities $\delta, \epsilon \in (0, 1/(L+1))$, such that $\delta > \epsilon$. We define the following functions for $l = 1, \ldots, L$:

$$\begin{aligned}
\mathcal{I}_{l,1} &= \mathbb{1}\{w_l \leq \rho^{\epsilon-1}\}\mathbb{1}\{y_l > \rho^{\delta-1}\} \\
\mathcal{I}_{l,2} &= \mathbb{1}\{w_l > \rho^{\epsilon-1}\}\mathbb{1}\{y_l \leq \rho^{\delta-1}\} \\
\mathcal{I}_{l,3} &= \mathbb{1}\{w_l > \rho^{\epsilon-1}\}\mathbb{1}\{y_l > \rho^{\delta-1}\} \\
\mathcal{I}_{l,4} &= \mathbb{1}\{w_l \leq \rho^{\epsilon-1}\}\mathbb{1}\{y_l \leq \rho^{\delta-1}\}
\end{aligned}$$

which in turn define the sets

$$\begin{aligned}
l \in \mathcal{I}_1 &\Leftrightarrow \mathcal{I}_{l,1} = 1 \\
l \in \mathcal{I}_2 &\Leftrightarrow \mathcal{I}_{l,2} = 1 \\
l \in \mathcal{I}_3 &\Leftrightarrow \mathcal{I}_{l,3} = 1 \\
l \in \mathcal{I}_4 &\Leftrightarrow \mathcal{I}_{l,4} = 1.
\end{aligned}$$

Observing that

$$\mathcal{I}_{l,1} + \mathcal{I}_{l,2} + \mathcal{I}_{l,3} + \mathcal{I}_{l,4} = 1,$$

we can write

$$\begin{aligned}
\mathbb{E}[\mathbb{1}\{I(x, \mathbf{w}, \mathbf{y}; \rho) < R\}] &= \mathbb{E}\Bigg[\mathbb{1}\{I(x, \mathbf{w}, \mathbf{y}; \rho) < R\} \cdot \\
&\qquad \cdot \prod_{l=1}^{L} (\mathcal{I}_{l,1} + \mathcal{I}_{l,2} + \mathcal{I}_{l,3} + \mathcal{I}_{l,4})\Bigg] \\
&= \sum_{\boldsymbol{\nu} \in \{1,2,3,4\}^L} \mathbb{E}\Bigg[\mathbb{1}\{I(x, \mathbf{w}, \mathbf{y}; \rho) < R\} \prod_{l=1}^{L} \mathcal{I}_{l,\nu(l)}\Bigg],
\end{aligned}$$

where the last sum is over all the possible L-tuples $\boldsymbol{\nu}$ with values in $\{1, 2, 3, 4\}$.

The following property will also be very useful in the rest of the proof:

Lemma 4.1. *The equality*

$$\mathbb{E}[\mathbb{1}\{I(x, \mathbf{w}, \mathbf{y}; \rho) < R\}] = \mathbb{E}\big[\mathbb{1}\{I(x, \mathbf{w}, \mathbf{y}; \rho) < R\}\mathbb{1}\{x \leq \rho^{\epsilon-1}\}\big],$$

holds true for sufficiently high ρ.

Proof. Since

$$I(x, \mathbf{w}, \mathbf{y}; \rho) > \frac{\alpha}{1+\alpha} \ln(1 + \rho x),$$

we see that

$$\mathbb{E}\big[\mathbb{1}\{I(x, \mathbf{w}, \mathbf{y}; \rho) < R\}\mathbb{1}\{x > \rho^{\epsilon-1}\}\big] \leq \mathbb{E}\big[\mathbb{1}\{x < Q/\rho\}\mathbb{1}\{x > \rho^{\epsilon-1}\}\big]$$

($Q = \exp\left(\frac{1+\alpha}{\alpha}R\right) - 1$ was defined in (4.8)). Denote by Z the supremum of the channel pdf, which is right-continuous at zero, in a region of zero. Then, we have

$$\mathbb{E}\big[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}\mathbb{1}\big\{x > \rho^{\epsilon-1}\big\}\big] \leq \frac{Z}{\rho}\big[Q - \rho^{\epsilon}\big]^{+},$$

where $[\cdot]^{+} = \max\{0, \cdot\}$. The right-hand side is identically zero for sufficiently high ρ (the argument of $[\cdot]^{+}$ becomes negative). Since

$$\begin{aligned}\mathbb{E}[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}] = \mathbb{E}\big[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}\mathbb{1}\big\{x \leq \rho^{\epsilon-1}\big\}\big] + \\ + \mathbb{E}\big[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}\mathbb{1}\big\{x > \rho^{\epsilon-1}\big\}\big],\end{aligned}$$

the lemma is proven. □

According to this result, we can always assume $x \leq \rho^{\epsilon-1}$ (or, equivalently, $a = \rho x \leq \rho^{\epsilon}$) when investigating the large-SNR outage probability.

4.B.2 Main results

As mentioned before, we consider now each term in

$$\Pr[I(x,\mathbf{w},\mathbf{y};\rho) < R] = \mathbb{E}[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}] = \sum_{\boldsymbol{\nu}\in\{1,2,3,4\}^{L}} p_{\boldsymbol{\nu}}, \quad (4.36)$$

where we introduced

$$p_{\boldsymbol{\nu}} = \mathbb{E}\left[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}\prod_{l=1}^{L}\mathcal{I}_{l,\nu(l)}\right],$$

the contribution of the L-tuple $\boldsymbol{\nu}$. Besides, we denote by $\Gamma(\boldsymbol{\nu}) = |\mathcal{I}_1| + |\mathcal{I}_2| + 2|\mathcal{I}_4|$ the total number of channels in a deep fading situation.

Observe that each term $p_{\boldsymbol{\nu}}$ converges to zero as $\rho \to +\infty$, although the convergence speed depends on the L-tuple $\boldsymbol{\nu}$. Fortunately, the L^4 different L-tuples can be grouped according to their characteristics, as summarized by Table 4.1. The situations marked with "—" are not possible: in particular, the first row does not make sense since it must be $|\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_3| + |\mathcal{I}_4| = L$.

The four lemmas describing the asymptotic behavior of the different L-tuples can be stated as follows:

Lemma 4.2. *Assume that $\boldsymbol{\nu} \in \{1,2\}^{L}$ and $1 - \alpha|\mathcal{I}_1| > 0$. In this situation, the limit*

$$\lim_{\rho\to+\infty} \rho^{1+\Gamma(\boldsymbol{\nu})} p_{\boldsymbol{\nu}}$$

exists and is positive. Observe that $\Gamma(\boldsymbol{\nu}) = L$ and we achieve full diversity order $L + 1$.

Table 4.1: Association between relay partitions and relative lemma (ML receiver).

#	$\lvert\mathcal{I}_3\rvert$	$\lvert\mathcal{I}_4\rvert$	$\lvert\mathcal{I}_2\rvert$	$\lvert\mathcal{I}_1\rvert$	*Case* $\alpha < \frac{1}{\lvert\mathcal{I}_1\rvert}$	*Case* $\alpha > \frac{1}{\lvert\mathcal{I}_1\rvert}$
0	0	0	0	0	—	—
1	0	0	0	>0	Lemma 4.2	Lemma 4.3
2	0	0	>0	0		—
3	0	0	>0	>0		Lemma 4.3
4	0	>0	0	0	Lemma 4.4	—
5	0	>0	0	>0		
6	0	>0	>0	0		—
7	0	>0	>0	>0		
8	>0	0	0	>0	Lemma 4.5	Lemma 4.3
9	>0	0	>0	>0		
10	>0	>0	0	>0		
11	>0	>0	>0	>0		
12	>0	0	0	0		—
13	>0	0	>0	0		—
14	>0	>0	0	0		—
15	>0	>0	>0	0		—

Lemma 4.3. *If* $\alpha|\mathcal{I}_1| - 1 > 0$ *(which implies* $|\mathcal{I}_1| > 0$*), then*

$$\rho^{|\mathcal{I}_1|+1} p_{\boldsymbol{\nu}} = \begin{cases} O(1) & \text{if } |\mathcal{I}_2| + |\mathcal{I}_4| = 0, \\ o(1) & \text{if } |\mathcal{I}_2| + |\mathcal{I}_4| > 0. \end{cases}$$

Lemma 4.4. *When* $|\mathcal{I}_4| > 0$ *and* $|\mathcal{I}_3| = 0$*, then*

$$\rho^{L+1} p_{\boldsymbol{\nu}} \to 0,$$

i.e. the corresponding contribution in (4.36) decays to zero faster than $\rho^{-(L+1)}$*.*

Note that $L+1$ is the maximum diversity order that can be provided by the system.

Lemma 4.5. *If* $|\mathcal{I}_3| > 0$ *and* $\alpha|\mathcal{I}_1| - 1 < 0$*, then*

$$p_{\boldsymbol{\nu}} = 0$$

for a sufficiently high ρ*. These terms will never contribute to the large-SNR outage probability.*

Table 4.2 summarizes the results above and shows the convergence speed of all the situations identified in Table 4.1.

Table 4.2: Convergence speed associated to each relay partition (ML receiver).

#	$\Gamma(p_{\boldsymbol{\nu}})$	*Case* $\alpha < \frac{1}{\lvert\mathcal{I}_1\rvert}$	*Case* $\alpha > \frac{1}{\lvert\mathcal{I}_1\rvert}$
0	—	—	—
1	$\lvert\mathcal{I}_1\rvert = L$	$p_{\boldsymbol{\nu}} = O\big(\rho^{-(L+1)}\big)$	$p_{\boldsymbol{\nu}} = O\big(\rho^{-(L+1)}\big)$
2	$\lvert\mathcal{I}_2\rvert = L$		—
3	$\lvert\mathcal{I}_1\rvert + \lvert\mathcal{I}_2\rvert = L$		$p_{\boldsymbol{\nu}} = o\big(\rho^{-(\lvert\mathcal{I}_1\rvert+1)}\big)$
4	$2\lvert\mathcal{I}_4\rvert > L$	$p_{\boldsymbol{\nu}} = o\big(\rho^{-(L+1)}\big)$	—
5	$\lvert\mathcal{I}_1\rvert + 2\lvert\mathcal{I}_4\rvert > L$		
6	$\lvert\mathcal{I}_2\rvert + 2\lvert\mathcal{I}_4\rvert > L$		—
7	$\lvert\mathcal{I}_1\rvert + \lvert\mathcal{I}_2\rvert + 2\lvert\mathcal{I}_4\rvert > L$		
8	$\lvert\mathcal{I}_1\rvert < L$	$p_{\boldsymbol{\nu}} = 0$ for $\rho > \rho_0$	$p_{\boldsymbol{\nu}} = O\big(\rho^{-(\lvert\mathcal{I}_1\rvert+1)}\big)$
9	$\lvert\mathcal{I}_1\rvert + \lvert\mathcal{I}_2\rvert < L$		$p_{\boldsymbol{\nu}} = o\big(\rho^{-(\lvert\mathcal{I}_1\rvert+1)}\big)$
10	$\lvert\mathcal{I}_1\rvert + 2\lvert\mathcal{I}_4\rvert$		
11	$\lvert\mathcal{I}_1\rvert + \lvert\mathcal{I}_2\rvert + 2\lvert\mathcal{I}_4\rvert$		
12	0		—
13	$\lvert\mathcal{I}_2\rvert < L$		—
14	$2\lvert\mathcal{I}_4\rvert$		—
15	$\lvert\mathcal{I}_2\rvert + 2\lvert\mathcal{I}_4\rvert$		—

Proof of Theorem 4.3

In this section, we split the results of Theorem 4.3 according to the value of the coding rate α.

Proposition 4.2. *If $\alpha < L^{-1}$, then*

$$\lim_{\rho\to+\infty} \rho^{L+1}\mathbb{E}[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}]$$

is finite and non-zero.

Proof. Consider the contribution of a generic L-tuple $\boldsymbol{\nu}$ in (4.36), that is $p_{\boldsymbol{\nu}}$. If $\alpha < L^{-1}$ we will always have $\alpha|\mathcal{I}_1| < 1$ (the first column in Table 4.2). Now, observe that $p_{\boldsymbol{\nu}} = O\big(\rho^{-(L+1)}\big)$ if $\boldsymbol{\nu}$ belongs to Situations 1–3 and $p_{\boldsymbol{\nu}} = o\big(\rho^{-(L+1)}\big)$ otherwise. Hence, when computing

$$\lim_{\rho\to+\infty} \rho^{L+1} \sum_{p_{\boldsymbol{\nu}}\in\{1,2,3,4\}^L} p_{\boldsymbol{\nu}},$$

the contributions from the L-tuples belonging to the Situations 4–15 vanish in the limit, while those related to the Situations 1–3 provide a finite non-zero limit, due to Lemma 4.2. □

Proposition 4.3. *If $M^{-1} \leq \alpha < (M-1)^{-1}$, with $M = 2,\dots,L$, then*

$$\lim_{\rho\to+\infty} \rho^{M+1}\mathbb{E}[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\}]$$

is finite and non-zero.

Proof. Once again, we consider each contribution $p_{\boldsymbol{\nu}}$.

Let $\boldsymbol{\nu}$ be a L-tuple such that $|\mathcal{I}_1| = 0$. In this case, it is always $\alpha|\mathcal{I}_1| < 1$ (the first column of Table 4.2). Then, the contribution always decays faster than $L+1 \geq M+1$ except when $|\mathcal{I}_2| = L$, i.e. when $\boldsymbol{\nu}$ belongs to Situation 2. In this last case $p_{\boldsymbol{\nu}}$ decays exactly as $\rho^{-(L+1)}$. In other words, we have

$$|\mathcal{I}_1| = 0 \Rightarrow \rho^{M+1}p_{\boldsymbol{\nu}} = \begin{cases} O\big(\rho^{-(L-M)}\big) & \boldsymbol{\nu} \text{ in Situation 2,} \\ o\big(\rho^{-(L-M)}\big) & \text{otherwise.} \end{cases}$$

This means that, when considering

$$\lim_{\rho\to+\infty} \rho^{M+1} \sum_{p_{\boldsymbol{\nu}}\in\{2,3,4\}^L} p_{\boldsymbol{\nu}},$$

all contributions will vanish if $M < L$. On the contrary, if $M = L$, the limit is finite and non-zero due to Situation 2 and Lemma 4.2.

Consider now the case where the L-tuple contains at least one element in $\mathcal{I}_1$, i.e. $|\mathcal{I}_1| > 0$. In this case, we may have either $|\mathcal{I}_1| \leq M-1$ or $|\mathcal{I}_1| \geq M$.

- $|\mathcal{I}_1| \leq M-1$. In this situation, we can write $\alpha|\mathcal{I}_1| \leq (M-1)\alpha < 1$, which implies that $1-\alpha|\mathcal{I}_1| > 0$ (the first column of Table 4.2). As before, the corresponding contribution decays to zero faster than (or as) $\rho^{-(L+1)}$. Since $L+1 \geq M+1$, the term $p_{\boldsymbol{\nu}}$ vanishes as $\rho \to +\infty$ in all cases except when $M=L$ and the L-tuple is in Situation 1 or 3. In other words, we have

$$0 < |\mathcal{I}_1| \leq M-1 \Rightarrow \rho^{M+1} p_{\boldsymbol{\nu}} = \begin{cases} O\big(\rho^{-(L-M)}\big) & \boldsymbol{\nu} \text{ in Sit. 1 and 3,} \\ o\big(\rho^{-(L-M)}\big) & \text{otherwise.} \end{cases}$$

- $|\mathcal{I}_1| \geq M$. In this situation, we can write $\alpha|\mathcal{I}_1| \geq M\alpha > 1$, which implies that $\alpha|\mathcal{I}_1| - 1 > 0$ (the second column of Table 4.2). Hence, any term $p_{\boldsymbol{\nu}}$ arising from these situations decays to zero as (or faster than) $\rho^{-(|\mathcal{I}_1|+1)}$, which means

$$M \leq |\mathcal{I}_1| \leq L \Rightarrow \rho^{M+1} p_{\boldsymbol{\nu}} = \begin{cases} O\big(\rho^{-(|\mathcal{I}_1|-M)}\big) & \boldsymbol{\nu} \text{ in Situations 1 and 8,} \\ o\big(\rho^{-(|\mathcal{I}_1|-M)}\big) & \text{otherwise.} \end{cases}$$

  In other words, the only finite contributions are those obtained from the L-tuples with $|\mathcal{I}_1| = M$ (due to Lemma 4.3), while the contributions vanish as $\rho \to +\infty$ if $|\mathcal{I}_1| > M$.

This concludes the proof of the proposition.

□

Proposition 4.4. *If $\alpha > 1$, then the following limit*

$$\lim_{\rho \to +\infty} \rho^2 \mathbb{E}[\mathbb{1}\{I(x, \mathbf{w}, \mathbf{y}; \rho) < R\}]$$

is finite and non-zero.

Proof. Observe that $\alpha > 1$ implies that $\alpha|\mathcal{I}_1| > 1$ if $|\mathcal{I}_1| > 0$ and $\alpha|\mathcal{I}_1| < 1$ if $|\mathcal{I}_1| = 0$. Hence, we have contributions from both columns of Table 4.2. Consider first the L-tuples such that $|\mathcal{I}_1| = 0$ (which correspond to situations with even number in the first column of the table). When $\rho \to +\infty$, their contribution can be summarized as follows:

$$|\mathcal{I}_1| = 0 \Rightarrow \rho^2 p_{\boldsymbol{\nu}} = \begin{cases} O\big(\rho^{-(L-1)}\big) & \boldsymbol{\nu} \text{ in Situation 2,} \\ o\big(\rho^{-(L-1)}\big) & \text{otherwise,} \end{cases}$$

which implies that all the L-tuples give rise to terms that converge to zero (except, perhaps, the case $L=1$).

Next, consider the contributions from L-tuples such that $|\mathcal{I}_1| > 0$ (in Table 4.2, the situations with odd number in the second column). In these cases, all the terms converge to zero faster than or as ρ^2, namely

$$|\mathcal{I}_1| > 0 \Rightarrow \rho^2 p_{\boldsymbol{\nu}} = \begin{cases} O\big(\rho^{-(|\mathcal{I}_1|-1)}\big) & \boldsymbol{\nu} \text{ in Situations 1 and 9,} \\ o\big(\rho^{-(|\mathcal{I}_1|-1)}\big) & \text{otherwise.} \end{cases}$$

Hence, the L-tuples $\boldsymbol{\nu}$ that have $|\mathcal{I}_1| = 1$ provide a non-zero contribution, whereas the rest of the L-tuples have a negligible contribution in asymptotic terms. □

4.B.3 Proofs of Lemmas 4.2 to 4.5

All the results presented so far are based on Lemmas 4.2 to 4.5. To provide a valid proof for each one of them, it is of utmost importance to establish the asymptotic behavior of ϕ_1 in the large-SNR domain.

Behavior of ϕ_1

First, we derive an expression for ϕ_1 which does not depend on the variables $\{\lambda_l\}_{l=1}^{L}$. For this purpose, rewrite (4.16) as

$$1 - \alpha L = z\phi_1 - \alpha \sum_{l=1}^{L} \frac{1}{1+\lambda_l \phi_1}$$

and observe that

$$\sum_{l=1}^{L} \frac{1}{1+\lambda_l \phi_1} = -\frac{1}{\phi_1} \frac{P'\left(-\frac{1}{\phi_1}\right)}{P\left(-\frac{1}{\phi_1}\right)},$$

where $P(\lambda)$ is the monic polynomial defined in (4.34) and $P'(\lambda)$ its first derivative, namely

$$P'(\lambda) = \sum_{l=1}^{L} \prod_{\substack{k=1 \\ k\neq l}}^{L} (\lambda - \gamma_k) - \sum_{l=1}^{L} \frac{b_l}{1+a} \gamma_l \sum_{\substack{k=1 \\ k\neq l}}^{L} \prod_{\substack{r\neq l \\ r\neq k}} (\lambda - \gamma_r).$$

Therefore, we can express

$$\begin{aligned}\sum_{l=1}^{L} \frac{1}{1+\lambda_l \phi_1} &= \frac{1}{1+\sum_{l=1}^{L} \frac{b_l}{1+a} \frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}} \left[\sum_{l=1}^{L} \frac{1}{1+\phi_1\gamma_l} + \right. \\ &\qquad \left. + \sum_{l=1}^{L} \frac{b_l}{1+a} \frac{\phi_1\gamma_l}{1+\phi_1\gamma_l} \sum_{\substack{k=1 \\ k\neq l}}^{L} \frac{1}{1+\phi_1\gamma_k} \right] \\ &= \sum_{l=1}^{L} \frac{1}{1+\phi_1\gamma_l} - \frac{1}{1+\sum_{l=1}^{L} \frac{b_l}{1+a} \frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}} \sum_{l=1}^{L} \frac{b_l}{1+a} \frac{\phi_1\gamma_l}{(1+\phi_1\gamma_l)^2}.\end{aligned}$$

The last result implies that ϕ_1 is the positive solution to the equation

$$1 = z\phi_1 + \alpha \sum_{l=1}^{L} \frac{\phi_1\gamma_l}{1+\phi_1\gamma_l} + \alpha \frac{\sum_{l=1}^{L} b_l \frac{\phi_1\gamma_l}{(1+\phi_1\gamma_l)^2}}{1+a+\sum_{l=1}^{L} b_l \frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}}$$

Remark 4.1. *The quantity ϕ_1 is always decreasing in the variables $\{c_l\}_{l=1}^L$.*

Proof. Consider λ_k as a function of $a = \rho x$ and $\{(b_l = z\rho w_l,\ c_l = z\rho y_l);\ l = 1, \ldots, L\}$, defined as the solution to

$$\sum_{l=1}^{L} b_l \frac{\gamma_l}{\lambda - \gamma_l} = 1 + a,$$

where, according to the previous definition, $\gamma_l = \frac{c_l}{\alpha(1+b_l)}$. Taking derivatives of the above expression, one can readily see that (now, we denote by $\lambda = \lambda(a, \mathbf{b}, \mathbf{c})$ a particular solution)

$$\frac{\partial \lambda}{\partial \gamma_k} \sum_{l=1}^{L} b_l \frac{\gamma_l}{(\lambda - \gamma_l)^2} = b_k \frac{\lambda}{(\lambda - \gamma_k)^2} \Rightarrow \frac{\partial \lambda}{\partial c_k} > 0$$

$$\frac{\partial \lambda}{\partial a} = -\left(\sum_{l=1}^{L} b_l \frac{\gamma_l}{(\lambda - \gamma_l)^2} \right)^{-1} < 0 \Rightarrow \frac{\partial \lambda}{\partial a} < 0$$

$$\frac{\partial \lambda}{\partial b_k} \alpha \sum_{l=1}^{L} \frac{c_l b_l (1 + b_l)}{(\alpha(1 + b_l)\lambda - c_l)^2} = \frac{(\alpha\lambda - c_k)c_k}{(\alpha(1 + b_l)\lambda - c_l)^2} \Rightarrow \frac{\partial \lambda}{\partial b_k} < 0 \text{ if } \lambda < \frac{c_k}{\alpha}.$$

This means that $\lambda = \lambda(a, \mathbf{b}, \mathbf{c})$ is monotonically decreasing with a and monotonically increasing with each c_k. Since ϕ_1 is a decreasing function of each λ_k, then ϕ_1 is a decreasing function of each c_k. □

Now, assuming that $1 - \alpha|\mathcal{I}_1| > 0$, the last remark implies that ϕ_1 is lower-bounded by the positive solution to the equation

$$1 - \alpha|\mathcal{I}_1| = z\phi_1 + \alpha \sum_{l \in \mathcal{I}_2 \cup \mathcal{I}_3\mathcal{I}_4} \frac{\phi_1 \gamma_l}{1 + \phi_1 \gamma_l} + \\ + \alpha \frac{\sum_{l \in \mathcal{I}_2 \cup \mathcal{I}_3\mathcal{I}_4} b_l \frac{\phi_1 \gamma_l}{(1+\phi_1\gamma_l)^2}}{1 + a + \sum_{l \in \mathcal{I}_1} b_l + \sum_{l \in \mathcal{I}_2 \cup \mathcal{I}_3\mathcal{I}_4} b_l \frac{\phi_1 \gamma_l}{1+\phi_1\gamma_l}},$$

obtained by letting $c_k \to +\infty$, for $k \in \mathcal{I}_1$.

Another lower-bound may be obtained by noticing that, for any non-empty set $\mathcal{I}$ of relays,

$$\frac{\sum_{l \in \mathcal{I}} b_l \frac{\phi_1 \gamma_l}{(1+\phi_1\gamma_l)^2}}{1 + a + \sum_{l \in \mathcal{I}} b_l \frac{\phi_1 \gamma_l}{1+\phi_1\gamma_l}} < \frac{\sum_{l \in \mathcal{I}} b_l \frac{\phi_1 \gamma_l}{(1+\phi_1\gamma_l)^2}}{\sum_{l \in \mathcal{I}} b_l \frac{\phi_1 \gamma_l}{1+\phi_1\gamma_l}} \leq 1.$$

Hence, when $1 - \alpha(|\mathcal{I}_1| + 1) > 0$ (which implies $\alpha < 1$), ϕ_1 is lower-bounded by the solution to the following equation

$$1 - \alpha(|\mathcal{I}_1| + 1) = z\phi_1 + \alpha \sum_{l \in \mathcal{I}_2 \cup \mathcal{I}_3\mathcal{I}_4} \frac{\phi_1 \gamma_l}{1 + \phi_1 \gamma_l}.$$

Similarly, when $\alpha|\mathcal{I}_1|-1>0$, ϕ_1 can be upper-bounded by the solution to the equation

$$1=z\phi_1+\alpha\sum_{l\in\mathcal{I}_1\cup\mathcal{I}_3}\frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}+\alpha\frac{\sum_{l\in\mathcal{I}_1\cup\mathcal{I}_3}b_l\frac{\phi_1\gamma_l}{(1+\phi_1\gamma_l)^2}}{1+a+\sum_{l\in\mathcal{I}_1\cup\mathcal{I}_3}b_l\frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}},$$

obtained by letting $c_k\to 0$, for $k\in\mathcal{I}_2\cup\mathcal{I}_4$. The last equation can also be rewritten as

$$\alpha|\mathcal{I}_1|-1=\alpha\sum_{l\in\mathcal{I}_1}\frac{1}{1+\phi_1\gamma_l}-\Bigg(z\phi_1+\alpha\sum_{l\in\mathcal{I}_3}\frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}+ \\ +\alpha\frac{\sum_{l\in\mathcal{I}_1\cup\mathcal{I}_3}b_l\frac{\phi_1\gamma_l}{(1+\phi_1\gamma_l)^2}}{1+a+\sum_{l\in\mathcal{I}_1\cup\mathcal{I}_3}b_l\frac{\phi_1\gamma_l}{1+\phi_1\gamma_l}}\Bigg).$$

The term between parenthesis is positive and, hence, we can write

$$\alpha|\mathcal{I}_1|-1<\alpha\sum_{l\in\mathcal{I}_1}\frac{1}{1+\phi_1\gamma_l}<\alpha|\mathcal{I}_1|\frac{\alpha(1+z\rho^{\epsilon})}{\alpha(1+z\rho^{\epsilon})+\phi_1 z\rho^{\delta}},$$

where we used the fact that

$$\gamma_l=\frac{c_l}{\alpha(1+b_l)}>\frac{z\rho^{\delta}}{\alpha(1+z\rho^{\epsilon})}\text{ for }l\in\mathcal{I}_1.$$

After some algebra, we get the upper-bound

$$\phi_1<\frac{1}{\alpha|\mathcal{I}_1|-1}\frac{\alpha(1+z\rho^{\epsilon})}{z\rho^{\delta}}.$$

Proof of Lemma 4.4

Let us consider an L-tuple $\boldsymbol{\nu}$ such that $|\mathcal{I}_4|>0$ and $|\mathcal{I}_3|=0$. We want to show that

$$\rho^{L+1}p_{\boldsymbol{\nu}}=\rho^{L+1}\mathbb{E}\Bigg[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho)<R\}\prod_{l=1}^{L}\mathcal{I}_{l,\nu(l)}\Bigg]\to 0$$

as $\rho\to+\infty$, i.e. the corresponding contribution in (4.36) decays to zero faster than $\rho^{-(L+1)}$. Indeed, observe that we can write

$$(1+\alpha)I\geq\alpha\ln(1+a)$$

from where we will have

$$\rho^{L+1}p_{\boldsymbol{\nu}}=\rho^{L+1}\mathbb{E}\Bigg[\mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho)<R\}\prod_{l=1}^{L}\mathcal{I}_{l,\nu(l)}\Bigg]\leq \\ \leq\rho^{L+1}\mathbb{E}\Bigg[\mathbb{1}\{a<Z_0\}\prod_{l=1}^{L}\mathcal{I}_{l,\nu(l)}\Bigg]=\rho^{L+1}\mathbb{E}[\mathbb{1}\{a<Q\}]\mathbb{E}\Bigg[\prod_{l=1}^{L}\mathcal{I}_{l,\nu(l)}\Bigg],$$

with Q as defined in (4.8).

On the other hand, we also have

$$\mathbb{E}[\mathbb{1}\{a < Q\}] \leq \frac{Z}{\rho}$$

(Z denotes a generic positive non-zero constant that may vary at any appearance) and

$$\begin{aligned}
\mathbb{E}[\mathcal{I}_{l,1}] &= \Pr\{w_l \leq \rho^{\epsilon-1}\} \Pr\{y_l > \rho^{\delta-1}\} \leq \Pr\{w_l \leq \rho^{\epsilon-1}\} \leq Z\rho^{\epsilon-1} \\
\mathbb{E}[\mathcal{I}_{l,2}] &= \Pr\{w_l > \rho^{\epsilon-1}\} \Pr\{y_l \leq \rho^{\delta-1}\} \leq \Pr\{y_l \leq \rho^{\delta-1}\} \leq Z\rho^{\delta-1} \\
\mathbb{E}[\mathcal{I}_{l,3}] &= \Pr\{w_l > \rho^{\epsilon-1}\} \Pr\{y_l > \rho^{\delta-1}\} \leq 1 \\
\mathbb{E}[\mathcal{I}_{l,4}] &= \Pr\{w_l \leq \rho^{\epsilon-1}\} \Pr\{y_l \leq \rho^{\delta-1}\} \leq Z\rho^{\delta+\epsilon-2}.
\end{aligned}$$

Hence, we can write

$$\rho^{L+1} p_{\boldsymbol{\nu}} \leq Z\rho^{L} \rho^{(\epsilon-1)|\mathcal{I}_1|} \rho^{(\delta-1)|\mathcal{I}_2|} \rho^{(\delta+\epsilon-2)|\mathcal{I}_4|}$$

and, since $|\mathcal{I}_3| = 0$ and $|\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_4| = L$, we can write

$$\rho^{L+1} p_{\boldsymbol{\nu}} \leq Z\rho^{\epsilon|\mathcal{I}_1| + \delta|\mathcal{I}_2| + (\delta+\epsilon-1)|\mathcal{I}_4|}.$$

Now, $\epsilon < \delta$ implies

$$\begin{aligned}
\epsilon|\mathcal{I}_1| + \delta|\mathcal{I}_2| + (\delta+\epsilon-1)|\mathcal{I}_4| < \delta|\mathcal{I}_1| + \delta|\mathcal{I}_2| + (2\delta+-1)|\mathcal{I}_4| = \\
= \delta(L + |\mathcal{I}_4|) - |\mathcal{I}_4|.
\end{aligned}$$

Now, since $|\mathcal{I}_4| \geq 1$ and $\delta < (L+1)^{-1}$, we know that

$$\delta(L + |\mathcal{I}_4|) - |\mathcal{I}_4| \leq \delta(L+1) - 1 < 0$$

and, consequently,

$$\rho^{L+1} p_{\boldsymbol{\nu}} \to 0$$

as $\rho \to +\infty$.

Considerations on Lemmas 4.2, 4.3 and 4.5

As mentioned in Section 4.2.1, Theorem 4.3 is not proven in full formality here. In particular, at the time of the writing of this dissertation we are still polishing a completely rigorous proof for Lemmas 4.2, 4.3 and 4.5. In what follows, we try to motivate our intuition behind the three results. More specifically, we compute the limit

$$\lim_{\rho \to +\infty} \rho^{1+\Gamma(\boldsymbol{\nu})} p_{\boldsymbol{\nu}} \tag{4.37}$$

by means of the LDCT, even though we only conjecture that its hypotheses are verified.

Recall that

$$p_{\boldsymbol{\nu}} = \int_{R_+^{2L+1}} \mathbb{1}\{I(x,\mathbf{w},\mathbf{y};\rho) < R\} \prod_{l=1}^{L} \mathcal{I}_{l,\nu(l)} \cdot \\ \cdot f_{|h_s|^2}(x) f_{|h_u|^2}(\mathbf{w}) f_{|h_d|^2}(\mathbf{y})\, \mathrm{d}x\, \mathrm{d}\mathbf{w}\, \mathrm{d}\mathbf{y},$$

where $I(x,\mathbf{w},\mathbf{y};\rho)$ is as defined in (4.33). Now, with the change of variables $a = \rho x$, $\{b_l = z\rho w_l,\, l \in \mathcal{I}_1 \cup \mathcal{I}_4\}$ and $\{c_l = z\rho y_l,\, l \in \mathcal{I}_2 \cup \mathcal{I}_4\}$, the previous expression can be rewritten as

$$\begin{aligned} p_{\boldsymbol{\nu}} = \frac{1}{z^{\Gamma(\boldsymbol{\nu})}\rho^{1+\Gamma(\boldsymbol{\nu})}} \int_{R_+^{2L+1}} & \mathbb{1}\{I(a,\mathbf{b},\mathbf{w},\mathbf{c},\mathbf{y};\rho) < R\} \mathbb{1}\{a \le \rho^{\epsilon}\} \cdot \\ & \cdot \mathbb{1}\{b_l \le z\rho^{\epsilon},\, l \in \mathcal{I}_1 \cup \mathcal{I}_4\} \mathbb{1}\{w_l > \rho^{\epsilon-1},\, l \in \mathcal{I}_2 \cup \mathcal{I}_3\} \cdot \\ & \cdot \mathbb{1}\{c_l \le z\rho^{\delta},\, l \in \mathcal{I}_2 \cup \mathcal{I}_4\} \mathbb{1}\{y_l > \rho^{\delta-1},\, l \in \mathcal{I}_1 \cup \mathcal{I}_3\} \cdot \\ & \cdot f_{|h_s|^2}\Big(\frac{a}{\rho}\Big) f_{|h_u|^2}\Big(\frac{\mathbf{b}}{z\rho}\Big) f_{|h_u|^2}(\mathbf{w}) f_{|h_d|^2}\Big(\frac{\mathbf{c}}{z\rho}\Big) f_{|h_d|^2}(\mathbf{y})\, \mathrm{d}a\, \mathrm{d}\mathbf{b}\, \mathrm{d}\mathbf{w}\, \mathrm{d}\mathbf{c}\, \mathrm{d}\mathbf{y}. \end{aligned}$$

In the likely case where $\mathbb{1}\{I(a,\mathbf{b},\mathbf{w},\mathbf{c},\mathbf{y};\rho) < R\}$ implies that a, $\{b_l\}$ and $\{c_l\}$ are all upper-bounded, the limit in (4.37) can be computed by means of the LDCT giving the outage gains in Section 4.2.1. Indeed, since a, $\{b_l\}$ and $\{c_l\}$ live on a compact set, the product $f_{|h_s|^2}\big(\frac{a}{\rho}\big) f_{|h_u|^2}\big(\frac{\mathbf{b}}{z\rho}\big) f_{|h_d|^2}\big(\frac{\mathbf{c}}{z\rho}\big)$ is upper-bounded due to (4.6), while the variables $\{w_l\}$ and $\{y_l\}$ do not rise any issues since the probability density functions $f_{|h_u|^2}(\cdot)$ and $f_{|h_d|^2}(\cdot)$ do not depend on ρ and are integrable on all $\mathbb{R}_+$.

To uphold the above assumption we need to compute the limit of the spectral efficiency $I(a,\mathbf{b},\mathbf{w},\mathbf{c},\mathbf{y};\rho)$ for $\rho \to +\infty$. The result is different according to how the relays are distributed over the sets $\mathcal{I}_1$, $\mathcal{I}_2$, $\mathcal{I}_3$ and $\mathcal{I}_4$. From Appendix 3.A.3, the coefficients $\{\lambda_l\}$ are the L positive solutions to

$$z\rho \Bigg[\sum_{l\in\mathcal{I}_1} \frac{b_l y_l}{\alpha(1+b_l)\lambda - z\rho y_l} + \sum_{l\in\mathcal{I}_2} \frac{c_l w_l}{\alpha(1+z\rho w_l)\lambda - c_l} + \\ + \sum_{l\in\mathcal{I}_3} \frac{z\rho w_l y_l}{\alpha(1+z\rho w_l)\lambda - z\rho y_l} + \frac{1}{z\rho} \sum_{l\in\mathcal{I}_4} \frac{b_l c_l}{\alpha(1+b_l)\lambda - c_l} \Bigg] = 1 + a. \quad (4.38)$$

Case $\mathcal{I}_3 \neq \emptyset$. When $\mathcal{I}_3$ is not empty, that is there exists at least a relay experiencing a good channel quality both from the source and towards the destination, (4.38) has the following set of solutions when $\rho \to +\infty$:

- $\lambda_1, \ldots, \lambda_{|\mathcal{I}_2|} \to 0$ and $\lambda_l \rho \to r_l$, with r_l one of the $|\mathcal{I}_2|$ positive solutions to

$$\sum_{l\in\mathcal{I}_2} \frac{c_l}{r - \frac{c_l}{\alpha z w_l}} = \alpha z \sum_{l\in\mathcal{I}_3} w_l;$$

- $\lambda_{|\mathcal{I}_2|+1}, \ldots, \lambda_{L-|\mathcal{I}_1|-1} \to \bar{\lambda}_{|\mathcal{I}_2|+1}, \ldots, \bar{\lambda}_{L-|\mathcal{I}_1|-1}$, where $\bar{\lambda}_l$ is either one of the $|\mathcal{I}_3|-1$ positive solutions to

$$\sum_{l\in\mathcal{I}_3} \frac{y_l}{\bar{\lambda} - \frac{y_l}{\alpha w_l}} = 0$$

or one of the $|\mathcal{I}_4|$ values $\left\{ \frac{c_l}{\alpha(1+b_l)},\, l \in \mathcal{I}_4 \right\}$;

- $\lambda_{L-|\mathcal{I}_1|}, \ldots, \lambda_L \to +\infty$ and $\frac{\lambda_l}{\rho} \to \mu_l$, with μ_l one of the $|\mathcal{I}_1|+1$ positive solutions to

$$\sum_{l\in\mathcal{I}_1} \frac{b_l}{1+b_l} \frac{y_l}{\mu - \frac{z y_l}{\alpha(1+b_l)}} + \frac{1}{\mu} \sum_{l\in\mathcal{I}_3} y_l = \frac{\alpha}{z}(1+a).$$

Then, from (4.16) and (4.17), the following results can be derived after some algebra.

- When $\alpha < (|\mathcal{I}_1|+1)^{-1}$,
  - ϕ_1 tends to ϕ_1^A, the unique positive solution to

$$\frac{1-\alpha(|\mathcal{I}_1|+1)}{\phi_1^A} = z + \alpha \sum_{l=|\mathcal{I}_2|+1}^{L-|\mathcal{I}_1|-1} \frac{\bar{\lambda}_l}{1+\bar{\lambda}_l \phi_1^A};$$

  - ϕ_2 tends to ϕ_2^A, the unique positive solution to

$$\frac{1-\alpha|\mathcal{I}_1|}{\phi_2^A} = z + \alpha \sum_{l\in\mathcal{I}_3} \frac{y_l}{\alpha w_l + y_l \phi_2^A} + \alpha \sum_{l\in\mathcal{I}_4} \frac{c_l}{\alpha(1+b_l) + c_l \phi_2^A};$$

  - the spectral efficiency can be written as

$$(1+\alpha) I = \alpha(|\mathcal{I}_1|+1)\ln\rho - \alpha|\mathcal{I}_1|\ln\rho + o(\ln\rho), \tag{4.39}$$

    which tends to infinity as $\rho \to +\infty$.

- When[8] $(|\mathcal{I}_1|+1)^{-1} < \alpha < |\mathcal{I}_1|^{-1}$,
  - $\rho\phi_1 \to \phi_1^B$, where ϕ_1^B is the unique positive solution to

$$\phi_1^B = \left(\alpha \sum_{l=L-|\mathcal{I}_1|}^{L} \frac{\mu_l}{1+\mu_l \phi_1^B} \right)^{-1};$$

[8] For simplicity, we do not treat here the boundaries between the different regions (e.g. $\alpha = (|\mathcal{I}_1|+1)^{-1}$), which need particular attention.

  - ϕ_2 tends to ϕ_2^A as before;
  - the spectral efficiency can be written as

$$(1+\alpha)I = \ln\rho - \alpha|\mathcal{I}_1|\ln\rho + o(\ln\rho), \tag{4.40}$$

    which tends to infinity as $\rho \to +\infty$.

- When $\mathcal{I}_1$ is not empty and $\alpha > |\mathcal{I}_1|^{-1}$,
  - $\rho\phi_1 \to \phi_1^B$ as before,
  - $\rho\phi_2 \to \phi_2^B$, where ϕ_2^B is the unique positive solution to

$$\phi_2^B = \left(\alpha \sum_{l\in\mathcal{I}_1} \frac{zy_l}{\alpha(1+b_l) + zy_l\phi_2^B}\right)^{-1},$$

  - the spectral efficiency can be written as

$$(1+\alpha)I = \alpha\ln(1+a) + \ln\frac{\phi_2^B}{\phi_1^B} + \alpha \sum_{l=L-|\mathcal{I}_1|}^{L} \ln(1+\mu_l\phi_1^B) - \alpha\sum_{l\in\mathcal{I}_1} \ln\left[1 + \frac{zy_l}{\alpha(1+b_l)}\phi_2^B\right] + o(1). \tag{4.41}$$

Case $\mathcal{I}_3 = \emptyset$, $\mathcal{I}_2 \neq \emptyset$. Note that the results derived so far hold for all $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_4$, as long as $\mathcal{I}_3$ is not empty. We analyze now the cases when all relays undergo a deep fade in at least one of their channels ($\mathcal{I}_3 = \emptyset$). When $\mathcal{I}_2 \neq \emptyset$, and in the high-SNR regime, (4.38) implies that the coefficients $\{\lambda_l\}$ behave as follows:

- $\lambda_1, \ldots, \lambda_{|\mathcal{I}_2|-1} \to 0$ and $\lambda_l\rho \to r_l$, with r_l one of the $|\mathcal{I}_2| - 1$ positive solutions to

$$\sum_{l\in\mathcal{I}_2} \frac{c_l}{r - \frac{c_l}{\alpha z w_l}} = 0;$$

- $\lambda_{|\mathcal{I}_2|}, \ldots, \lambda_{L-|\mathcal{I}_1|} \to \bar{\lambda}_{|\mathcal{I}_2|}, \ldots, \bar{\lambda}_{L-|\mathcal{I}_1|}$, with $\bar{\lambda}_l$ one of the $|\mathcal{I}_4|+1$ positive solutions to

$$\frac{1}{\bar{\lambda}}\sum_{l\in\mathcal{I}_2} c_l + \sum_{l\in\mathcal{I}_4} \frac{b_l}{1+b_l}\frac{c_l}{\bar{\lambda} - \frac{c_l}{\alpha(1+b_l)}} = \alpha(1 + a + \sum_{l\in\mathcal{I}_1} b_l);$$

- $\lambda_{L-|\mathcal{I}_1|+1}, \ldots, \lambda_L \to +\infty$ and $\frac{\lambda_l}{\rho} \to \mu_l$, with μ_l one of the $|\mathcal{I}_1|$ positive solutions to

$$\sum_{l\in\mathcal{I}_1} \frac{b_l}{1+b_l}\frac{y_l}{\mu - \frac{zy_l}{\alpha(1+b_l)}} = \frac{\alpha}{z}(1+a).$$

Then, similarly to the previous case, the limit of the spectral efficiency depends on the value of the coding rate α.

- When $\alpha < |\mathcal{I}_1|^{-1}$,
  - ϕ_1 tends to ϕ_1^A, where ϕ_1^A is now the unique positive solution to
$$\frac{1-\alpha|\mathcal{I}_1|}{\phi_1^A} = z + \alpha \sum_{l=|\mathcal{I}_2|}^{L-|\mathcal{I}_1|} \frac{\bar{\lambda}_l}{1+\bar{\lambda}_l\phi_1^A},$$
  - ϕ_2 tends to ϕ_2^A, where ϕ_2^A is the unique positive solution to
$$\frac{1-\alpha|\mathcal{I}_1|}{\phi_2^A} = z + \alpha \sum_{l\in\mathcal{I}_4} \frac{c_l}{\alpha(1+b_l)+c_l\phi_2^A},$$
  - the spectral efficiency can be written as
$$\begin{aligned}(1+\alpha)I = \alpha\ln(1+a) + (1-\alpha|\mathcal{I}_1|)\ln\frac{\phi_2^A}{\phi_1^A} + z(\phi_1^A-\phi_2^A) + \\ + \alpha\sum_{l=|\mathcal{I}_2|}^{L-|\mathcal{I}_1|}\ln(1+\bar{\lambda}_l\phi_1^A) + \alpha\sum_{l=L-|\mathcal{I}_1|+1}^{L}\ln\mu_l + \\ - \alpha\sum_{l\in\mathcal{I}_1}\ln\frac{zy_l}{\alpha(1+b_l)} - \alpha\sum_{l\in\mathcal{I}_4}\ln\Big[1+\frac{c_l}{\alpha(1+b_l)}\phi_2^A\Big] + o(1). \quad (4.42)\end{aligned}$$
- When $\mathcal{I}_1$ is not empty and $\alpha > |\mathcal{I}_1|^{-1}$,
  - $\rho\phi_1 \to \phi_1^B$, where ϕ_1^B is the unique positive solution to
$$\phi_1^B = \left(\alpha\sum_{l=L-|\mathcal{I}_1|+1}^{L}\frac{\mu_l}{1+\mu_l\phi_1^B}\right)^{-1},$$
  - $\rho\phi_2 \to \phi_2^B$, where ϕ_2^B is the unique positive solution to
$$\phi_2^B = \left(\alpha\sum_{l\in\mathcal{I}_1}\frac{zy_l}{\alpha(1+b_l)+zy_l\phi_2^B}\right)^{-1},$$
  - the spectral efficiency can be written as
$$\begin{aligned}(1+\alpha)I = \alpha\ln(1+a) + \ln\frac{\phi_2^B}{\phi_1^B} + \alpha\sum_{l=L-|\mathcal{I}_1|+1}^{L}\ln(1+\mu_l\phi_1^B) \\ - \alpha\sum_{l\in\mathcal{I}_1}\ln\Big[1+\frac{zy_l\phi_2^B}{\alpha(1+b_l)}\Big] + o(1). \quad (4.43)\end{aligned}$$

Case $\mathcal{I}_3 = \emptyset$, $\mathcal{I}_2 = \emptyset$. When $\mathcal{I}_2$ is also empty, the coefficients $\{\lambda_l\}$ take on the following values in the limit for $\rho \to +\infty$:

- $\lambda_1, \ldots, \lambda_{|\mathcal{I}_4|} \to \bar{\lambda}_1 \ldots \bar{\lambda}_{|\mathcal{I}_4|}$, with $\bar{\lambda}_l$ one of the $|\mathcal{I}_4|$ positive solutions to

$$\sum_{l \in \mathcal{I}_4} \frac{b_l}{1+b_l} \frac{c_l}{\bar{\lambda} - \frac{c_l}{\alpha(1+b_l)}} = \alpha(1 + a + \sum_{l \in \mathcal{I}_1} b_l);$$

- $\lambda_{|\mathcal{I}_4|+1}, \ldots, \lambda_L \to +\infty$ and $\frac{\lambda_l}{\rho} \to \mu_l$, with μ_l one of the $|\mathcal{I}_1|$ positive solutions to

$$\sum_{l \in \mathcal{I}_1} \frac{b_l}{1+b_l} \frac{y_l}{\mu - \frac{z y_l}{\alpha(1+b_l)}} = \frac{\alpha}{z}(1+a).$$

The effect on the spectral efficiency is as listed hereafter.

- When $\alpha < |\mathcal{I}_1|^{-1}$,
  - ϕ_1 tends to ϕ_1^A, where ϕ_1^A is now the unique positive solution to

$$\frac{1 - \alpha|\mathcal{I}_1|}{\phi_1^A} = z + \alpha \sum_{l=1}^{|\mathcal{I}_4|} \frac{\bar{\lambda}_l}{1 + \bar{\lambda}_l \phi_1^A},$$

  - ϕ_2 tends to ϕ_2^A, where ϕ_2^A is now the unique positive solution to

$$\frac{1 - \alpha|\mathcal{I}_1|}{\phi_2^A} = z + \alpha \sum_{l \in \mathcal{I}_4} \frac{c_l}{\alpha(1+b_l) + c_l \phi_2^A},$$

  - the spectral efficiency can be written as

$$\begin{aligned}(1+\alpha)I = \alpha \ln(1+a) + (1 - \alpha|\mathcal{I}_1|) \ln \frac{\phi_2^A}{\phi_1^A} + z(\phi_1^A - \phi_2^A) + \\ + \alpha \sum_{l=1}^{|\mathcal{I}_4|} \ln(1 + \bar{\lambda}_l \phi_1^A) + \alpha \sum_{l=|\mathcal{I}_4|+1}^{L} \ln \mu_l + \\ - \alpha \sum_{l \in \mathcal{I}_1} \ln \frac{z y_l}{\alpha(1+b_l)} - \alpha \sum_{l \in \mathcal{I}_4} \ln\Big[1 + \frac{c_l}{\alpha(1+b_l)} \phi_2^A\Big] + o(1). \quad (4.44)\end{aligned}$$

- When $\mathcal{I}_1$ is not empty $\alpha > |\mathcal{I}_1|^{-1}$,
  - $\rho\phi_1 \to \phi_1^B$, where ϕ_1^B is the unique positive solution to

$$\phi_1^B = \left(\alpha \sum_{l=|\mathcal{I}_4|+1}^{L} \frac{\mu_l}{1 + \mu_l \phi_1^B} \right)^{-1},$$

– $\rho\phi_2 \to \phi_2^B$, where ϕ_2^B is the unique positive solution to

$$\phi_2^B = \left(\alpha \sum_{l \in \mathcal{I}_1} \frac{z y_l}{\alpha(1+b_l) + z y_l \phi_2^B} \right)^{-1},$$

– the spectral efficiency can be written as

$$(1+\alpha) I = \alpha \ln(1+a) + \ln \frac{\phi_2^B}{\phi_1^B} + \alpha \sum_{l=|\mathcal{I}_4|+1}^{L} \ln(1+\mu_l \phi_1^B) - \alpha \sum_{l \in \mathcal{I}_1} \ln\Big[1 + \frac{z y_l \phi_2^B}{\alpha(1+b_l)}\Big] + o(1). \quad (4.45)$$

Comments. The intuition behind Lemmas 4.2, 4.3 and 4.5 comes from a thorough analysis of the limiting behavior of the spectral efficiency. For example, equations (4.39) and (4.40) readily suggest Lemma 4.5.

Next, let $\mathcal{I}_4 = \emptyset$ in (4.42) and (4.44). Then, under the assumptions of Lemma 4.2, the large-SNR spectral efficiency is a finite quantity that diverges as any b_l, $l \in \mathcal{I}_1$, or any c_l, $l \in \mathcal{I}_2$, grows large. This fact brings the intuition behind Lemma 4.2.

With a similar analysis of (4.41), (4.43) and (4.45), we can reasonably conjecture that Lemma 4.3 holds true.

Appendix 4.C Proof for the AF relaying strategy with LMMSE receiver

In this appendix we prove Theorem 4.4. More specifically, we give a large-SNR approximation of the outage probability of

$$\mathit{SINR}(x, \mathbf{w}, \mathbf{y}; \rho) = x + z^2 \rho^2 \sum_{l=1}^{L} \frac{w_l y_l}{z\rho y_l + \alpha(1 + z\rho w_l)\frac{1}{\phi_1}}.$$

The proof follows the same guidelines as the previous one, that is each contribution

$$p_{\boldsymbol{\nu}} = \mathbb{E}\left[\mathbb{1}\{\mathit{SINR}(x, \mathbf{w}, \mathbf{y}; \rho) < Q\} \prod_{l=1}^{L} \mathcal{I}_{l,\nu(l)} \right]$$

of

$$\Pr[\mathit{SINR}(x, \mathbf{w}, \mathbf{y}; \rho) < Q] = \mathbb{E}[\mathbb{1}\{\mathit{SINR}(x, \mathbf{w}, \mathbf{y}; \rho) < Q\}] = \sum_{\boldsymbol{\nu} \in \{1,2,3,4\}^L} p_{\boldsymbol{\nu}} \quad (4.46)$$

Table 4.3: Association between relay partitions and relative lemma (LMMSE receiver).

<table>
<tr><th>#</th><th>$|\mathcal{I}_3|$</th><th>$|\mathcal{I}_4|$</th><th>$|\mathcal{I}_2|$</th><th>$|\mathcal{I}_1|$</th><th>Case $\alpha < \frac{1}{|\mathcal{I}_1|}$</th><th>Case $\alpha > \frac{1}{|\mathcal{I}_1|}$</th></tr>
<tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td colspan="2">—</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>0</td><td>>0</td><td rowspan="3">Lemma 4.6</td><td>Lemma 4.7</td></tr>
<tr><td>2</td><td>0</td><td>0</td><td>>0</td><td>0</td><td>—</td></tr>
<tr><td>3</td><td>0</td><td>0</td><td>>0</td><td>>0</td><td>Lemma 4.7</td></tr>
<tr><td>4</td><td>0</td><td>>0</td><td>0</td><td>0</td><td rowspan="4">Lemma 4.8</td><td>—</td></tr>
<tr><td>5</td><td>0</td><td>>0</td><td>0</td><td>>0</td><td></td></tr>
<tr><td>6</td><td>0</td><td>>0</td><td>>0</td><td>0</td><td>—</td></tr>
<tr><td>7</td><td>0</td><td>>0</td><td>>0</td><td>>0</td><td></td></tr>
<tr><td colspan="5"></td><td>Case $\alpha < \frac{1}{|\mathcal{I}_1|+1}$</td><td>Case $\alpha > \frac{1}{|\mathcal{I}_1|+1}$</td></tr>
<tr><td>8</td><td>>0</td><td>0</td><td>0</td><td>>0</td><td rowspan="4">Lemma 4.9</td><td rowspan="4">Lemma 4.10</td></tr>
<tr><td>9</td><td>>0</td><td>0</td><td>>0</td><td>>0</td></tr>
<tr><td>10</td><td>>0</td><td>>0</td><td>0</td><td>>0</td></tr>
<tr><td>11</td><td>>0</td><td>>0</td><td>>0</td><td>>0</td></tr>
<tr><td>12</td><td>>0</td><td>0</td><td>0</td><td>0</td><td colspan="2" rowspan="4">Lemma 4.11</td></tr>
<tr><td>13</td><td>>0</td><td>0</td><td>>0</td><td>0</td></tr>
<tr><td>14</td><td>>0</td><td>>0</td><td>0</td><td>0</td></tr>
<tr><td>15</td><td>>0</td><td>>0</td><td>>0</td><td>0</td></tr>
</table>

is considered and its vanishing speed is analyzed. The functions $\{\mathcal{I}_{l,i} : i = 1,2,3,4\}$ and the sets $\{\mathcal{I}_i\}$ are the same as above. Moreover, Lemma 4.1 holds true in this case also and, hence, we can always assume $a = \rho x \leq \rho^{\epsilon}$.

4.C.1 Main results

The asymptotic contribution of the L-tuple $\boldsymbol{\nu}$ can be described by one of the six lemmas below. Table 4.3 shows which partitions of the relay set each lemma applies to.

Lemma 4.6. *Assume that $\boldsymbol{\nu} \in \{1,2\}^L$ and $1 - \alpha|\mathcal{I}_1| > 0$. In this situation, the limit*

$$\lim_{\rho \to +\infty} \rho^{1+\Gamma(\boldsymbol{\nu})} p_{\boldsymbol{\nu}}$$

exists and is positive. Observe that $\Gamma(\boldsymbol{\nu}) = L$ and we achieve full diversity order $L+1$.

Lemma 4.7. *If $|\mathcal{I}_3| = |\mathcal{I}_4| = 0$ and $\alpha|\mathcal{I}_1| - 1 > 0$ (which implies $|\mathcal{I}_1| > 0$), then*

$$\rho^{|\mathcal{I}_1|+1} p_{\boldsymbol{\nu}} = \begin{cases} O(1) & \text{if } |\mathcal{I}_2| = 0, \\ o(1) & \text{if } |\mathcal{I}_2| > 0. \end{cases}$$

Lemma 4.8. *When* $|\mathcal{I}_4| > 0$ *and* $|\mathcal{I}_3| = 0$*, then*

$$\rho^{L+1} p_{\boldsymbol{\nu}} \to 0,$$

i.e. the corresponding contribution in (4.46) decays to zero faster than $\rho^{-(L+1)}$*.*

Lemma 4.9. *If* $|\mathcal{I}_3| > 0$ *and* $1 - \alpha(|\mathcal{I}_1| + 1) > 0$*, then*

$$p_{\boldsymbol{\nu}} = 0$$

for a sufficiently high ρ*. These terms will never contribute to the large-SNR outage probability.*

Lemma 4.10. *If* $|\mathcal{I}_3| > 0$*,* $|\mathcal{I}_1| > 0$ *and* $\alpha(|\mathcal{I}_1| + 1) - 1 > 0$*, then*

$$\rho^{|\mathcal{I}_1|+1} p_{\boldsymbol{\nu}} = \begin{cases} O(1) & \text{if } |\mathcal{I}_2| + |\mathcal{I}_4| = 0, \\ o(1) & \text{if } |\mathcal{I}_2| + |\mathcal{I}_4| > 0. \end{cases}$$

Lemma 4.11. *If* $|\mathcal{I}_3| > 0$ *and* $|\mathcal{I}_1| = 0$*, then*

$$\rho p_{\boldsymbol{\nu}} = \begin{cases} 0 & \text{if } \alpha < 1 + \dfrac{1}{Q}, \\ o(1) & \text{if } \alpha > 1 + \dfrac{1}{Q} \text{ and } |\mathcal{I}_2| + |\mathcal{I}_4| > 0, \\ O(1) & \text{if } \alpha > 1 + \dfrac{1}{Q} \text{ and } |\mathcal{I}_2| + |\mathcal{I}_4| = 0. \end{cases}$$

According to these lemmas, the speed of convergence to zero is reported in Table 4.4 for all the possible situations.

Proof of Theorem 4.4

As for the ML case, the results of Theorem 4.4 are proven separately according to the value taken by the coding rate α.

Proposition 4.5. *If* $\alpha < L^{-1}$ *then*

$$\lim_{\rho \to +\infty} \rho^{L+1} \mathbb{E}[\mathbb{1}\{SINR(x, \mathbf{w}, \mathbf{y}; \rho) < Q\}]$$

is finite and non-zero.

Proof. To start, note that Situations 12–15 in Table 4.4 can be readily discarded since $\alpha < 1$. Now, consider a L-tuple $\boldsymbol{\nu}$ such that $|\mathcal{I}_3| = 0$ (Situations 1–7 in Table 4.4). Since $|\mathcal{I}_1| \leq L$, it is $\alpha < \frac{1}{L} \leq \frac{1}{|\mathcal{I}_1|}$ always. On the other hand, when $|\mathcal{I}_3| > 0$, $|\mathcal{I}_1| < L$ implies $\alpha < \frac{1}{L} \leq \frac{1}{|\mathcal{I}_1|+1}$. Thus, we only have to look at the first column of Table 4.4. Due to the different speeds of convergence to zero, the only L-tuples contributing to the limit

$$\lim_{\rho \to +\infty} \rho^{L+1} \sum_{p_{\boldsymbol{\nu}} \in \{1,2,3,4\}^L} p_{\boldsymbol{\nu}},$$

are those belonging to Situations 1–3. According to Lemma 4.6, the limit is finite and non-zero. □

Table 4.4: Convergence speed associated to each relay partition (LMMSE receiver).

<table>
<tr><th>#</th><th>$\Gamma(p_{\boldsymbol{\nu}})$</th><th>Case $\alpha < \frac{1}{|\mathcal{I}_1|}$</th><th>Case $\alpha > \frac{1}{|\mathcal{I}_1|}$</th></tr>
<tr><td>0</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td>1</td><td>$|\mathcal{I}_1| = L$</td><td rowspan="3">$p_{\boldsymbol{\nu}} = O\big(\rho^{-(L+1)}\big)$</td><td>$p_{\boldsymbol{\nu}} = O\big(\rho^{-(L+1)}\big)$</td></tr>
<tr><td>2</td><td>$|\mathcal{I}_2| = L$</td><td>—</td></tr>
<tr><td>3</td><td>$|\mathcal{I}_1| + |\mathcal{I}_2| = L$</td><td>$p_{\boldsymbol{\nu}} = o\big(\rho^{-(|\mathcal{I}_1|+1)}\big)$</td></tr>
<tr><td>4</td><td>$2|\mathcal{I}_4| > L$</td><td rowspan="4">$p_{\boldsymbol{\nu}} = o\big(\rho^{-(L+1)}\big)$</td><td>—</td></tr>
<tr><td>5</td><td>$|\mathcal{I}_1| + 2|\mathcal{I}_4| > L$</td><td rowspan="3"></td></tr>
<tr><td>6</td><td>$|\mathcal{I}_2| + 2|\mathcal{I}_4| > L$</td><td>—</td></tr>
<tr><td>7</td><td>$|\mathcal{I}_1| + |\mathcal{I}_2| + 2|\mathcal{I}_4| > L$</td></tr>
<tr><td></td><td></td><td>Case $\alpha < \frac{1}{|\mathcal{I}_1|+1}$</td><td>Case $\alpha > \frac{1}{|\mathcal{I}_1|+1}$</td></tr>
<tr><td>8</td><td>$|\mathcal{I}_1| < L$</td><td rowspan="4">$p_{\boldsymbol{\nu}} = 0$ for $\rho > \rho_0$</td><td>$p_{\boldsymbol{\nu}} = O\big(\rho^{-(|\mathcal{I}_1|+1)}\big)$</td></tr>
<tr><td>9</td><td>$|\mathcal{I}_1| + |\mathcal{I}_2| < L$</td><td rowspan="3">$p_{\boldsymbol{\nu}} = o\big(\rho^{-(|\mathcal{I}_1|+1)}\big)$</td></tr>
<tr><td>10</td><td>$|\mathcal{I}_1| + 2|\mathcal{I}_4|$</td></tr>
<tr><td>11</td><td>$|\mathcal{I}_1| + |\mathcal{I}_2| + 2|\mathcal{I}_4|$</td></tr>
<tr><td>12</td><td>0</td><td colspan="2" rowspan="4">$p_{\boldsymbol{\nu}} = O(\rho)\mathbb{1}\{\alpha > 1 + \frac{1}{Q}\}$</td></tr>
<tr><td>13</td><td>$|\mathcal{I}_2| < L$</td></tr>
<tr><td>14</td><td>$2|\mathcal{I}_4|$</td></tr>
<tr><td>15</td><td>$|\mathcal{I}_2| + 2|\mathcal{I}_4|$</td></tr>
</table>

Proposition 4.6. *If* $(M+1)^{-1} < \alpha \leq M^{-1}$*, with* $M = 1, \ldots, L-1$*, then*

$$\lim_{\rho \to +\infty} \rho^{M+1} \mathbb{E}[\mathbb{1}\{SINR(x, \mathbf{w}, \mathbf{y}; \rho) < Q\}]$$

is finite and non-zero.

Proof. As for the previous case, we can discard Situations 12–15. First, consider the L-tuples $\boldsymbol{\nu}$ such that $|\mathcal{I}_1| \leq M-1$. Then $\alpha \leq \frac{1}{M} \leq \frac{1}{|\mathcal{I}_1|+1} < \frac{1}{|\mathcal{I}_1|}$ and, according to the first column of Table 4.4, the contribution $p_{\boldsymbol{\nu}}$ tends to zero faster than or as fast as $\rho^{-(L+1)}$. This means

$$\lim_{\rho \to +\infty} \rho^{M+1} p_{\boldsymbol{\nu}} = 0.$$

Consider now the cases where $|\mathcal{I}_1| \geq M+1$, meaning $\alpha > \frac{1}{M+1} \geq \frac{1}{|\mathcal{I}_1|} > \frac{1}{|\mathcal{I}_1|+1}$. The results summarized in the second column of Table 4.4 imply

that the contributions of these cases converge to zero faster than or as fast as $\rho^{-(|\mathcal{I}_1|+1)}$. As a consequence, since $\rho^{|\mathcal{I}_1|+1} > \rho^{M+1}$

$$\lim_{\rho\to+\infty} \rho^{M+1} p_{\boldsymbol{\nu}} = 0.$$

Finally, when $|\mathcal{I}_1| = M$ and $\frac{1}{|\mathcal{I}_1|+1} < \alpha \leq \frac{1}{|\mathcal{I}_1|}$, we observe that Situations 1–7 (i.e. $|\mathcal{I}_3| = 0$) can be discarded since they converge to zero faster than or as fast as $\rho^{-(L+1)}$ (first column of Table 4.4) and, hence,

$$\lim_{\rho\to+\infty} \rho^{M+1} p_{\boldsymbol{\nu}} = 0.$$

To conclude, we only need to apply Lemma 4.10 to the cases with $|\mathcal{I}_1| = M$ and $|\mathcal{I}_3| > 0$. As a result,

$$\lim_{\rho\to+\infty} \rho^{M+1} p_{\boldsymbol{\nu}}$$

is finite and non-zero for all relay partitions belonging to Situation 8 and vanishes otherwise. □

Proposition 4.7. *If $1 < \alpha \leq \alpha_{th}$, then*

$$\lim_{\rho\to+\infty} \rho^2 \mathbb{E}[\mathbb{1}\{SINR(x, \mathbf{w}, \mathbf{y}; \rho) < Q\}]$$

is finite and non-zero.

Proof. Recall that $Q = Q(\alpha) = \exp\left(\frac{1+\alpha}{\alpha} R\right) - 1$ and $\alpha_{th} = 1 + 1/Q(\alpha_{th})$. Then, Situations 12–15 can be neglected in this case also.

When $|\mathcal{I}_1| > 0$, we have $\alpha > \frac{1}{|\mathcal{I}_1|} > \frac{1}{|\mathcal{I}_1|+1}$. Thus, we are interested in the second column of Table 4.4 only. Observe that, in all situations, the contribution $p_{\boldsymbol{\nu}}$ converges to zero faster than or as fast as $\rho^{-(|\mathcal{I}_1|+1)}$. Then, we always have

$$\lim_{\rho\to+\infty} \rho^2 p_{\boldsymbol{\nu}} = 0$$

except when $|\mathcal{I}_1| = 1$ and $|\mathcal{I}_3| = L - 1$. In this last case, indeed, $\rho^2 p_{\boldsymbol{\nu}}$ converges to a finite non-zero quantity due to Lemma 4.10.

On the other hand, when $|\mathcal{I}_1| = 0$, it is $\alpha|\mathcal{I}_1| < 1$ and the contribution to the outage probability vanishes faster than or as fast as $\rho^{-(L+1)}$. This implies that

$$\lim_{\rho\to+\infty} \rho^2 p_{\boldsymbol{\nu}} = 0.$$

Note that the previous limit is finite and non-zero in the special case where $|\mathcal{I}_2| = L = 1$. □

Proposition 4.8. *If $\alpha > \alpha_{th}$, then*

$$\lim_{\rho\to+\infty} \rho^2 \mathbb{E}[\mathbb{1}\{SINR(x, \mathbf{w}, \mathbf{y}; \rho) < Q\}]$$

is finite and non-zero.

Proof. Since for $\alpha > \alpha_{th}$ it holds $\alpha > 1 + \frac{1}{Q}$, Lemma 4.11 implies that

$$\lim_{\rho \to +\infty} \rho \sum p_{\boldsymbol{\nu}}$$

is finite and non-zero when the sum is done over all the possible L-tuples such that $|\mathcal{I}_1| = 0$ and $|\mathcal{I}_3| > 0$ (Situations 12–15 in Table 4.4). According to the proofs of the previous results, the contribution of all the other L-tuples converges to zero faster than ρ^{-1} and, hence, they are negligible in terms of outage probability. □

Proofs of Lemmas 4.6 to 4.11

To conclude the proof, we still need to demonstrate all the lemmas appearing in Table 4.3. Unfortunately, as for the ML case in Appendix 4.B, at this point only Lemma 4.8 has a rigorous proof (which is basically the same as that of Lemma 4.4). The results stated by the other lemmas may be reasonably assumed to be true after computing the limit for $\rho \to +\infty$ of the SINR in all the possible relay partitions over the sets $\mathcal{I}_1, \ldots, \mathcal{I}_4$. The analysis is not reported here since very similar to what has been explained in Section 4.B.3 for the ML receiver.

Chapter 5

Reducing Interference: Isometric LD-STBC

The results of the two previous chapters have shown that distributed space-time block coding based on linear-dispersion matrices with random i.i.d. entries is an interesting solution for relay networks. More specifically, relays implementing the i.i.d. LD-STBC are capable of improving the spectral efficiency of the direct link and, more important, they introduce spatial diversity.

In this chapter, we try to improve the performance of the randomized LD-STBC. For this purpose, we present a new code whose linear-dispersion matrices are still independently generated. The entries of each relay matrix, however, are not independent of one another anymore. To understand how we can obtain some benefit by introducing some structure in the coding matrices, let us recall that the symbol vector transmitted by relay l is

$$\mathbf{C}_l\mathbf{s} = \sum_{k=1}^{K} \mathbf{c}_{l,k} s_k,$$

where $\mathbf{c}_{l,k}$ is the k-th column of the $N \times K$ linear-dispersion matrix $\mathbf{C}_l$ and s_k the k-th symbol of the source vector $\mathbf{s}$. Thus, a straightforward analogy with DS/CDMA systems [57] suggests that each relay can spread and multiplex the source symbols free of interference if the columns of the linear-dispersion matrix are mutually orthogonal.

Thus, when $N \geq K$ or, equivalently, when the coding rate $\alpha = K/N$ is lower than or equal to one, we constrain the columns of the coding matrices to be orthogonal to one another. More specifically, for the l-th relay, the signatures $\{\mathbf{c}_{l,k}\}_{k=1}^{K}$ are K columns of an $N \times N$ random unitary matrix. Then, the analysis is extended to the case $\alpha > 1$ by building the matrices $\mathbf{C}_l$ with N rows of a $K \times K$ random unitary matrix.

In what follows, the performance of the new code is evaluated in terms of spectral efficiency and compared to the performance of the i.i.d. code. For

simplicity's sake, only the DF relaying strategy is considered, thus avoiding the issues with the forwarded noise. As usual in this dissertation, the comparison is carried out in the asymptotic regime for $K, N \to +\infty$, $K/N \to \alpha$, where the spectral efficiency depends on the code through its coding rate α only and not on the specific linear-dispersion matrices. Note, however, that the random-matrix-theory tools employed in the previous chapters for i.i.d. codes do not apply here. Section 5.3 introduces the *rectangular free additive convolution*, a tool of free-probability theory which allows to compute the distribution of the singular values of a sum of rectangular matrices.

Synchronicity issues

Before starting the analysis, it is worth to mention a drawback of replacing i.i.d. codes for more structured ones.

Even though the issue was not addressed in the previous chapters, i.i.d. LD-STBC's are robust to little misalignment in relay transmissions. In other words, as long as code chips stay synchronized, small relative transmission delays are tolerated. Indeed, as depicted in Figure 5.1, it is enough to apply a detection window to obtain truncated versions of the linear-dispersion matrices with the same statistical properties. Moreover, if the signature length N is much larger than the maximum delay Δ, the previous results hold directly since $\frac{K}{N-\Delta} \approx \frac{K}{N} = \alpha$. Conversely, if Δ and N have the same order of magnitude, the results in Chapters 3 and 4 give a straightforward lower-bound on spectral efficiency and outage probability. It suffices to fix $\alpha = \frac{K}{N-\Delta}$ and remember that K information symbols are sent over $K+N+\Delta$ channel accesses. In this situation, however, further investigation is needed to understand how to extract information from the 2Δ samples removed by the detection window, since that information would not be negligible.

It can be readily seen that this property is lost when adding the new orthogonality constraint to the linear-dispersion matrices. Indeed, the sequences obtained by truncating orthogonal signatures do not need to be orthogonal anymore.

5.1 Signal model

The new choice for the coding matrices does not modify the signal model, which is exactly the same as the one presented in Section 4.1 for the DF relay channel. Observe that we will limit our analysis to the spectral efficiency, so that it is not interesting to consider all the possible decoding sets. For this reason, we assume in what follows that L relays are able to decode the source message and, thus, are allowed to join the second transmission phase.

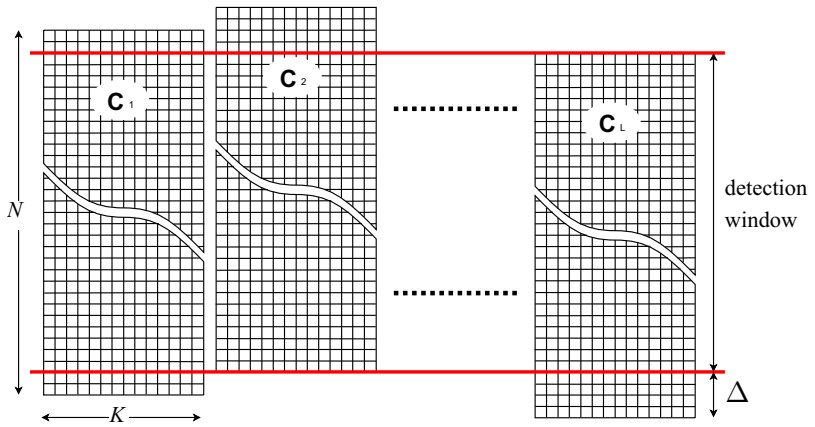


Figure 5.1: Misalignment of linear-dispersion matrices with asynchronous relay transmissions.

The resulting signal model can hence be written as

$$\mathbf{d} = \begin{bmatrix} h_s \mathbf{I}_K \\ \sum_{l=1}^{L} g_l h_{dl} \mathbf{C}_l \end{bmatrix} \mathbf{s} + \mathbf{n} = \begin{bmatrix} h_s \mathbf{I}_K \\ \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}} \end{bmatrix} \mathbf{s} + \mathbf{n}, \tag{5.1}$$

where, as before, $\mathbf{s} = [s_1 \ \cdots \ s_K]^T$ denotes the vector containing the message from the source and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_d^2 \mathbf{I}_N)$ represents the additive white Gaussian noise. The symbols s_k, $k = 1, \ldots, K$, are assumed i.i.d., with zero mean ($\mathbb{E}[s_k] = 0$) and variance P_s ($\mathbb{E}[|s_k|^2] = P_s$). Observe that, to simplify the subsequent notation, we prefer to gather the channel gains and the linear-dispersion matrices in a different way with respect to Section 3.1.1, namely

$$\widetilde{\mathbf{\Psi}} = \begin{bmatrix} g_1 h_{d1} \mathbf{I}_N & \cdots & g_L h_{dL} \mathbf{I}_N \end{bmatrix} \qquad \widetilde{\mathbf{C}} = \begin{bmatrix} \mathbf{C}_1^T & \cdots & \mathbf{C}_L^T \end{bmatrix}^T. \tag{5.2}$$

5.1.1 The coding matrices

As mentioned in the introduction, the intention here is to reduce the total interference by eliminating the interference generated within each relay. For this purpose, DS/CDMA experience suggests to use linear-dispersion matrices with orthogonal columns. Note that interference will not completely vanish since orthogonality is required only within each individual relay. Extending the constraint across all the relays would imply a very significant loss of flexibility, since all matrices would have to be jointly designed to be mutually orthogonal. Here, we are more interested in a dynamic system where the number of active terminals can vary without significantly jeopardizing the global coding scheme. Furthermore, global orthogonality is equivalent to TDMA, which is shown to be outperformed by STC in [32] (see also Section 4.1.4).

As a result, we model the $N \times K$ matrices $\{\mathbf{C}_l : l = 1, \ldots, L\}$ as mutually independent random matrices with orthogonal columns. More specifically, each coding matrix is constructed by selecting K different columns of a $N \times N$ Haar-distributed unitary matrix, i.e. a random unitary matrix whose distribution is invariant by left- or right-multiplication by a constant unitary matrix (we say that it is bi-unitarily invariant). We will refer to this model as Haar-distributed codes or, equivalently, as random isometric codes.

A matrix with orthogonal columns must be such that $K \leq N$, i.e. $\alpha = K/N \leq 1$. For completeness, the analysis is extended to the case $\alpha > 1$ by considering coding matrices with orthogonal rows (N rows of a $K \times K$ Haar-distributed unitary matrix). A scaling factor $\sqrt{\alpha}$, such that $\mathbf{C}_l \mathbf{C}_l^H = \alpha \mathbf{I}_N$, must be included to guarantee that the same power constraint as in the i.i.d. case is satisfied, which leads to a fair comparison of the results for the two different choices of the dispersion matrices.

To conclude this introduction, let us mention that Haar-distributed random unitary matrices are quite simple to generate (see Appendix 5.C) and that isometric codes have been extensively studied in the literature (see, for instance, [91–93]).

5.2 Spectral efficiency

As we did for the i.i.d. codes, we assume that the receiver has perfect channel state information of all the links and that it knows the relay linear-dispersion matrices. Under these hypotheses, we analyze again the spectral efficiency of the LMMSE and the ML receivers.

5.2.1 The LMMSE receiver

Following the same reasoning as in Section 3.3, the SINR at the output of an LMMSE receiver can be expressed by

$$SINR = \frac{P_s}{\sigma_d^2}|h_s|^2 + \\ + \frac{P_s}{\sigma_d^2}\mathbf{c}_1^H \widetilde{\mathbf{\Psi}}^H \left(\frac{P_s/\sigma_d^2}{1 + P_s|h_s|^2/\sigma_d^2} \widetilde{\mathbf{\Psi}} \mathbf{D}\mathbf{D}^H \widetilde{\mathbf{\Psi}}^H + \mathbf{I}_N \right)^{-1} \widetilde{\mathbf{\Psi}} \mathbf{c}_1, \quad (5.3)$$

where we focus on symbol s_1 without loss of generality and where we introduce the column vector $\mathbf{c}_1$ and the matrix $\mathbf{D}$ such that $\widetilde{\mathbf{C}} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{D} \end{bmatrix}$.

As for the i.i.d. case, the above SINR is a random quantity, since it intrinsically depends on the randomly generated coding matrices $\{\mathbf{C}_l\}$. In other words, for each realization of the code, the system performs differently. However, it will be shown in Section 5.3 that, if we increase the dimensions of the linear-dispersion matrices and keep constant the coding rate α, the

SINR in (5.3) converges to the deterministic value $SINR^{\text{Haar}}$, namely

$$\lim_{\substack{K,N\to+\infty \\ \frac{K}{N}\to\alpha}} SINR = SINR^{\text{Haar}}$$

in probability. The convergence is fast also in this case and $SINR^{\text{Haar}}$ can be used as an accurate approximation of the SINR of finite systems, even for not-so-large linear dispersion matrices (see Section 5.2.4 for details).

The formal proof of the convergence result is given in Section 5.3. Here, we only sketch its main points to give the expression of $SINR^{\text{Haar}}$. Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_K$ be the K real non-negative eigenvalues of the $K \times K$ interference matrix $\widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$. Then, the empirical distribution of the eigenvalues defined in (2.1) can be written as follows[1]:

$$\nu_N^2 = \frac{1}{K} \sum_{k=1}^{K} \delta_{\lambda_k} \tag{5.4}$$

where δ_λ is the Dirac distribution (mass point) at λ. Since the eigenvalues are generally random, their empirical distribution is also random. However, in Section 5.3 we show that ν_N^2 converges weakly in probability to a new deterministic distribution ν^2 when $K = \alpha N \to +\infty$, i.e. when K and N grow indefinitely while their ratio tends to α. More specifically, ν_N^2 converges weakly in probability to ν^2 (and we write $\nu_N^2 \xrightarrow{\text{P}} \nu^2$) if

$$\lim_{\substack{K,N\to+\infty \\ \frac{K}{N}\to\alpha}} \int f(t)\nu_N^2(\,\mathrm{d}t) = \int f(t)\nu^2(\,\mathrm{d}t) \tag{5.5}$$

in probability for any measurable function $f(\cdot)$.

In order to prove the convergence of the SINR, we will use the moment generating function[2] of ν^2, defined as

$$M_{\nu^2}(z) = \int \frac{zt}{1-zt} \nu^2(\,\mathrm{d}t), \quad z \in \mathbb{C} \setminus \mathbb{R}_+. \tag{5.6}$$

As proven in Section 5.3, the asymptotic SINR can be written as

$$SINR^{\text{Haar}} = \frac{P_s}{\sigma_d^2}|h_s|^2 + \frac{P_s}{\sigma_d^2} \frac{\eta^{\text{Haar}}}{1-\chi\eta^{\text{Haar}}}, \tag{5.7}$$

where

$$\chi = \frac{P_s/\sigma_d^2}{1 + P_s|h_s|^2/\sigma_d^2}.$$

[1] We use the square in ν_N^2 in order to emphasize that it is a law on the eigenvalues, and not on the singular values.

[2] Note that this is not the classical moment generating function $\mathbb{E}_X[e^{zX}]$.

and

$$\eta^{\text{Haar}} = -\frac{1}{\chi} M_{\nu^2}(-\chi).$$

Since the expression of the SINR given in (5.7) is independent of the actual intended symbol s_k, we conclude that all symbols experiment the same asymptotic SINR. Hence, by applying Shannon's formula and due to the continuity of the logarithm, we can state that the spectral efficiency (measured in nat/s/Hz) of the LMMSE receiver tends to

$$I_{\text{LMMSE}}^{\text{Haar}} = \frac{\alpha}{1+\alpha} \ln\big(1 + SINR^{\text{Haar}}\big) \tag{5.8}$$

in probability when $K = \alpha N \to +\infty$. The factor $\frac{\alpha}{1+\alpha} = \frac{K}{K+N}$ takes into account the fact that a total of $(K+N)$ channel accesses are employed to transmit only K information symbols.

5.2.2 The ML receiver

From the signal model in (5.1) it is straightforward to realize that the system is characterized by colored interference also when employing Haar-distributed codes, and then linear filters are sub-optimal receivers. To extract all the information contained in the received signal $\mathbf{d}$, the ML receiver is needed. Assuming independent Gaussian coding at the source, the spectral efficiency of the ML receiver in our scenario is known to be [10]

$$\begin{aligned} I_{\text{ML}} &= \frac{1}{K+N} \ln\det\left(\mathbf{I}_{K+N} + \frac{P_s}{\sigma_d^2}\begin{bmatrix} h_s\mathbf{I}_K \\ \widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}} \end{bmatrix}\begin{bmatrix} h_s^*\mathbf{I}_K & \widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H \end{bmatrix}\right) \\ &= \frac{\alpha}{1+\alpha}\ln\Big(1+\frac{P_s}{\sigma_d^2}|h_s|^2\Big) + \frac{1}{K+N}\ln\det\Big(\mathbf{I}_K + \chi\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}\Big), \end{aligned}$$

in nats per degree of freedom. As before, this is a random quantity depending on the random coding matrix $\widetilde{\mathbf{C}}$. However, it can be shown that it tends in probability to the deterministic spectral efficiency

$$I_{\text{ML}}^{\text{Haar}} = \frac{\alpha}{1+\alpha}\ln\Big(1+\frac{P_s}{\sigma_d^2}|h_s|^2\Big) + \frac{\alpha}{1+\alpha}\int_{-\chi}^{0}\frac{M_{\nu^2}(z)}{z}\,\mathrm{d}z \tag{5.9}$$

when the dimensions K and N of the matrices $\mathbf{C}_l$ tend to infinity at the same rate, i.e. with constant ratio $\alpha = K/N$.

The previous expression can be easily derived by recalling that

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln\det(\mathbf{I}+x\mathbf{A}) = \frac{K}{x} - \frac{1}{x^2}\mathrm{tr}\left\{\left(\mathbf{A}+\frac{1}{x}\mathbf{I}\right)^{-1}\right\} = \frac{K}{x}\left[1 - \frac{1}{K}\mathrm{tr}\Big\{(\mathbf{I}+x\mathbf{A})^{-1}\Big\}\right]$$

for any square matrix $\mathbf{A}$. Then, the spectral efficiency can be written as

$$I_{\mathrm{ML}} = \frac{\alpha}{1+\alpha}\ln\Big(1+\frac{P_s}{\sigma_d^2}|h_s|^2\Big) + \\ + \frac{\alpha}{1+\alpha}\int_{-\chi}^{0}\frac{1}{z}\Big[\frac{1}{K}\mathrm{tr}\Big\{\Big(\mathbf{I}_K - z\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}\Big)^{-1}\Big\} - 1\Big]\,\mathrm{d}z,$$

since $\ln\det\mathbf{I} = 0$. The asymptotic spectral efficiency (5.9) can be obtained[3] by noting that

$$\begin{aligned}\frac{1}{K}\mathrm{tr}\Big\{\Big(\mathbf{I}_K - z\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}\Big)^{-1}\Big\} - 1 \\ &= \frac{1}{K}\mathrm{tr}\Big\{z\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}\Big(\mathbf{I}_K - z\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}\Big)^{-1}\Big\} \\ &= \frac{1}{K}\sum_{k=1}^{K}\frac{z\lambda_k}{1-z\lambda_k} \\ &= \int\frac{zt}{1-zt}\nu_N^2(\,\mathrm{d}t),\end{aligned}$$

and that the last expression tends in probability to the moment generating function (5.6) when $K = \alpha N \to +\infty$, as proven in the next section.

5.2.3 The i.i.d. coding scheme

As explained above, the main motivation behind Haar-distributed random coding is the interference reduction with respect to i.i.d. coding analyzed in the previous chapters. When L relays participate to the second transmission phase, the DF spectral efficiencies derived in Section 4.1 can be written as follows:

$$I_{\mathrm{LMMSE}} \overset{\text{a.s.}}{\to} I_{\mathrm{LMMSE}}^{\mathrm{iid}} = \frac{\alpha}{1+\alpha}\ln\bigg(1+\frac{P_s}{\sigma_d^2}|h_s|^2+\frac{P_s\sum_{l=1}^{L}|g_l h_{dl}|^2}{\beta\sigma_d^2}\bigg), \quad (5.10a)$$

$$\begin{aligned}I_{\mathrm{ML}} \overset{\text{a.s.}}{\to} I_{\mathrm{ML}}^{\mathrm{iid}} = \frac{\alpha}{1+\alpha}\ln\bigg(1+\frac{P_s}{\sigma_d^2}|h_s|^2+\frac{P_s\sum_{l=1}^{L}|g_l h_{dl}|^2}{\beta\sigma_d^2}\bigg) \\ + \frac{1}{1+\alpha}\bigg(\ln\beta+\frac{1}{\beta}-1\bigg),\end{aligned} \quad (5.10b)$$

[3]Formally, one should show that the argument of the integral is upper-bounded by a positive integrable function before taking the limit. However, this step is a straightforward consequence of, e.g., Montel's theorem [94] and is therefore omitted.

where β is the positive solution to $\beta = 1 + \alpha\beta\frac{\chi\sum_{l=1}^{L}|g_l h_{dl}|^2}{\beta+\chi\sum_{l=1}^{L}|g_l h_{dl}|^2}$, namely

$$\beta = \frac{1}{2}\left[1-(1-\alpha)\chi\sum_{l=1}^{L}|g_l h_{dl}|^2 + \right. \\ \left. + \sqrt{\left(1-(1-\alpha)\chi\sum_{l=1}^{L}|g_l h_{dl}|^2\right)^2 + 4\chi\sum_{l=1}^{L}|g_l h_{dl}|^2}\right].$$

Observe that the limits in (5.10) hold almost surely and not only in probability as those derived above for the isometric coding scheme. It seems only natural to conjecture that these convergence results also hold in the almost-sure sense. However, the mathematical background used here has only been able to establish convergence in probability; besides, the difference between the two convergence modes has no real importance in practical aspects.

5.2.4 Finite-dimensional systems vs. large systems

Summarizing the results above, and similarly to what happened with i.i.d. linear dispersion matrices, the Haar-distributed coding scheme behaves (converges in probability to) a deterministic system when the size of the randomly-generated coding matrices grows large keeping constant the coding rate α. The asymptotic spectral efficiency is given by (5.8) or (5.9), according to the chosen receiver. Observe that both expressions depend on α only and not on K or N directly. Once again, these limiting values are excellent approximations of the finite-dimensional codes, even for not-so-large linear-dispersion matrices. For example, in Figure 5.2, we represent the average spectral efficiency over one thousand different realizations of the codes, together with the corresponding standard deviation. All the values are normalized with respect to the asymptotic spectral efficiency. The coding rate α is fixed to $4/5$, but the dimensions of the code increase with M, namely $K = 4M$ and $N = 5M$. Note that for $M = 10$, which corresponds to $K = 40$ and $N = 50$, the error is lower than 2%.

5.3 General case and asymptotic results

In the previous section, the spectral efficiency of the considered system was said to converge in probability to a deterministic constant when the dimensions K and N of the coding matrices grow indefinitely but with constant ratio $\alpha = K/N$. Furthermore, the limit was expressed in terms of the asymptotic distribution ν^2 of the eigenvalues of the interference matrix $\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}$. In what follows, we will first discuss the convergence of

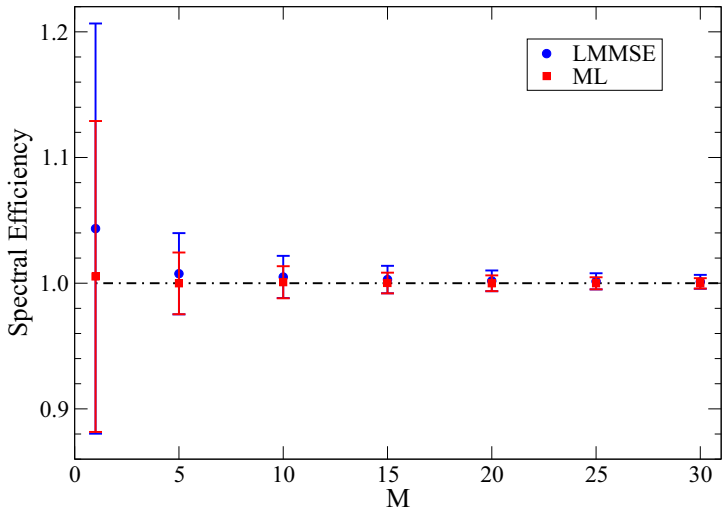


Figure 5.2: Simulation results: average spectral efficiency and relative standard deviations. System assumptions: $P_s/\sigma_d^2 = 1$, $h_s = 0$, $L = 2$, $\{|g_l h_{dl}|^2\} = \{1, 1\}$, $\alpha = 4/5$ and $K = 4M$, $N = 5M$. The ordinates are normalized with respect to the asymptotic spectral efficiency at $\alpha = 4/5$, see Figure 5.4 and Figure 5.5.

the empirical eigenvalue distribution ν_N^2 to ν^2 and show how to compute the asymptotic distribution ν^2. Then, we will prove the results stated in Section 5.2.

The results below are based on *free-probability theory* [75, 78]. As introduced in Section 2.3, free-probability theory describes the behavior of random variables defined on non-commutative algebras. In this context, free random variables are the equivalent of independent random variables in classical commutative probability, meaning that the distribution of sums and products of free non-commutative random variables can be expressed in terms of the singular distributions of the original variables.

It is a well-known result of random-matrix theory and free probability that large $N \times N$ independent Hermitian unitarily invariant random matrices can be seen as almost sure models for free non-commutative random variables (see Theorem 2.2 and, for more details, [77]). Now, assume that $\mathbf{M}_1$ and $\mathbf{M}_2$ are two such matrices and that the distributions of their eigenvalues (which are real, due to Hermiticity) tend to the measures μ_1 and μ_2, respectively, as $N \to +\infty$. Then, the eigenvalue distribution of $\mathbf{M}_1 + \mathbf{M}_2$ tends to $\mu_1 \boxplus \mu_2$, the free additive convolution of μ_1 and μ_2 defined in Section 2.3.4 (see also [77, 78]). As a consequence of Theorem 2.3 and [95, Proposition 3.5], a similar result can be stated for bi-unitarily invariant matrices and the distributions of their singular values (singular law or singular distribution).

Let $\tilde{\mu}_1$ and $\tilde{\mu}_2$ be the asymptotic singular laws[4] of two $N \times N$ independent bi-unitarily invariant random matrices $\mathbf{M}_1$ and $\mathbf{M}_2$. Denote by μ_1 and μ_2 their symmetrization, i.e. $\mu_i = \frac{1}{2}(\tilde{\mu}_i(-B) + \tilde{\mu}_i(B))$ for any Borel set B. Then, the symmetrization of the singular law of $\mathbf{M}_1 + \mathbf{M}_2$ tends[5] to $\mu_1 \boxplus \mu_2$.

Now, consider the interference matrix $\widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$. Denote by ν the asymptotic distribution of its singular values (or, equivalently, of the singular values of $\widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$), which is related to the asymptotic eigenvalue distribution by the identity

$$\int f(t)\nu^2(\,\mathrm{d}t) = \int f(t^2)\nu(\,\mathrm{d}t),$$

for any measurable function $f(\cdot)$. Since $\widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}} = \sum_{l=1}^{L} g_l h_{dl} \mathbf{C}_l$, intuition suggests that the distribution ν may be computed from the singular value distributions of the bi-unitarily invariant matrices $\{g_l h_{dl} \mathbf{C}_l\}$.

Unfortunately, the traditional free additive convolution is not helpful in the general case where $K \neq N$, since the matrices $\mathbf{C}_l$ are not square and, thus, not covered by the result above. An analogous theory for rectangular matrices has been developed by F. Benaych-Georges in [96] (see also [97]). Since these concepts are very recent and probably not widespread through the technical community, we summarize here the main points and refer the interested readers to the cited papers for a more detailed analysis of the topic. Due to the fact that the matrices $\mathbf{C}_l$ and $\mathbf{C}_l^H$ have the same non-zero singular values, we will only consider the case $\alpha \leq 1$, the extension to the case $\alpha > 1$ being straightforward.

5.3.1 Preliminaries

Let us focus on a symmetric distribution μ (recall that a generic distribution $\tilde{\mu}$ can be symmetrized by taking $\mu = \frac{1}{2}(\tilde{\mu}(-B) + \tilde{\mu}(B))$ for any Borel set B) and define the probability measure μ^2 on $\mathbb{R}_+$ to satisfy $\int f(t^2)\mu(\,\mathrm{d}t) = \int f(t)\mu^2(\,\mathrm{d}t)$ for any positive measurable function $f(\cdot)$. The moment generating series of μ^2 is defined as

$$M_{\mu^2}(z) = \sum_{n=1}^{+\infty} m_n z^n \tag{5.11}$$

where $m_n = \int t^n \mu^2(\,\mathrm{d}t)$ is the n-th moment of μ^2.

For a given $\alpha \in (0, 1]$, we denote by $H_\mu(z)$ the rectangular Cauchy transform with ratio α of the distribution μ, defined as

$$H_\mu(z) = zT \circ M_{\mu^2}(z) = z(\alpha M_{\mu^2}(z) + 1)(M_{\mu^2}(z) + 1), \tag{5.12}$$

[4] The empirical distribution of the singular values of a $N \times N$ matrix is defined as in (5.4), replacing eigenvalues with singular values. Analogously, its limit for $N \to +\infty$ is intended as in (5.5).

[5] In these cases, convergence is proven almost surely and not only in probability. See also the discussion in Section 5.2.3.

where $\circ$ denotes composition in the ring of the formal power series, and where we introduced the function

$$T(z) = (\alpha z + 1)(z + 1).$$

Let us denote by $H_\mu^{-1}(z)$ the formal inverse of the power series $H_\mu(z)$, that is $H_\mu(H_\mu^{-1}(z)) = H_\mu^{-1}(H_\mu(z)) = z$, which exists since $H_\mu(0) = 0$ and $H'_\mu(0) = 1$.

Similarly we define $U(z)$ as the formal inverse of $T(z) - 1$ (observe that $T(0) = 1$ and $T'(0) = \alpha + 1$) and write the *rectangular R-transform with ratio* α of μ as

$$C_\mu(z) = U\left(\frac{z}{H_\mu^{-1}(z)} - 1\right), \tag{5.13}$$

which is well-defined since $z^{-1}H_\mu^{-1}(z)$ is invertible with respect to multiplication ($\frac{1}{z}H_\mu^{-1}(z) = \frac{1}{H'_\mu(0)} + \dots$). The formal power series $C_\mu(z)$ identifies unambiguously the underlying probability measure μ.

To recover μ from $C_\mu(z)$, we may proceed as follows. Noting that $U(z)$ is the formal inverse of $T(z) - 1$, from (5.13) one can write:

$$\frac{z}{H_\mu^{-1}(z)} = T(C_\mu(z)).$$

Since $H_\mu(H_\mu^{-1}(z)) = H_\mu^{-1}(H_\mu(z)) = z$, the last equation implies

$$H_\mu(z) = zT[C_\mu(H_\mu(z))].$$

By comparing this with (5.12), we readily see that we can compute $M_{\mu^2}(z)$ as the formal power series satisfying the fixed point equation

$$M_{\mu^2}(z) = C_\mu[zT(M_{\mu^2}(z))]. \tag{5.14}$$

Recall that $M_{\mu^2}(z)$ univocally identifies the underlying distribution μ^2 by means of its moments, which can always be computed from (5.14). Further details are given in Section 5.5.

From the moment power series $M_{\mu^2}(z)$ we can analytically recover the actual density μ^2 as follows. Note that the moment generating series (5.11) can be written as

$$M_{\mu^2}(z) = -1 + \frac{1}{z}G_{\mu^2}\left(\frac{1}{z}\right),$$

where

$$G_{\mu^2}(z) = z^{-1} + \sum_{n=1}^{+\infty} m_n z^{-n-1}$$

is the Stieltjes transform (written as formal power series) of the distribution μ^2 according to Definition 2.2. Hence, μ^2 (and μ) can be recovered by means of the Stieltjes inversion formula (2.4), that is

$$\int_a^b \mu^2(\,\mathrm{d}t) = -\frac{1}{\pi}\lim_{y\to 0}\int_a^b \Im[G_{\mu^2}(x + jy)]\,\mathrm{d}x.$$

However, this is not necessary for our purposes since the expressions of the asymptotic spectral efficiency in this chapter are given in terms of the moment generating function.

5.3.2 Main results

We are now ready to restate [96, Theorems 3.12 and 3.13], which are at the basis of the technical results below.

First, let us recall that the singular law of a $K \times N$ ($K \leq N$) matrix $\mathbf{X}$ is $\frac{1}{K}\sum_{k=1}^{K} \delta_{\varsigma_k}$, where $\{\varsigma_k\}_{k=1}^{K}$ are the singular values of $\mathbf{X}$. Also, a random matrix is called bi-unitarily invariant if its probability measure is invariant by left- and right-multiplication by constant unitary matrices.

Theorem 5.1. *Let $K(N)$ be a sequence of integers such that $K(N)/N \to \alpha \in (0,1]$. Let $\mathbf{X}_N$ and $\mathbf{Y}_N$ be two sequences of $K \times N$ independent random matrices, one of them being bi-unitarily invariant. Assume that the symmetrizations of their singular laws converge in probability towards the probability measures μ_X and μ_Y, respectively. Then, the symmetrization of the singular law of $\mathbf{X}_N + \mathbf{Y}_N$ converges in probability to a new distribution that we denote by $\mu_X \boxplus_\alpha \mu_Y$, the* rectangular-free additive convolution with ratio α *of the measures μ_X and μ_Y.*

Using the rectangular R-transform introduced in Section 5.3.1, the resulting distribution can be computed by means of the following theorem.

Theorem 5.2. *Given $\alpha \in (0,1]$ and the two symmetric probability measures μ_X and μ_Y on the real line, the function $C_{\mu_X}(z) + C_{\mu_Y}(z)$ is the rectangular R-transform with ratio α of the symmetric probability measure $\mu_X \boxplus_\alpha \mu_Y$. Equivalently*

$$C_{\mu_X \boxplus_\alpha \mu_Y} = C_{\mu_X}(z) + C_{\mu_Y}(z),$$

which also means that the binary operator $\boxplus_\alpha$ is commutative and associative.

By associativity, Theorem 5.1 can be readily extended to L (sequences of) bi-unitarily invariant matrices $\mathbf{X}_N^{(1)}, \ldots, \mathbf{X}_N^{(L)}$ with asymptotic singular laws $\mu_1, \ldots, \mu_L$: the symmetrization of the singular law of $\mathbf{X}_N^{(1)} + \cdots + \mathbf{X}_N^{(L)}$ is

$$\nu = \mu_1 \boxplus_\alpha \cdots \boxplus_\alpha \mu_L$$

and can be evaluated following the algorithm summarized in Figure 5.3.

5.3.3 Asymptotic spectral efficiencies

The proof of the results in Section 5.2 follow from a direct application of the rectangular-free convolution. Let γ_N denote the relay contribution to the

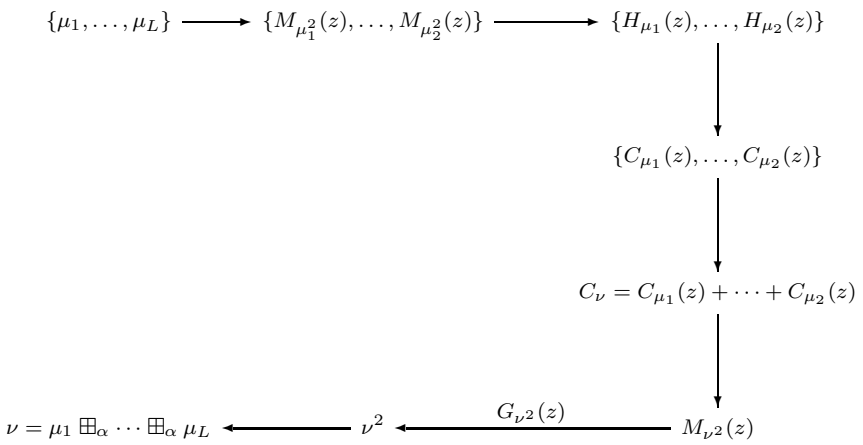


Figure 5.3: Algorithm for computing $\nu = \mu_1 \boxplus_\alpha \cdots \boxplus_\alpha \mu_L$.

SINR (5.3), namely

$$\gamma_N = \mathbf{c}_1^H \widetilde{\boldsymbol{\Psi}}^H \Big(\chi \widetilde{\boldsymbol{\Psi}} \mathbf{D}\mathbf{D}^H \widetilde{\boldsymbol{\Psi}}^H + \mathbf{I}_N\Big)^{-1} \widetilde{\boldsymbol{\Psi}} \mathbf{c}_1.$$

Recalling that $\widetilde{\mathbf{C}} = [\mathbf{c}_1\ \mathbf{D}]$, the matrix inversion lemma implies that

$$\gamma_N = \frac{\eta_N}{1-\chi\eta_N},$$

where

$$\eta_N = \mathbf{c}_1^H \widetilde{\boldsymbol{\Psi}}^H \Big(\chi \widetilde{\boldsymbol{\Psi}} \widetilde{\mathbf{C}} \widetilde{\mathbf{C}}^H \widetilde{\boldsymbol{\Psi}}^H + \mathbf{I}_N\Big)^{-1} \widetilde{\boldsymbol{\Psi}} \mathbf{c}_1.$$

Let $\mathbf{A}_N = \Big(\chi \widetilde{\mathbf{C}}^H \widetilde{\boldsymbol{\Psi}}^H \widetilde{\boldsymbol{\Psi}} \widetilde{\mathbf{C}} + \mathbf{I}_K\Big)^{-1} \widetilde{\mathbf{C}}^H \widetilde{\boldsymbol{\Psi}}^H \widetilde{\boldsymbol{\Psi}} \widetilde{\mathbf{C}}$. Then, the following result holds true:

Proposition 5.1. *Consider η_N and $\mathbf{A}_N$ as defined above. Assume that K/N converges to α as N tends to infinity. Then*

$$\lim_{N\to+\infty} \left(\eta_N - \frac{1}{K}\mathrm{tr}\{\mathbf{A}_N\}\right) = 0$$

almost surely.

Proof. This result is a direct consequence of the symmetric distribution of the columns of $\widetilde{\mathbf{C}}$. The formal proof below follows the same guidelines as that of [91, Proposition 3]. More specifically, it shows that the fourth-order moment of $\tau_N = \eta_N - \frac{1}{K}\mathrm{tr}\{\mathbf{A}_N\}$ vanishes at least as fast as N^{-2}, namely

$$\mathbb{E}\left[\left|\eta_N - \frac{1}{K}\mathrm{tr}\{\mathbf{A}_N\}\right|^4\right] = O\big(N^{-2}\big).$$

Indeed, when this condition is verified, Markov's inequality implies that

$$\Pr\Big[|\tau_N| < \epsilon\Big] \leq \frac{\mathbb{E}\Big[|\tau_N|^4\Big]}{\epsilon^4} = O\big(N^{-2}\big),$$

for any $\epsilon > 0$. Then, the series $\sum_{N=1}^{+\infty} \Pr\Big[|\tau_N| < \epsilon\Big]$ converges and, from the first Borel-Cantelli lemma (see, e.g., [59, Theorem 4.3]), $\tau_N \to 0$ almost surely as $K, N \to +\infty$ and $K/N \to \alpha$.

Let $\mathbf{B}_N = \widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H \Big(\chi \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}} \widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H + \mathbf{I}_N\Big)^{-1} \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$. Then $\mathrm{tr}\{\mathbf{A}_N\} = \mathrm{tr}\{\mathbf{B}_N\}$. Given a $K \times K$ Haar-distributed random unitary matrix $\mathbf{\Omega}_K$ independent of $\widetilde{\mathbf{C}}$ and $\widetilde{\mathbf{\Psi}}$, the distributions of $\mathbf{B}_N$ and $\mathbf{\Omega}_K^H \mathbf{B}_N \mathbf{\Omega}_K$ are the same. This is a consequence of $\{\mathbf{C}_l\}$ (and hence $\widetilde{\mathbf{C}}$) being unitarily invariant. Similarly, denoting $\mathbf{e}_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$ and by $\boldsymbol{\omega}_1 = \mathbf{\Omega}_K \mathbf{e}_1$ the first column of $\mathbf{\Omega}_K$, $\eta_N = \mathbf{e}_1^H \mathbf{B}_N \mathbf{e}_1$ and $\boldsymbol{\omega}_1^H \mathbf{B}_N \boldsymbol{\omega}_1$ have the same distribution and, in particular,

$$\mathbb{E}\left[\left|\eta_N - \frac{1}{K}\mathrm{tr}\{\mathbf{A}_N\}\right|^4\right] = \mathbb{E}\left[\left|\boldsymbol{\omega}_1^H \mathbf{B}_N \boldsymbol{\omega}_1 - \frac{1}{K}\mathrm{tr}\{\mathbf{B}_N\}\right|^4\right]. \tag{5.15}$$

Now, the vector $\boldsymbol{\omega}_1$ is uniformly distributed on the unit sphere of $\mathbb{C}^K$ [77] and, thus, may be written as $\boldsymbol{\omega}_1 = \frac{\mathbf{x}}{\|\mathbf{x}\|}$, where $\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I}_K)$ is a vector with K i.i.d. normal Gaussian entries. Let us write

$$f_{1,N} = \frac{\mathbf{x}^H \mathbf{B}_N \mathbf{x}}{\|\mathbf{x}\|^2} - \frac{\mathbf{x}^H \mathbf{B}_N \mathbf{x}}{K}$$
$$f_{2,N} = \frac{1}{K}\big(\mathbf{x}^H \mathbf{B}_N \mathbf{x} - \mathrm{tr}\{\mathbf{B}_N\}\big)$$

and

$$f_N = f_{1,N} + f_{2,N} = \boldsymbol{\omega}_1^H \mathbf{B}_N \boldsymbol{\omega}_1 - \frac{1}{K}\mathrm{tr}\{\mathbf{B}_N\}.$$

The term $f_{2,N}$ can be easily shown to vanish as $K, N \to +\infty$ and $K/N \to \alpha$. First, note that $\mathbf{x}$ is independent of $\widetilde{\mathbf{\Psi}}$ and $\widetilde{\mathbf{C}}$ and, thus, of $\mathbf{B}_N$. Also, $\mathbf{B}_N$ is uniformly bounded since

$$\left\|\Big(\chi \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}} \widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H + \mathbf{I}_N\Big)^{-1}\right\| < 1$$

and

$$\begin{aligned} \|\mathbf{B}_N\| &\leq \left\|\Big(\chi \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}} \widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H + \mathbf{I}_N\Big)^{-1}\right\| \left\|\widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}} \widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H\right\| \\ &\leq \left\|\widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{\Psi}}\right\| \left\|\widetilde{\mathbf{C}}^H \widetilde{\mathbf{C}}\right\|, \end{aligned}$$

where we have repeatedly used the inequality $\|\mathbf{MN}\| \leq \|\mathbf{M}\|\|\mathbf{N}\|$, which holds true for the spectral norm and any two square matrices $\mathbf{M}$ and $\mathbf{N}$. Finally, recalling the definition of $\widetilde{\mathbf{\Psi}}$ and $\widetilde{\mathbf{C}}$ in (5.2), one has

$$\left\|\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\right\| = \sum_{l=1}^{L}|g_l h_{dl}|^2,$$

which we can consider bounded, and

$$\left\|\widetilde{\mathbf{C}}^H\widetilde{\mathbf{C}}\right\| \leq \sum_{l=1}^{L}\left\|\mathbf{C}_l^H\mathbf{C}_l\right\| = L\max\{1,\alpha\},$$

due to the triangular inequality. Then, the assumptions of Lemma 2.1 are satisfied and $f_{2,N} \to 0$ almost surely in the large-code domain. Note that the proof of the Lemma 2.1 is based on the fact that $\mathbb{E}\big[|f_{2,N}|^N\big] = O\big(N^{-2}\big)$, which implies convergence due to the first Borel-Cantelli lemma, as seen above.

On the other hand, the term $f_{1,N}$ can be written as

$$f_{1,N} = \frac{\mathbf{x}^H\mathbf{B}_N\mathbf{x}}{K}\left(\frac{K}{\|\mathbf{x}\|^2}-1\right).$$

The above convergence result on $f_{2,N}$ implies that

$$\frac{\mathbf{x}^H\mathbf{B}_N\mathbf{x}}{K} < 2\frac{\mathrm{tr}\{\mathbf{B}_N\}}{K} \leq 2\|\mathbf{B}_N\|,$$

almost surely when the dimensions of $\mathbf{B}_N$ are large enough. The last inequality is a consequence of the fact that $\mathrm{tr}\{\mathbf{B}_N\} \leq K\|\mathbf{B}_N\|$ (recall that the spectral norm of the $K \times K$ Hermitian positive semi-definite matrix $\mathbf{B}_N$ is its maximum eigenvalue). Since $\mathbf{B}_N$ is uniformly bounded, $\mathbf{x}^H\mathbf{B}_N\mathbf{x}/K$ is bounded almost everywhere.

For the second term of $f_{1,N}$ we have

$$\mathbb{E}\left[\left(\frac{K}{\|\mathbf{x}\|^2}-1\right)^4\right] = O\big(K^{-2}\big) = O\big(N^{-2}\big)$$

when $K, N \to +\infty$ and $K/N \to \alpha$. This follows from the fact that $\|\mathbf{x}\|^2$ is χ^2 distributed with $2K$ degrees of freedom and its probability density function is $\frac{t^K}{K!}e^{-t}$. Summarizing,

$$\mathbb{E}\Big[|f_{1,N}|^4\Big] = \mathbb{E}\left[\left(\frac{\mathbf{x}^H\mathbf{B}_N\mathbf{x}}{K}\right)^4\left(\frac{K}{\|\mathbf{x}\|^2}-1\right)^4\right] = O\big(N^{-2}\big).$$

Finally, by applying the Cauchy-Schwarz inequality to the convex function $f(x) = x^4$, one has $\mathbb{E}\big[|f_N|^4\big] \leq 8\big(\mathbb{E}\big[|f_{1,N}|^4\big] + \mathbb{E}\big[|f_{2,N}|^4\big]\big)$ and, hence, $\mathbb{E}\big[|f_N|^4\big] = O\big(N^{-2}\big)$. From (5.15), this proves the lemma. □

Now, recall that ν_N^2 denotes the empirical eigenvalue distribution of $\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}$, namely

$$\nu_N^2 = \frac{1}{K}\sum_{k=1}^{K}\delta_{\lambda_k},$$

where $\{\lambda_k : k = 1,\dots,K\}$ are the K positive eigenvalues of the matrix. Then, we can write

$$\frac{1}{K}\mathrm{tr}\{\mathbf{A}_N\} = \int \frac{t}{1+\chi t}\nu_N^2(\,\mathrm{d}t).$$

Note that $\{\sqrt{\lambda_k} : k = 1,\dots,K\}$ is the set of singular values of $\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}} = \sum_{l=1}^{L} g_l h_{dl}\mathbf{C}_l$. According to the isometric coding scheme described in Section 5.1.1, each matrix $\mathbf{C}_l$ is built by extracting K columns of a $N \times N$ ($K < N$) Haar-distributed unitary random matrix. Then, each matrix $g_l h_{dl}\mathbf{C}_l$ is bi-unitarily invariant (see Appendix 5.C) and the symmetrization of its singular law is $\mu_l = \frac{1}{2}(\delta_{-|g_l h_{dl}|} + \delta_{|g_l h_{dl}|})$, independently of N. Theorem 5.1 implies that the singular law of $\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}$ converges in probability to $\nu = \mu_1 \boxplus_\alpha \cdots \boxplus_\alpha \mu_L$ and, equivalently, that $\nu_N^2 \xrightarrow{\mathrm{P}} \nu^2$ when $K = \alpha N \to +\infty$. Since $\frac{t}{1+\chi t}$ is a bounded function of $t > 0$, we can state that

$$\frac{1}{K}\mathrm{tr}\{\mathbf{A}_N\} \to \int \frac{t}{1+\chi t}\nu^2(\,\mathrm{d}t) \text{ in probability.}$$

By noting that the moment generating series $M_{\nu^2}(z) = \sum_{n=1}^{+\infty} m_n z^n$ can be equivalently written as the analytic function

$$M_{\nu^2}(z) = \int \frac{zt}{1-zt}\nu^2(\,\mathrm{d}t)$$

for $z \in \mathbb{C} \setminus \mathbb{R}_+$, Proposition 5.1 implies

$$\lim_{K=\alpha N\to+\infty} \eta_N = \eta^{\mathrm{Haar}} = -\frac{1}{\chi}M_\nu^2(-\chi).$$

The asymptotic expressions of the SINR (5.7) and of the spectral efficiency (5.8) follow from continuity of $\gamma_N = \gamma_N(\eta_N)$ and of the logarithm function.

We still need to compute the moment power series of the probability measure ν^2, according to Theorem 5.2. Since the analytic form of the moment generating function associated to $\mu_l = \frac{1}{2}(\delta_{-|g_l h_{dl}|} + \delta_{|g_l h_{dl}|})$ is very simple, namely

$$M_{\mu_l^2}(z) = \frac{z|g_l h_{dl}|^2}{1 - z|g_l h_{dl}|^2},$$

and since we are interested in the point $z = -\chi$, one can be tempted to derive $M_{\nu^2}(z)$ as a real function by simply considering $M_{\mu_l^2}(z)$ as a function

of z on the real negative axis. Unfortunately, this approach may not lead to the desired result. Indeed, by following the algorithm in Figure 5.3 in the real analytic domain, the rectangular R-transform with ratio α of μ_l can be analytically expressed as (see Appendix 5.A)

$$C_{\mu_l}(z) = \frac{\sqrt{1+4\alpha|g_l h_{dl}|^2 z}-1}{2\alpha},$$

which is defined only for $z \geq -\big(4\alpha|g_l h_{dl}|^2\big)^{-1}$. This implies that the rectangular R-transform of $\nu = \mu_1 \boxplus_\alpha \cdots \boxplus_\alpha \mu_L$, namely

$$C_\nu(z) = \sum_{l=1}^{L} C_{\mu_l}(z) = \frac{1}{2\alpha}\sum_{l=1}^{L}\Big[\sqrt{1+4\alpha|g_l h_{dl}|^2 z}-1\Big], \tag{5.16}$$

only exists for $z \geq -\big(4\alpha \max\{|g_l h_{dl}|^2\}\big)^{-1}$. When introducing this constraint, identity (5.14) may not be satisfied at $z = -\chi$.

Here, we show a possible approach for finding an algebraic solution to (5.14). In the following section, this method is proven to give a closed form expression of $M_{\nu^2}(z)$, $z \in \mathbb{R}_-$, when $L = 2$.

By definition, we know that the rectangular R-transform of ν is the sum of the rectangular R-transforms of the original distributions $\{\mu_l\}$, namely $C_\nu = \sum_{l=1}^{L} C_{\mu_l}$. We assume now that there exist L functions $M_l(z)$ such that $M_{\nu^2}(z) = \sum_{l=1}^{L} M_l(z) = \sum_{l=1}^{L} C_{\mu_l}\Big[zT\Big(\sum_{l=1}^{L} M_l(z)\Big)\Big]$, where the second equality yields from (5.14). Furthermore, they are the solutions to the following system of equations:

$$\begin{bmatrix} M_1(z) \\ \vdots \\ M_L(z) \end{bmatrix} = \begin{bmatrix} C_{\mu_1}\Big[z\Big(1+\alpha\sum_{l=1}^{L} M_l(z)\Big)\Big(1+\sum_{l=1}^{L} M_l(z)\Big)\Big] \\ \vdots \\ C_{\mu_L}\Big[z\Big(1+\alpha\sum_{l=1}^{L} M_l(z)\Big)\Big(1+\sum_{l=1}^{L} M_l(z)\Big)\Big] \end{bmatrix},$$

which is equivalent to

$$\begin{bmatrix} M_1(z)(1+\alpha M_1(z)) \\ \vdots \\ M_L(z)(1+\alpha M_L(z)) \end{bmatrix} = \\ = z\Big(1+\alpha\sum_{l=1}^{L} M_l(z)\Big)\Big(1+\sum_{l=1}^{L} M_l(z)\Big) \begin{bmatrix} |g_1 h_{d1}|^2 \\ \vdots \\ |g_L h_{dL}|^2 \end{bmatrix}. \tag{5.17}$$

The transformation $x \to x(1+\alpha x)$ has been applied to both sides of the identity and the right-hand side has been simplified knowing that $\alpha C_{\mu_l}^2(z) + C_{\mu_l}(z) = |g_l h_{dl}|^2 z$ (see (5.31) in Appendix 5.A).

Now, consider $M_{\nu^2}(z)$ as a function of z on the negative real axis $\mathbb{R}_-$. From its analytic form (5.6), it is straightforward to prove that $M_{\nu^2}(z)$ is monotonically increasing and bounded between the values -1 and zero. This fact implies that each function $M_l(z)$ is negative and lower-bounded by -1. Indeed, for each individual equation of the system (5.17), i.e.

$$M_l(z)(1+\alpha M_l(z)) = z(1+\alpha M_{\nu^2}(z))(1+M_{\nu^2}(z))|g_l h_{dl}|^2,$$

one realizes that the right-hand side is always negative (recall that we consider $\alpha \leq 1$). Then, it must be $M_l(z) \in (-1/\alpha, 0]$. Now, since $M_l(z) \leq 0$, the equality $M_{\nu^2}(z) = \sum_{l=1}^{L} M_l(z)$ implies that $-1 < M_l(z) \leq 0$. Thus, if there exists a set of functions $\{M_l(z) : l = 1, \ldots, L\}$ such that $-1 < M_l(z) \leq 0$ and $-1 < \sum_{l=1}^{L} M_l(z) \leq 0$, the resulting $M_{\nu^2}(z) = \sum_{l=1}^{L} M_l(z)$ is the desired moment generating function. This is certainly the case when $L = 2$.

Alternatively, Section 5.5 explains a low-complexity approximation method for the general case.

5.4 Special cases and examples

In this section, we will derive the exact closed-form solution of two special cases, namely (i) when there are only $L = 2$ relays in the system and (ii) when all the relay–destination channels are equal. In particular, the second case is especially simple and offers some more insight into the problem.

Note. As explained above, the asymptotic spectral efficiency of the considered relay channel depends on the limit distribution ν^2 of the eigenvalues of $\widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$ as $K = \alpha N \to +\infty$. This matrix is full rank only if $K \leq N$ ($\alpha \leq 1$). Conversely, when $K > N$ the matrix has $K - N$ null eigenvalues and N positive eigenvalues that are equal to those of $\widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}} \widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H$. Then, when $\alpha > 1$, let μ^2 be the eigenvalue distribution of the full rank matrix $\widetilde{\mathbf{\Psi}} \mathbf{W} \mathbf{W}^H \widetilde{\mathbf{\Psi}}^H$, where we have defined $\mathbf{W} = \frac{1}{\sqrt{\alpha}} \widetilde{\mathbf{C}}$ (recall that $\mathbf{C}_l \mathbf{C}_l^H = \alpha \mathbf{I}_N$ when $\alpha > 1$). Denoting by δ_0 the Dirac delta distribution, the distribution ν^2 is related to μ^2 by the following identity:

$$\nu^2(\mathrm{d}t) = \frac{\alpha - 1}{\alpha} \delta_0(\mathrm{d}t) + \frac{1}{\alpha} \mu^2\left(\frac{\mathrm{d}t}{\alpha}\right)$$

or, equivalently,

$$M_{\nu^2}(z) = \frac{1}{\alpha} M_{\mu^2}(\alpha z).$$

Since it is totally equivalent to characterize the distribution ν^2 when $\alpha \leq 1$ or the distribution μ^2 when $\alpha > 1$, all the results below apply only to the case $\alpha \leq 1$.

5.4.1 The two-relay case

For $L = 2$ relays, the system in (5.17) can be solved as follows. To simplify the notation, we make the dependence on z implicit and write $M_l = M_l(z)$, $l \in \{1, 2\}$, and $M = M_{\nu^2}(z)$. Furthermore, we denote $\gamma^{(+)} = |g_1 h_{d1}|^2 + |g_2 h_{d2}|^2$ and $\gamma^{(-)} = |g_1 h_{d1}|^2 - |g_2 h_{d2}|^2$. Then, the system of equations can be written as

$$\begin{cases} M_1(1 + \alpha M_1) = z(1 + \alpha M)(1 + M)|g_1 h_{d1}|^2 \\ M_2(1 + \alpha M_2) = z(1 + \alpha M)(1 + M)|g_2 h_{d2}|^2. \end{cases} \tag{5.18}$$

By subtracting the two equations and recalling that we must have $1+\alpha M \neq 0$, we get

$$M_1 - M_2 = z(1 + M)\gamma^{(-)}. \tag{5.19}$$

On the other hand, adding the two equations of (5.18) leads to the new identity

$$M - \frac{1}{2}\Big(M^2 + (M_1 - M_2)^2\Big) = z(1 + \alpha M)(1 + M)\gamma^{(+)}.$$

By inserting (5.19), we get the following second order equation in M:

$$\begin{aligned} \alpha\Big[1 + \big(z\gamma^{(-)}\big)^2 - 2z\gamma^{(+)}\Big]M^2 + \\ + 2\Big[1 + \alpha\big(z\gamma^{(-)}\big)^2 - z(1 + \alpha)\gamma^{(+)}\Big]M + \\ + \alpha\big(z\gamma^{(-)}\big)^2 - 2z\gamma^{(+)} = 0, \end{aligned}$$

which has the two solutions

$$\begin{aligned} M \in \big\{M^{(+)}, M^{(-)}\big\} = -\frac{\Big[1 + \alpha\big(z\gamma^{(-)}\big)^2 - z(1 + \alpha)\gamma^{(+)}\Big]}{\alpha\Big[1 + \big(z\gamma^{(-)}\big)^2 - 2z\gamma^{(+)}\Big]} \cdot \\ \cdot \left[1 \pm \sqrt{1 - \alpha\frac{\Big[\alpha\big(z\gamma^{(-)}\big)^2 - 2z\gamma^{(+)}\Big]\Big[1 + \big(z\gamma^{(-)}\big)^2 - 2z\gamma^{(+)}\Big]}{\Big[1 + \alpha\big(z\gamma^{(-)}\big)^2 - z(1 + \alpha)\gamma^{(+)}\Big]^2}}\right]. \end{aligned} \tag{5.20}$$

Basic algebra shows that the discriminant is positive, meaning that the two solutions exist and are different to one another. However, since

$$\frac{\Big[1 + \alpha\big(z\gamma^{(-)}\big)^2 - z(1 + \alpha)\gamma^{(+)}\Big]}{\alpha\Big[1 + \big(z\gamma^{(-)}\big)^2 - 2z\gamma^{(+)}\Big]} > 1$$

for $z < 0$, one has $M^{(+)} < -1$ (the second factor of the right-hand side of (5.20) is also larger than one when the plus sign is chosen) and has to

be discarded. On the contrary, it is trivial to show that $M^{(-)} \in (-1, 0]$, meaning that the moment generating function is

$$M_{\nu^2}(z) = -\frac{\left[1+\alpha\left(z\gamma^{(-)}\right)^2 - z(1+\alpha)\gamma^{(+)}\right]}{\alpha\left[1+\left(z\gamma^{(-)}\right)^2 - 2z\gamma^{(+)}\right]} \cdot \cdot \left[1 - \sqrt{1-\alpha\frac{\left[\alpha\left(z\gamma^{(-)}\right)^2 - 2z\gamma^{(+)}\right]\left[1+\left(z\gamma^{(-)}\right)^2 - 2z\gamma^{(+)}\right]}{\left[1+\alpha\left(z\gamma^{(-)}\right)^2 - z(1+\alpha)\gamma^{(+)}\right]^2}}\right]. \quad (5.21)$$

This is the function that has to be used to compute the asymptotic LMMSE SINR (5.7) (recall that $\eta^{\text{Haar}} = -M_{\nu^2}(-\chi)/\chi$) and the asymptotic ML spectral efficiency (5.9). Unfortunately, when increasing the number L of relays and, hence, the number of functions $M_l(z)$, the solution of the system in (5.17) is not as straightforward as the present one.

5.4.2 Equal channels

In this second example, we assume that all the equivalent downlink channels are equal, i.e. $|g_l h_{dl}|^2 = 1$, $l = 1, \ldots, L$.

Under this hypothesis, we can work directly in the analytic domain. Indeed, the rectangular R-transform of ν, given by (5.16), now takes the form

$$C_\nu(z) = \frac{L}{2\alpha}\left[\sqrt{1+4\alpha z} - 1\right].$$

Then, after some algebra, the function $M_{\nu^2}(z)$ that satisfies $M_{\nu^2}(z) = C_\nu(z)[z(\alpha M_{\nu^2}(z)+1)(M_{\nu^2}(z)+1)]$ is equal to

$$M_{\nu^2}(z) = \frac{L}{2\alpha}\frac{L(\alpha+1)z - 1 + \sqrt{(L(1-\alpha)z+1)^2 - 4(L-\alpha)z}}{1-L^2 z}, \quad (5.22)$$

implying

$$\eta^{\text{Haar}} = \frac{L}{2\alpha\chi}\frac{1+L(\alpha+1)\chi - \sqrt{(1-L(1-\alpha)\chi)^2 + 4(L-\alpha)\chi}}{1+L^2\chi}.$$

As a remark, it is straightforward to verify that (5.22) and (5.21) represent the same function when setting $L = 2$ and $|g_1 h_{d1}|^2 = |g_2 h_{d2}|^2 = 1$ (i.e. $\gamma^{(-)} = 0$ and $\gamma^{(+)} = 2$).

Figure 5.4 depicts the asymptotic spectral efficiency of the LMMSE filter, namely

$$I^{\text{Haar}}_{\text{LMMSE}} = \frac{\alpha}{1+\alpha}\ln\left(1 + \frac{P_s}{\sigma_d^2}|h_s|^2 + \frac{P_s}{\sigma_d^2}\frac{\eta^{\text{Haar}}}{1-\chi\eta^{\text{Haar}}}\right),$$

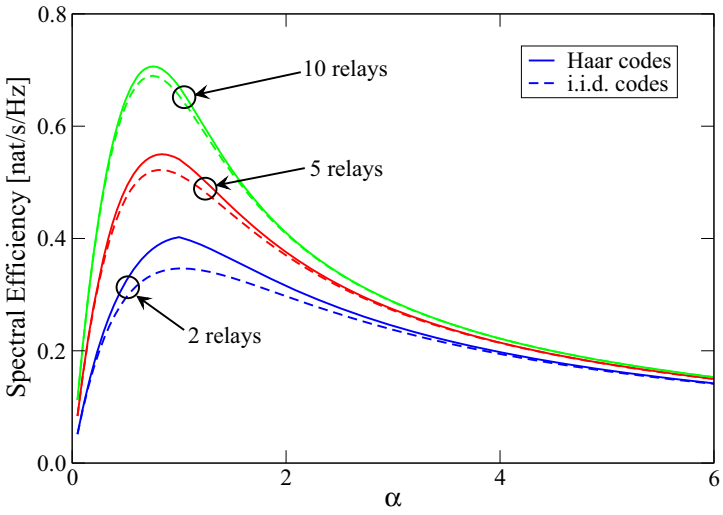


Figure 5.4: Equal-channel case: spectral efficiency as a function of α for isometric (solid line) and i.i.d. (dashed line) codes, LMMSE receiver, $|h_s|^2 = 0$, $P_s/\sigma_d^2 = 1$ and different numbers of relays.

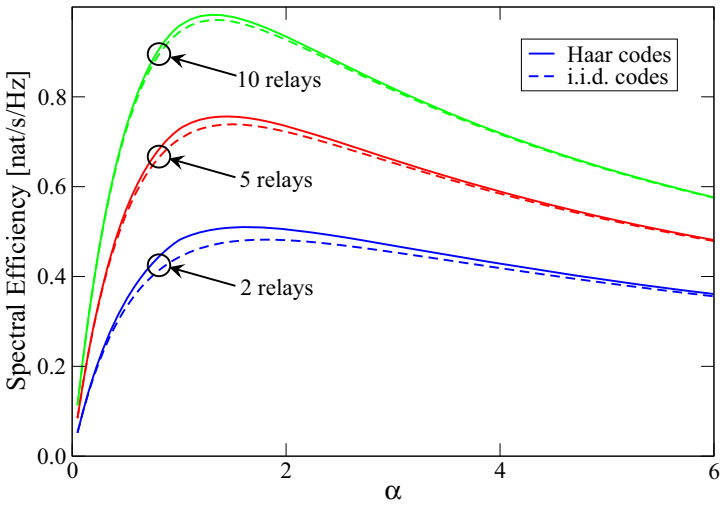


Figure 5.5: Equal-channel case: spectral efficiency as a function of α for isometric (solid line) and i.i.d. (dashed line) codes, ML receiver, $|h_s|^2 = 0$, $P_s/\sigma_d^2 = 1$ and different numbers of relays.

as a function of α for different values of the number L of relays. To calculate the asymptotic spectral efficiency $I_{\mathrm{ML}}^{\mathrm{Haar}}$ of the ML receiver (5.9), depicted in Figure 5.5, Appendix 5.B shows how to compute the antiderivative of $M_{\nu^2}(z)/z$. To focus on the effect of the codes, all the curves refer to the case $h_s = 0$.

For both receivers, the figures also show the asymptotic spectral effi-

ciencies corresponding to the use of i.i.d. codes as obtained from (5.10). By direct comparison of the two coding schemes, one notices that isometric codes introduce some benefits, as we had anticipated. However, the gain over the i.i.d. scheme — which is around 17% in spectral efficiency (comparing maxima) when considering two relays and the LMMSE filter — decays fast as the number of relays increases. Indeed, Haar codes only cancel the interference generated within each relay; interference among different relays, which becomes predominant when the number of relays increases, is not attenuated by the use of Haar coding matrices. Besides, note that the benefits are less important with the ML receiver (only around 6% with two relays), which is less sensible to colored interference.

The curves in the two graphs also highlight the fact that the coding rate α should be tuned to maximize the spectral efficiency. Unfortunately, analytically locating the maximum is unfeasible, due to the complexity of the expressions involved. The considered situations offer, nevertheless, a clear counterexample that the trivial choice $\alpha = 1$ is not always the best one: maxima can be located both at α lower than 1 (LMMSE example) and at α larger than 1 (ML example).

5.5 A moment-based approximation

In Section 5.3, the spectral efficiency obtained with isometric random codes has been shown to converge to a deterministic quantity. However, conversely to the i.i.d. case, the limit can be computed only in some particular cases as those presented in the previous section. For this reason, we propose here a low-complexity approximation of the general asymptotic spectral efficiency.

So far, no particular attention has been paid to the support of the eigenvalue distribution ν^2. However, it can also be characterized from the supports of the singular laws of the matrices $\mathbf{C}_l$. Take two symmetric compactly-supported probability measures μ_1 and μ_2. Then, another result of [96] states that the support of $\mu_1 \boxplus_\alpha \mu_2$ is contained in the union of the convex hulls of the supports of μ_1 and μ_2. This fact implies that the distribution ν^2 of the eigenvalues of $\widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$ has a compact support contained in $[0, \max_l\{|g_l h_{dl}|^2\}]$. Then, ν^2 is totally defined by its moments $\{m_i : i = 1, 2, \dots\}$. The basic idea behind the approximation below is to replace the measure ν^2 by a new measure $\bar{\nu}_n^2$, such that their first $2n-1$ moments are respectively equal, where $n \in \mathbb{N}_+$ will depend on the allowed complexity. Before presenting the algorithm with more detail, we show how to compute the moments of ν^2.

5.5.1 The moments of ν^2

The moments m_i, $i = 1, 2, \dots$, of the distribution ν^2 can be easily computed from (5.14). Indeed, from (5.16), the coefficients of the rectangular

R-transform of ν, that is of the power series

$$C_\nu(z) = \sum_{i=1}^{+\infty} c_i z^i,$$

can be written as

$$c_i = \frac{(2\alpha)^{i-1}}{i!}\left(\prod_{k=0}^{i-1}(1-2k)\right)\sum_{l=1}^{L}|g_l h_{dl}|^{2i}.$$

Then, (5.14) is equivalent to

$$M_{\nu^2}(z) = \sum_{i=1}^{+\infty} c_i z^i (\alpha M_{\nu^2}^2(z) + (\alpha+1)M_{\nu^2}(z) + 1)^i.$$

Now, it suffices to recall that $M_{\nu^2}(z) = \sum_{i=1}^{+\infty} m_i z^i$ and compare corresponding coefficients of equal powers of z. The first two moments are

$$m_1 = c_1 = \sum_{l=1}^{L}|g_l h_{dl}|^2 \text{ and} \tag{5.23a}$$

$$m_2 = (\alpha+1)c_1^2 + c_2 = (\alpha+1)\left(\sum_{l=1}^{L}|g_l h_{dl}|^2\right)^2 - \alpha\sum_{l=1}^{L}|g_l h_{dl}|^4. \tag{5.23b}$$

Any symbolic computation software can help in writing the expressions of higher order moments.

5.5.2 Approximation of the Stieltjes transform

As mentioned before, the approximation algorithm consists in replacing the real eigenvalue distribution ν^2 by a new discrete one of the form $\bar{\nu}_n^2 = \sum_{k=1}^{n}\gamma_{k,n}\delta_{\lambda_{k,n}}$, such that the first $2n-1$ moments of ν^2 and $\bar{\nu}_n^2$ coincide. For every $n \in \mathbb{N}_+$, the points $\{\lambda_{k,n}\}_{k=1}^n$ should fall in the support of ν^2 and, obviously, $\gamma_{k,n} > 0, \forall k \in \{1,\ldots,n\}$, with $\sum_{k=1}^{n}\gamma_{k,n} = 1$. The motivation behind this choice is that well-known results on the moment problem [98–100] tell us that the point-wise convergence

$$\bar{G}_n(z) \to G_{\nu^2}(z)$$

is exponential in n, where we denote by

$$\bar{G}_n(z) = \sum_{k=1}^{n}\frac{\gamma_{k,n}}{z-\lambda_{k,n}}$$

and

$$G_{\nu^2}(z) = \int \frac{1}{z-t}\nu^2(\,\mathrm{d}t)$$

the Stieltjes transforms of $\bar{\nu}_n^2$ and ν^2, respectively (see Definition 2.2). Since $M_{\nu^2}(z) = 1 - z^{-1}G_{\nu^2}(z^{-1})$, we propose the following approximations:

$$\eta^{\text{Haar}} \approx -\frac{1}{\chi}\left(1 - \sum_{k=1}^{n} \frac{\gamma_{k,n}}{1+\chi\lambda_{k,n}}\right) = \sum_{k=1}^{n} \frac{\gamma_{k,n}\lambda_{k,n}}{1+\chi\lambda_{k,n}},$$

$$I_{\text{ML}}^{\text{Haar}} \approx \frac{\alpha}{1+\alpha}\ln\left(1+\frac{P_s}{\sigma_d^2}|h_s|^2\right) + \frac{\alpha}{1+\alpha}\sum_{k=1}^{n}\gamma_{k,n}\ln(1+\chi\lambda_{k,n}).$$

Observe that, according to the last equation, the proposed approximation is equivalent to splitting the transmission over n parallel channels, the k-th one having channel gain $\lambda_{k,n}$ and carrying a fraction $\gamma_{k,n}$ of the total information.

The Gauss-Jacobi mechanical quadrature

It remains to explain how to compute the coefficients $\gamma_{k,n}$ and $\lambda_{k,n}$. The problem of approximating $G_{\nu^2}(z)$ by $\bar{G}_n(z)$ is known in the literature as the Gauss-Jacobi mechanical quadrature and makes use of the theory of orthogonal polynomials [98, 99]. We summarize hereafter its main points.

For the probability measure ν^2 with moments $m_i = \int t^i \nu^2(\,\mathrm{d}t)$, we define the scalar product

$$\langle f, g\rangle = \int f(\lambda)g(\lambda)\nu^2(\,\mathrm{d}\lambda)$$

on the space[6] $L^2(\nu^2)$. Then, the Gram-Schmidt orthogonalization procedure can be applied to the sequence $\{\lambda^n : n = 1, 2, \dots\}$ of non-negative powers of λ. As a result, we get a sequence $\{p_n(\lambda)\}_{n\geq 0}$ such as

- the polynomial $p_n(\lambda)$ has degree n and positive leading coefficient;
- the polynomials are orthonormal, i.e. $\langle p_n, p_q\rangle = 1$ if and only if $n = q$ and zero otherwise.

Equivalently, the polynomials $p_n(\lambda)$ can be computed recursively, thanks to the following result.

Proposition 5.2 (The three terms recursion relation)**.** *The family of polynomials $\{p_k\}$ satisfies the relation*

$$\lambda p_k(\lambda) = b_{k-1}p_{k-1}(\lambda) + a_k p_k(\lambda) + b_k p_{k+1}(\lambda),$$

where the coefficients a_k and b_k, defined by $a_k = \langle \lambda p_k(\lambda), p_k(\lambda)\rangle$ and $b_k = \langle \lambda p_k(\lambda), p_{k+1}(\lambda)\rangle$, are positive. The recurrence formula is initiated by $b_{-1} = 0$ and $p_0(\lambda) = 1$.

[6] Recall that, given a measure space (S, Σ, μ), the space $L^2(\mu)$ is, roughly speaking, the vector space of all functions $f(\cdot)$ such that $\int_S |f(t)|^2\mu(\,\mathrm{d}t) < +\infty$.

The coefficient $\{a_k\}$ and $\{b_k\}$ will be functions of the moments of ν^2 as, for example:

$$a_0 = m_1 \qquad b_0 = \sqrt{m_2 - m_1^2} \qquad a_1 = \frac{m_3 - 2m_1 m_2 + m_1^3}{m_2 - m_1^2} \qquad b_1 = \dots$$

Furthermore, the Stieltjes transform $\bar{G}_n(z)$ can be expressed directly in their terms, namely as the following continuous fraction:

$$\bar{G}_n(z) = \cfrac{1}{z - a_0 - \cfrac{b_0^2}{z - a_1 - \cfrac{b_1^2}{\dots - \cfrac{b_{n-2}^2}{z - a_{n-1}}}}}.$$

Finally, for a given n, the points $\{\lambda_{k,n}\}_{k=1}^n$ simply are the n zeros of $p_n(\lambda)$. The Christoffel-Darboux formula permits to compute the coefficients $\{\gamma_{k,n}\}_{k=1}^n$:

$$\gamma_{k,n} = \frac{1}{\sum_{i=0}^{n-1} |p_i(\lambda_{k,n})|^2}.$$

Figure 5.6 shows a comparison between simulation results and asymptotic approximations, for different values of n. Observe that matching three moments (i.e. $n = 2$) of the asymptotic eigenvalue distribution of the interference matrix $\widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$ suffices to obtain a good deterministic approximation of a randomly generated code of length $N = 100$, which is realistic in practical applications.

5.6 The low-power regime

As mentioned in the introduction, probably the main motivation behind the introduction of relays is the desire of achieving high data rates by means of distributed space-diversity techniques. However, relays may also be helpful in systems where the received Signal-to-Noise Ratio (SNR) is very low, because of strict energy requirements (e.g. sensor networks) or large source–destination distances (e.g. satellite communications). By improving the quality of the link, relays may reduce power consumption at the source or increase the communications range.

For this reason, in this section we describe the low-power (or wideband) regime of the considered relay channel. More specifically, we compute the minimum normalized energy per bit that allows reliable transmission, namely [89]

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{\ln 2}{\dot{I}(0)}, \tag{5.24}$$

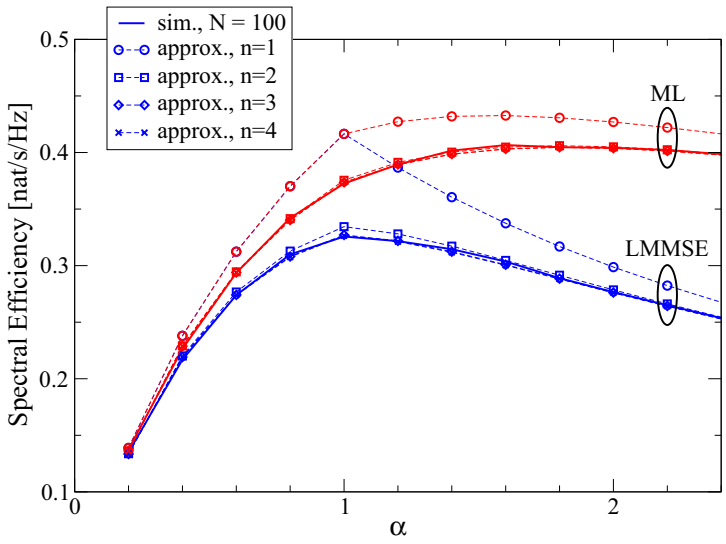


Figure 5.6: Comparison between simulation curve and approximations for $n = 1, \ldots, 4$. Systems assumptions: $P_s/\sigma_d^2 = 1$, $|h_s|^2 = 0$, $L = 2$, $\{|g_l h_{dl}|^2\} = \{0.5, 0.8\}$ and $N = 100$.

where $\dot{I}(0)$ is the first derivative of the spectral efficiency in the limit for the SNR tending to zero, expressed in nats per degree of freedom. N_0 denotes the noise power spectral density. Besides, as the energy increases from $(E_b/N_0)_{\min}$, the spectral efficiency presents a slope given by [89]

$$S_0 = -\frac{2[\dot{I}(0)]^2}{\ddot{I}(0)} \quad \text{(in bits per degree of freedom per 3 dB)}, \tag{5.25}$$

being $\ddot{I}(0)$ the limit for the SNR tending to zero of the second derivative of the spectral efficiency.

In other words, for the reference SNR $\rho = P_s/\sigma_d^2$ tending to zero, we need to compute the limit of the first- and second-order derivatives of the spectral efficiency. Let m_1 and m_2 be the first two moments of the eigenvalue distribution ν^2 of the interference matrix $\widetilde{\mathbf{C}}^H \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{C}}$, which are given by (5.23) in Section 5.5.1. Then, the first two derivatives of the spectral efficiency in $\rho = 0$ are

$$\left.\frac{\partial I_{\text{LMMSE}}^{\text{Haar}}}{\partial \rho}\right|_{\rho=0} = \frac{\alpha}{1+\alpha}\left(|h_s|^2 + m_1\right), \tag{5.26a}$$

$$\left.\frac{\partial^2 I_{\text{LMMSE}}^{\text{Haar}}}{\partial \rho^2}\right|_{\rho=0} = -\frac{\alpha}{1+\alpha}\left(|h_s|^4 + 2|h_s|^2 m_1 + 2m_2 - m_1^2\right), \tag{5.26b}$$

for the LMMSE filter and

$$\left.\frac{\partial I_{\mathrm{ML}}^{\mathrm{Haar}}}{\partial \rho}\right|_{\rho=0} = \frac{\alpha}{1+\alpha}\big(|h_s|^2 + m_1\big) \tag{5.27a}$$

$$\left.\frac{\partial^2 I_{\mathrm{ML}}^{\mathrm{Haar}}}{\partial \rho^2}\right|_{\rho=0} = -\frac{\alpha}{1+\alpha}\big(|h_s|^4 + 2|h_s|^2 m_1 + m_2\big). \tag{5.27b}$$

for the ML receiver.

These results can be proven as follows. Recall that the moment generating series of ν^2 is

$$M_{\nu^2}(z) = \sum_{i=1}^{+\infty} m_i z^i.$$

Now, since $\rho \to 0$ implies $\chi = \rho - |h_s|^2\rho^2 + o(\rho^2)$, one has

$$\eta^{\mathrm{Haar}} = -\frac{1}{\chi} M_{\nu^2}(-\chi) = m_1 - m_2\rho + o(\rho)$$

and, after some algebra,

$$\begin{aligned}
I_{\mathrm{LMMSE}}^{\mathrm{Haar}} &= \frac{\alpha}{1+\alpha}\Big[(|h_s|^2 + m_1)\rho + (m_1^2 - m_2)\rho^2 - \frac{1}{2}(|h_s|^2 + m_1)^2\rho^2\Big] + o(\rho^2),\\
\frac{\partial I_{\mathrm{ML}}^{\mathrm{Haar}}}{\partial \rho} &= \frac{\alpha}{1+\alpha}\frac{|h_s|^2}{1+\rho|h_s|^2} - \frac{\alpha}{1+\alpha}\frac{M_{\nu^2}(-\chi)}{\chi}\frac{\partial \chi}{\partial \rho}\\
&= \frac{\alpha}{1+\alpha}\Big[|h_s|^2 + m_1 - (|h_s|^4 + 2|h_s|^2 m_1 - m_2)\rho\Big] + o(\rho),
\end{aligned}$$

all for ρ small enough. The results in (5.26) and (5.27) follow from inspection once recalling the general Maclaurin expansion $I(\rho) = I(0) + \sum_{i=1}^{+\infty} \frac{1}{i!}\left(\left.\frac{\partial^i I}{\partial \rho^i}\right|_{\rho=0}\right)\rho^i$.

As before, it is worth comparing isometric and i.i.d. codes. The same technique can be applied to the i.i.d. spectral efficiencies in (5.10), leading to similar results. The first and second derivatives at $\rho = 0$ are given by:

$$\begin{aligned}
\left.\frac{\partial I^{\mathrm{iid}}}{\partial \rho}\right|_{\rho=0} &= \frac{\alpha}{1+\alpha}\left(|h_s|^2 + \sum_{l=1}^{L}|g_l h_{dl}|^2\right), \text{ in both cases,}\\
\left.\frac{\partial^2 I_{\mathrm{LMMSE}}^{\mathrm{iid}}}{\partial \rho^2}\right|_{\rho=0} &= -\frac{\alpha}{1+\alpha}\Bigg[|h_s|^4 + 2|h_s|^2\sum_{l=1}^{L}|g_l h_{dl}|^2 +\\
&\qquad + (2\alpha+1)\left(\sum_{l=1}^{L}|g_l h_{dl}|^2\right)^2\Bigg],\\
\left.\frac{\partial^2 I_{\mathrm{ML}}^{\mathrm{iid}}}{\partial \rho^2}\right|_{\rho=0} &= -\frac{\alpha}{1+\alpha}\Bigg[|h_s|^4 + 2|h_s|^2\sum_{l=1}^{L}|g_l h_{dl}|^2 +
\end{aligned}$$

$$+ (\alpha+1)\left(\sum_{l=1}^{L}|g_l h_{dl}|^2\right)^2\Bigg].$$

Inserting these results into (5.24) and (5.25), one readily obtains $(E_b/N_0)_{\min}$ and the slopes $S_0^{\text{LMMSE,iid}}$, $S_0^{\text{ML,iid}}$.

5.6.1 Slope comparison

Since the four schemes (two possible receivers and two possible codes) present the same minimum energy-per-bit, it is interesting to compare the slopes of the spectral efficiency as E_b/N_0 approaches $(E_b/N_0)_{\min}$ from above. From the expressions of the second-order derivatives, it is straightforward to verify that Haar codes outperform i.i.d. ones for both the receivers. Indeed:

$$\frac{S_0^{\text{LMMSE,Haar}}}{S_0^{\text{LMMSE,iid}}} = 1 + \frac{2\alpha\sum_{l=1}^{L}|g_l h_{dl}|^4}{|h_s|^4 + 2|h_s|^2 m_1 + 2m_2 - m_1^2}, \tag{5.28}$$

$$\frac{S_0^{\text{ML,Haar}}}{S_0^{\text{ML,iid}}} = 1 + \frac{\alpha\sum_{l=1}^{L}|g_l h_{dl}|^4}{|h_s|^4 + 2|h_s|^2 m_1 + m_2}. \tag{5.29}$$

More meaningful is the comparison between the two receivers when employing isometric codes. By replacing the expressions of the second derivatives, one obtains:

$$\frac{S_0^{\text{ML,Haar}}}{S_0^{\text{LMMSE,Haar}}} = 1 + \frac{m_2 - m_1^2}{|h_s|^4 + 2|h_s|^2 m_1 + m_2}. \tag{5.30}$$

It is straightforward to show that

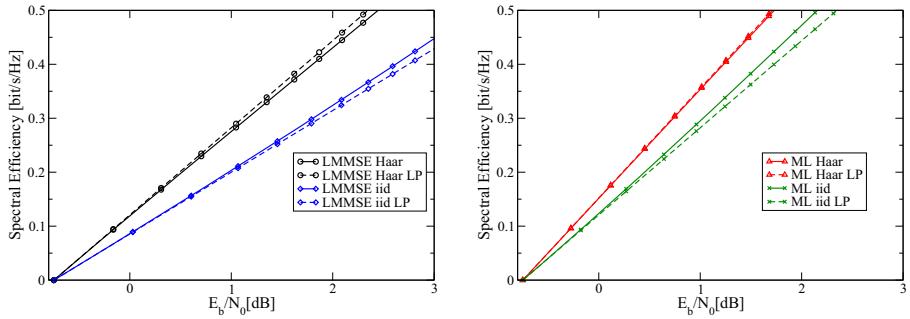
$$1 \le \frac{S_0^{\text{ML,Haar}}}{S_0^{\text{LMMSE,Haar}}} < 2.$$

Observe that the case $S_0^{\text{ML,Haar}}/S_0^{\text{LMMSE,Haar}} = 1$ arises only when the variance of the distribution ν^2 vanishes, i.e. when $m_2 - m_1^2 = 0$. This condition implies that the matrix $\widetilde{\mathbf{C}}^H\widetilde{\mathbf{\Psi}}^H\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{C}}$ is, up to a constant factor, an identity matrix. This is another evidence of the optimality of the LMMSE receiver in the white-interference signal model. Nevertheless, since

$$m_2 - m_1^2 = \alpha\left[\left(\sum_{l=1}^{L}|g_l h_{dl}|^2\right)^2 - \sum_{l=1}^{L}|g_l h_{dl}|^4\right],$$

the interference can never be whitened, except for the trivial case $\alpha = 0$.

Note that the three ratios (5.28), (5.29) and (5.30) tend to one as $|h_s|^2$ increases, meaning that all the coding/receiver schemes are equivalent in that



(a) LMMSE receiver (b) ML receiver

Figure 5.7: Spectral efficiency vs. E_b/N_0: comparison between real curves and low-power (LP) approximations for the LMMSE (a) and the ML (b) receivers.

situation. The reason is that the relay contribution becomes less important when the quality of the direct link is high.

Figure 5.7 compares simulation curves with the approximations just derived, both for the LMMSE receiver (see Figure 5.7(a)) and for the ML receiver (see Figure 5.7(b)). The gain of Haar coding over i.i.d. coding is evident. Besides, as commented in Section 5.4.2, we can notice once again that Haar signatures are especially useful with the LMMSE receiver, due to higher sensitivity of the linear receiver to colored interference.

5.7 Conclusions

This chapter has presented a randomized distributed linear-dispersion space-time block code for the relay channel which is based on isometric matrices. These codes show some gain with respect to the i.i.d.-based ones presented in Chapters 3 and 4. This advantage is due to the orthogonal structure of the coding matrices, which removes intra-relay interference. Intuition and simulation results suggest that isometric codes are more suitable in system with a low number of relays. Indeed, as we add more terminals, the interference generated within each relay becomes negligible with respect to the one due to the superposition of all relay transmissions. Furthermore, the difference between the two coding schemes is more significant when employing a LMMSE receiver, which is more sensible to colored interference than the ML receiver.

The analysis has been carried out in the asymptotic domain, i.e. when both dimensions of the coding matrices grow indefinitely but keeping constant the coding rate α. Indeed, as in the i.i.d. case, large enough random isometric codes show a deterministic behavior, independent of the specific realization of the matrices. Results have been derived by resorting to the

rectangular R-transform, a recent result of probability theory that allows to estimate the distribution of the singular values of a sum of rectangular matrices.

Appendix 5.A The rectangular R-transform of δ_a

According to the algorithm depicted in Figure 5.3, we compute here the rectangular R-transform with ratio α corresponding to the symmetrized distribution $\mu = \frac{1}{2}\delta_{-\sqrt{a}} + \delta_{\sqrt{a}}$, that is μ^2 is the distribution of the deterministic constant $a > 0$.

First, the moment generating function is given by

$$M_{\mu^2}(z) = \frac{az}{1-az},$$

which implies

$$H_\mu(z) = \frac{z}{(1-az)^2}[1-(1-\alpha)az],$$

according to (5.12). Recalling that $T(z) = (\alpha z + 1)(z+1)$ and that $z = T(U(z)) - 1$, from (5.13) we know that $C_\mu(z)$ is a solution to

$$H_\mu^{-1}(z) = \frac{z}{(1+C_\mu)(1+\alpha C_\mu)},$$

or, equivalently, to

$$\begin{aligned} z &= H\Big(\frac{z}{(1+C_\mu)(1+\alpha C_\mu)}\Big) \\ &= \frac{z}{(1+C_\mu)(1+\alpha C_\mu) - az}\,\frac{(1+C_\mu)(1+\alpha C_\mu) - (1-\alpha)az}{(1+C_\mu)(1+\alpha C_\mu) - az}. \end{aligned}$$

The last identity can be rewritten as

$$\left[(1+C_\mu)(1+\alpha C_\mu) - az\right]^2 = (1+C_\mu)(1+\alpha C_\mu) - (1-\alpha)az$$

and, after some algebra, as

$$\left[(2\alpha C_\mu + 1 + 2\alpha)^2 - (1+4\alpha az)\right]\left[(2\alpha C_\mu + 1)^2 - (1+4\alpha az)\right] = 0.$$

Since it must be $C_\mu(0) = 0$, the first term can be discarded and $C_\mu(z)$ is a solution to

$$(2\alpha C_\mu + 1)^2 = 1 + 4\alpha az,$$

and, thus, to

$$\alpha C_\mu^2 + C_\mu - az = 0. \tag{5.31}$$

When $z \in \mathbb{R}$ and $z \geq -a$, $C_\mu(z)$ is given by

$$C_\mu(z) = \frac{\sqrt{1+4\alpha az} - 1}{2\alpha}.$$

Appendix 5.B The antiderivative of $M_{\nu^2}(z)/z$

In this appendix we briefly explain how to compute a primitive of the function

$$\frac{M_{\nu^2}(x)}{x} = \frac{L}{2\alpha}\frac{L(\alpha+1)x - 1 + \sqrt{(L(1-\alpha)x-1)^2 - 4(L-\alpha)x}}{(1-L^2x)x},$$

defined for $x \leq 0$ (for continuity extension, we fix $f(0) = L$). The resulting function can be used to compute the asymptotic spectral efficiency in the case of ML receiver, equal channels and $\alpha \leq 1$ (see Section 5.4.2 and Figure 5.5).

Let us define

$$\begin{aligned}
R &= \sqrt{(L(1-\alpha)x-1)^2 - 4(L-\alpha)x};\\
A &= L(1-\alpha);\\
B &= L + \alpha L - 2\alpha;\\
E &= \frac{2B}{L^2} - \frac{A^2}{L^4} = \frac{(1+\alpha)(2L-1-\alpha)}{L^2};\\
b &= \frac{B}{L^2} - \frac{A^2}{L^4};\\
c^2 &= \frac{A^2}{L^4} - \frac{2B}{L^2} + 1 = \Big(\frac{1+\alpha-L}{L}\Big)^2 = 1 - E.
\end{aligned}$$

Then, $\frac{M_{\nu^2}(x)}{x}$ can be decomposed in terms of the form $\frac{p(x)}{q(x)R}$, where the coefficients of the polynomials $p(x)$ and $q(x)$ are functions of A, B, E, b, c^2, α and L. By applying [101, Formulas 2.261, 2.264-2 and 2.266], the antiderivative can be written as

$$\begin{aligned}
\int f(x)\,\mathrm{d}x = \frac{L}{2\alpha}\ln(R - Bx + 1) - \frac{1-\alpha}{2\alpha}\ln(AR - A^2x + B)\\
- \frac{L-1-\alpha}{2\alpha}\ln(cR + b(1-L^2x) + c^2) + C,
\end{aligned}$$

where C is a generic constant and $c = |c| = \frac{L-\alpha-1}{L}$, since we consider $\alpha \leq 1$ and we can assume $L \geq 2$ without loss of generality (the case with one single relay is trivial).

Appendix 5.C Haar-distributed unitary matrices

It is well known that a unitary matrix $\mathbf{U}$ is a square matrix with complex entries such that $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$. In other words, the inverse and the Hermitian transpose of the matrix $\mathbf{U}$ coincide.

It is straightforward to prove that the set $\mathcal{U}(N)$ of $N \times N$ unitary matrices forms a multiplicative compact group. The unitary matrix $\mathbf{U} \in \mathcal{U}(N)$ is said to be Haar distributed[7] if the probability distribution of $\mathbf{U}$ is invariant by left- or right-multiplication by constant unitary matrices. In other words, the probability distribution of $\mathbf{U}$ coincides with the Haar measure on $\mathcal{U}(N)$ (see, e.g., [102]).

Haar-distributed unitary matrices have a lot of nice properties about, for instance, the distribution of their eigenvalues or the asymptotic behavior of their trace [77, 103]. Here, however, we limit our interest to the fact that all subsets of $K \leq N$ columns (or rows) of a Haar-distributed random unitary matrix have the same distribution. Indeed, it is enough to recall that a permutation matrix $\mathbf{P}$ is also unitary and, thus, a permutation $\mathbf{UP}$ of the columns (or $\mathbf{PU}$ of the rows) of $\mathbf{U}$ does not change its distribution. As a consequence, the random isometric linear-dispersion matrices employed in this chapter are also bi-unitarily invariant.

To conclude, we explain how Haar-distributed random unitary matrices can be generated from Gaussian random matrices (see also [77, 91, 103]).

First method. Take an $N \times N$ random matrix $\mathbf{X}$ filled with i.i.d. Gaussian entries with zero mean and unitary variance. Then the unitary matrix $\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1/2}$ is Haar distributed. Indeed, for any constant unitary matrix $\mathbf{U}$, one has

$$\mathbf{UX}(\mathbf{X}^H\mathbf{X})^{-1/2} = \mathbf{UX}((\mathbf{UX})^H\mathbf{UX})^{-1/2}.$$

Since the Gaussian multi-variate distribution is completely defined by its mean and variance, $\mathbf{X}$ and $\mathbf{UX}$ are equally distributed. So are the unitary matrices $\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1/2}$ and $\mathbf{UX}(\mathbf{X}^H\mathbf{X})^{-1/2}$. Invariance by right-multiplication follows from the compactness of $\mathcal{U}(N)$.

Second method. Once again, consider the $N \times N$ Gaussian random matrix $\mathbf{X}$ defined above. The unitary matrix resulting from a QR decomposition of $\mathbf{X}$ is Haar distributed. Indeed, let

$$\mathbf{X} = \mathbf{Q}(\mathbf{X})\mathbf{R}(\mathbf{X}),$$

where $\mathbf{Q}(\mathbf{X})$ is unitary, $\mathbf{R}(\mathbf{X})$ is upper triangular and we force the diagonal elements of $\mathbf{R}(\mathbf{X})$ to be positive (note that this decomposition can be easily obtained by means of the Gram-Schmidt orthogonalization). Since $\mathbf{X}^H\mathbf{X} = \mathbf{R}(\mathbf{X})^H\mathbf{R}(\mathbf{X})$, it is obvious that $\mathbf{R}(\mathbf{X}) = \mathbf{R}(\mathbf{UX})$ for any constant unitary matrix $\mathbf{U}$. This fact implies

$$\mathbf{UX} = \mathbf{Q}(\mathbf{UX})\mathbf{R}(\mathbf{UX}) = \mathbf{Q}(\mathbf{UX})\mathbf{R}(\mathbf{X})$$

and, hence,

[7] The expression *standard unitary random matrix* is also used in the literature.

1. $\mathbf{Q}(\mathbf{X})$ and $\mathbf{Q}(\mathbf{UX})$ have the same probability distribution, since the Gaussian matrices $\mathbf{X}$ and $\mathbf{UX}$ are identically distributed;

2. $\mathbf{UQ}(\mathbf{X}) = \mathbf{Q}(\mathbf{UX})$, since we can also write $\mathbf{UX} = \mathbf{UQ}(\mathbf{X})\mathbf{R}(\mathbf{X})$.

The two results together prove that $\mathbf{Q}(\mathbf{X})$ and $\mathbf{UQ}(\mathbf{X})$ have the same probability distribution or, equivalently, that $\mathbf{Q}(\mathbf{X})$ is a Haar-distributed random unitary matrix.

Third method. Consider a $N \times K$ matrix $\mathbf{X}$ with Gaussian zero-mean unitary-variance i.i.d. entries. The unitary matrices resulting from the singular value decomposition (SVD) of $\mathbf{X}$ are Haar distributed. Indeed, let $\mathbf{U}$ and $\mathbf{V}$ be two unitary matrices of dimensions $N \times N$ and $K \times K$, respectively, and let $\mathbf{S}$ be a $N \times K$ diagonal matrix with positive entries such that

$$\mathbf{X} = \mathbf{USV},$$

that is $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{S}$ are the SVD factors of $\mathbf{X}$. Note that the three matrices can be made independent of one another. Moreover, let $\mathbf{Q}$ be a $N \times N$ Haar-distributed random unitary matrix independent of $\mathbf{U}$ and $\mathbf{S}$. Then, $\mathbf{QU}$ is Haar distributed. Since $\mathbf{XX}^H = \mathbf{USS}^H\mathbf{U}^H$ and $\mathbf{QXX}^H\mathbf{Q}^H$ are identically distributed, we can assume $\mathbf{U}$ to be Haar distributed without loss of generality. Similar reasoning can be followed for $\mathbf{V}$.

Chapter 6

Conclusions

This dissertation has proposed the use of randomized distributed LD-STBC for the relay channel. The underlying motivation was the need for a diversity-achieving distributed space-time coding scheme which is suitable for a dynamic network with a (possibly) large and variable number of transmitters. Randomized distributed LD-STBC turned out to be a good candidate, since it simply requires assigning each relay a specific linear-dispersion matrix. The matrices are independently generated and do not depend on the total number of transmitters. The technical chapters of this thesis have hence been focused on the analysis of the code, investigating whether it introduces diversity. When possible, its performance have been compared to those of existing schemes.

6.1 Summary of the presented results

Probably the most intuitive way to generate random matrices is to fill them with i.i.d. random entries. In Chapter 3 the randomized i.i.d. LD-STBC has been employed in an AF multiple-relay channel. The resulting spectral efficiency has been shown to converge almost surely to a deterministic quantity when the dimensions of the code grow large while keeping constant the coding rate. The limiting spectral efficiency turns out to be an excellent estimate of the spectral efficiency obtained by practical large-enough codes (the approximation is already accurate for code lengths around 30 or 40 symbols). As a consequence, all the analysis carried out in Chapters 3 and 4 has been based on the deterministic asymptotic spectral efficiency to avoid dealing with code randomness.

In terms of spectral efficiency, a direct comparison of the AF multiple-relay channel with the direct source–destination link has raised an intuitive result: relaying is superior to point-to-point transmission only when the source–destination channel is too weak to support the communication alone. Sufficient conditions for the superiority of relaying have been derived for both

the ML receiver and the LMMSE receiver. In the ML-receiver case, the condition turns out to be also necessary when one single relay is considered.

Note that these results on spectral efficiency are not enough to investigate the diversity offered by the system. For this purpose, Chapter 4 deals with outage probability, i.e. the probability that the spectral efficiency is lower than a target rate due to channel fading. The diversity order is then given by the speed at which the outage probability vanishes as we increase the transmitted power.

To start with a simple case, in Chapter 4 we have first replaced AF relays by DF ones. In this way, no noise is forwarded by the relays to the destination and the outage analysis is less involved. The results are quite intuitive: the system undergoes an outage only when the direct link and all the source–relay–destination paths are deeply faded. In other words, one single good link is enough to sustain communications. The only exception arises when the LMMSE receiver is employed together with high coding rates. In this case, the relays excessively compress the information and the linear receiver is not capable of exploiting their contribution. Note, however, that a good choice of the coding rate is important also in the ML-receiver case, since it has a strong influence on the outage gain. Figures 4.1 and 4.2 (the latter refers to the LMMSE receiver) show that the minimum outage gain may not be achieved by setting the coding rate to the trivial value one. Working with the best coding rate results in significant improvements in terms of outage probability, as depicted in Figure 4.5.

Conversely, AF relays offer full diversity only when the coding rate is low enough. Then, by increasing the coding rate, the diversity order decays as depicted in Figure 4.7. The fact that AF relays employing randomized LD-STBC do not achieve full diversity is a consequence of the non-orthogonality of the codes. At the receiver, the colored structure of the equivalent interference is an effect of the correlation among relay contributions and reduces the number of equivalent parallel channels. Obviously, the correlation increases as we reduce the length of the code.

Summarizing, randomized LD-STBC is more suitable for DF relaying than for AF relaying, since full diversity order can be achieved while working with higher coding rates, which better exploits the system degrees of freedom. Nevertheless, by comparing systems with one single relay (see Section 4.2.3), intuition suggests that AF relays outperform DF ones when the coding rate is low enough. In this case, indeed, the equivalent noise is almost white and the receiver can always extract some information from all the noisy relay messages. Conversely, when the coding rate increases, the AF equivalent noise becomes more colored, making the relay contribution useless. It is then preferable to use DF relays: each message is relayed by a lower number of terminals (only those which were able to decode the source message), but no noise is forwarded.

The results of Chapter 3 and, especially, Chapter 4 have revealed the

potentialities of randomized LD-STBC by analyzing linear-dispersion matrices filled with i.i.d. random entries. These matrices are very simple to build, but they do not allow designing any smart transformation of the information vector. In Chapter 5, we have hence proposed to use isometric linear-dispersion matrices with orthogonal columns (or rows, according to the coding rate), which intuitively reduce interference. Note that this choice does not spoil the flexibility of the code: matrices assigned to different relays are still independent of each other and the resulting code can be employed in dynamic networks as the previous one. Unfortunately, matrix independence is also the weak point of isometric codes: the orthogonal structure of the linear-dispersion matrices reduces interference inside each relay but does not affect the one among different relays. As a consequence, the gain of isometric codes over i.i.d. ones decreases as we add relays to the system.

6.2 Future work

A number of research lines are left open by this dissertation. First, to simplify the analysis, all the results refer to orthogonal relaying, i.e. the source remains silent during the relaying phase. Nevertheless, non-orthogonal relaying protocols are probably more interesting in practice, since the source continuously transmits new symbols without being aware of the relays and their functioning. As mentioned in Section 3.5, some preliminary work has been done in [55], where non-orthogonal AF relaying has been shown to outperform orthogonal AF relaying in terms of spectral efficiency. The outage characterization of both AF and DF non-orthogonal relaying is still missing.

The results about isometric codes presented in Chapter 5 also leave some interesting questions unsolved. Indeed, very recent mathematical tools have been used to give an asymptotic characterization of the codes. The resulting expressions are more involved than those for the i.i.d. codes. For this reason, Chapter 5 has only dealt with the spectral efficiency of DF relays (recall that DF relays do not forward noise from the source–relay channel). Furthermore, exact expressions have been derived only for some special cases. On the one hand, it would hence be interesting to extend those results to the generic DF relay channel and, from there, study the outage behavior of the system.

On the other hand, isometric codes are probably more suitable for AF relays (as compared to DF relays). This intuition comes from the fact that, apparently, the main problem with AF relays is the colored noise they generate. Let us consider the covariance matrix of the equivalent received noise (see Section 3.1.1), namely

$$\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \begin{bmatrix} \sigma_d^2 \mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix};$$
$$\text{with } \mathbf{R} = \sigma_u^2 \widetilde{\mathbf{C}} \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{\Psi}}^H \widetilde{\mathbf{C}}^H + \sigma_d^2 \mathbf{I}_N.$$

Now, observe that the matrix $\mathbf{R}$ can be rewritten as follows:

$$\mathbf{R} = \sigma_u^2 \sum_{l=1}^{L} |g_l h_{dl}|^2 \mathbf{C}_l \mathbf{C}_l^H + \sigma_d^2 \mathbf{I}_N.$$

Then, with a coding rate $\alpha \geq 1$, isometric linear-dispersion matrices imply that $\mathbf{C}_l \mathbf{C}_l^H = \alpha \mathbf{I}_N$ and the resulting noise is white:

$$\mathbf{R} = \left(\alpha \sigma_u^2 \sum_{l=1}^{L} |g_l h_{dl}|^2 + \sigma_d^2 \right) \mathbf{I}_N.$$

When $\alpha < 1$, the matrices $\{\mathbf{C}_l : l = 1, \ldots, L\}$ are tall with orthogonal columns and the equivalent noise is not white anymore. Still, we can reasonably believe that the noise structure is less colored as compared to the one generated by i.i.d. codes.

To conclude, let us mention that the main limitation of the presented scheme is probably the information assumed at the receiver. Aiming at a completely dynamic and distributed system, indeed, it would be interesting for each relay to generate its specific linear-dispersion matrix without exchanging any set-up information with the destination. A possible solution to this problem could be the estimation, at the receiver side, of an equivalent channel that absorbs the linear dispersion into it.

Appendix A

Notation

In general, uppercase boldface letters ($\mathbf{A}$) denote matrices, lowercase boldface letters ($\mathbf{a}$) denote (column) vectors and italics (a) denote scalars.

$\mathbf{A}^T, \mathbf{A}^*, \mathbf{A}^H$	Transpose, complex conjugate and transpose conjugate (Hermitian transpose) of a matrix $\mathbf{A}$, respectively.
$\mathbf{A}^{-1}$	Inverse of $\mathbf{A}$.
$\mathbf{A}^{1/2}$	Positive definite Hermitian square root of $\mathbf{A}$, i.e. $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$.
$\det(\mathbf{A})$	The determinant of the matrix $\mathbf{A}$.
$\mathrm{tr}\{\mathbf{A}\}$	The trace of the matrix $\mathbf{A}$.
$\|\mathbf{A}\|$	The spectral norm of the matrix $\mathbf{A}$, i.e. the largest singular value of $\mathbf{A}$.
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product between $\mathbf{A}$ and $\mathbf{B}$. If $\mathbf{A} = [a_{m,n} : m = 1, \ldots, M; n = 1, \ldots, N]$, $\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{B} & \cdots & a_{1,N}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{M,1}\mathbf{B} & \cdots & a_{M,N}\mathbf{B} \end{bmatrix}$.
$\mathbf{I}_N$	The $N \times N$ identity matrix.
$\mathbf{0}$	A vector or matrix with all-zero entries.
$\mathrm{diag}\{a_1, \ldots, a_N\}$	A diagonal $N \times N$ matrix with entries a_i.
$\mathbb{R}^N, \mathbb{C}^N$	The set of vectors with N real or complex valued elements, respectively.
$\mathbb{R}^{M \times N}, \mathbb{C}^{M \times N}$	The set of $M \times N$ matrices with real or complex valued elements, respectively.
$\mathbb{R}[x], \mathbb{C}[x]$	The set of polynomials in x with real and complex valued coefficients, respectively.
$\mathbb{R}[[x]], \mathbb{C}[[x]]$	The set of formal power series in x with real and complex valued coefficients, respectively.
j	The immaginary unit, i.e. $j = \sqrt{-1}$.
$\Re(z)$	The real part of the complex number z.

$\Im(z)$	The immaginary part of the complex number z.
$[a]^+$	Maximum between the real quantity a and zero.
$\Pr[\mathcal{A}]$	Probability of the event $\mathcal{A}$.
$\mathbb{E}[\cdot]$	Expected value.
$\mathcal{CN}(\cdot,\cdot)$	The (multivariate) complex Gaussian distribution.
P_{out}	Outage probability.
$I(X,Y)$	Mutual information between signals X and Y.
$\lvert\mathcal{L}\rvert$	The cardinality of the set $\mathcal{L}$.
$\bar{\mathcal{L}}$	Complement of the set $\mathcal{L}$.
$\mathbb{1}\{\mathcal{A}\}$	The indicator function of the event $\mathcal{A}$, i.e. $\mathbb{1}\{\mathcal{A}\} = \begin{cases} 1 & \text{if } \mathcal{A} \text{ is true,} \\ 0 & \text{if } \mathcal{A} \text{ is false.} \end{cases}$
$\mathbf{1}$	Algebra unity.
$\delta_a(\cdot), \delta_a$	Dirac delta centered in a.
$\delta_{i,k}$	Kronecker delta.
$\mathrm{d}\mathbf{x}_i^j$	For any $i \leq j$, $\mathrm{d}\mathbf{x}_i^j = \mathrm{d}x_i \dots \mathrm{d}x_j$.
$\mathrm{d}\mathbf{x}_\mathcal{L}$	For a given set of indexes $\mathcal{L}$, $\mathrm{d}\mathbf{x}_\mathcal{L} = \prod_{l\in\mathcal{L}} \mathrm{d}x_l$.
$f(z) \doteq z^d$	Exponential equality, i.e. $\lim_{z\to+\infty} \frac{\ln f(z)}{lnz} = d$.
ln	The natural logarithm.
$f \boxplus g$	Additive free convolution of the distributions f and g.
$f \boxplus_\alpha g$	Rectangular additive free convolution with ratio α of the distributions f and g.
$f \boxtimes g$	Multiplicative free convolution of the distributions f and g.
$\xrightarrow{\text{a.s.}}$	Almost sure convergence.
$\xrightarrow{\text{D}}$	Convergence in distribution.
$\xrightarrow{\text{m.s.}}$	Convergence in mean square.
$\xrightarrow{\text{P}}$	Convergence in probability.
$O(\cdot), o(\cdot)$	Landau symbols for order of convergence.
$\max, \min$	Maximum and minimum, respectively.
sup	Supremum (lowest upper-bound).

Appendix B

Acronyms

AF	Amplify and Forward.
AWGN	Additive White Gaussian Noise.
CDMA	Code Division Multiple Access.
CF	Compress and Forward.
CSI	Channel State Information.
DF	Decode and Forward.
DMT	Diversity–Multiplexing Tradeoff.
DS/CDMA	Direct Sequence Code Division Multiple Access.
GHz	Giga Hertz.
GSM	Global System for Mobile Communications (originally *Groupe Spécial Mobile*).
Hz	Hertz.
i.i.d.	independent and identically distributed.
IP	Internet Protocol.
IS-95	Interim Standard 95.
LAN	Local Area Network.
LD-STBC	Linear-Dispersion Space-Time Block Code (or Coding).
LDCT	Lebesgue's Dominated Convergence Theorem.
LMMSE	Linear Minimum Mean Square Error.
MAC	Media Access Control.
MIMO	Multiple Input Multiple Output.
ML	Maximum Likelihood.
pdf	probability density function.
RMT	Random Matrix Theory.
s	second.
SINR	Signal-to-Interference-plus-Noise Ratio.
SNR	Signal-to-Noise Ratio.
STC	Space-Time Code (or Coding).
STBC	Space-Time Block Code (or Coding).

SVD	Singular Value Decomposition.
TDMA	Time Division Multiple Access.
UMTS	Universal Mobile Telecommunications System.

Bibliography

[1] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.

[2] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. New York, NY, USA: Cambridge University Press, 2005.

[3] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. New York, NY, USA: Kluwer Academic / Plenum Publishers, 1999.

[4] A. Kaye and D. George, "Transmission of multiplexed PAM signals over multiple channel and diversity systems," *IEEE Trans. Commun. [legacy, pre-1988]*, vol. 18, no. 5, pp. 520–526, Oct. 1970.

[5] W. van Etten, "An optimum linear receiver for multiple channel digital transmission systems," *IEEE Trans. Commun. [legacy, pre-1988]*, vol. 23, no. 8, pp. 828–834, Aug. 1975.

[6] ——, "Maximum likelihood receiver for multiple channel transmission systems," *IEEE Trans. Commun. [legacy, pre-1988]*, vol. 24, no. 2, pp. 276–283, Feb. 1976.

[7] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[8] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[9] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[10] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1999, invited paper.

[11] F. H. P. Fitzek and M. D. Katz, Eds., *Cooperation in Wireless Networks: Principles and Applications – Real Egoistic Behavior is to Cooperate!* Dordrecht, The Netherlands: Springer, 2006.

[12] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, Department of Statistics, University of California, Berkeley, CA, USA, Jun. 1968.

[13] ——, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

[14] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[15] J. N. Laneman, "Cooperative diversity in wireless networks: Algorithms and architectures," Ph.D. dissertation, Massachusetts Institute of Technology, Sep. 2002.

[16] M. Dohler, "Virtual antenna arrays," Ph.D. dissertation, King's College London, University of London, Strand, London, UK, Nov. 2003.

[17] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity–Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[18] ——, "User cooperation diversity–Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[19] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[20] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[21] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.

[22] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[23] M. Yuksel and E. Erkip, "Diversity-multiplexing tradeoff in cooperative wireless systems," in *Proc. Conference in Information Sciences and Systems (CISS)*, Princeton, NJ, USA, Mar. 22–24, 2006.

[24] P. Elia, K. Vinodh, M. Anand, and P. Vijay Kumar, "D-MG tradeoff and optimal codes for a class of AF and DF cooperative communication protocols," 2006. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:cs/0611156

[25] S. Berger and A. Wittneben, "Cooperative distributed multiuser MMSE relaying in wireless ad-hoc networks," in *Proc. Asilomar Conference on Signals, Systems, and Computers 2005*, Pacific Grove, CA, USA, Nov. 2005.

[26] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[27] P. Mitran, H. Ochiai, and V. Tarokh, "Space-time diversity enhancements using collaborative communications," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2041–2057, Jun. 2005.

[28] O. Simeone and U. Spagnolini, "Capacity region of wireless ad hoc networks using opportunistic collaborative communications," in *Proc. IEEE ICC 2006*, Istanbul, Turkey, Jun. 11–15, 2006.

[29] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[30] V. I. Morgenshtern and H. Bölcskei, "Random matrix analysis of large relay networks," in *Proc. Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, Sep. 27–29, 2006.

[31] I. Hammerström, M. Kuhn, and A. Wittneben, "Cooperative diversity by relay phase rotations in block fading environments," in *Proc. IEEE SPAWC*, Lisbon, Portugal, Jul. 11–14, 2004.

[32] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[33] W. Hachem, P. Bianchi, and Ph. Ciblat, "Outage probability-based power and time optimization for relay networks," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 764–782, Feb. 2009.

[34] R. U. Nabar, H. Bölcskei, and F. W. Kneubühler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[35] A. Reznik, S. R. Kulkarni, and S. Verdú, "Degraded Gaussian multirelay channel: Capacity and optimal power allocation," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3037–3046, Dec. 2004.

[36] B. Sirkeci-Mergen and A. Scaglione, "Randomized space-time coding for distributed cooperative communications," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5003–5017, Oct. 2007.

[37] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[38] Y. Cao and B. Vojcic, "MMSE multiuser detection for cooperative diversity CDMA systems," in *Proc. IEEE WCNC 2004*, Atlanta, GA, USA, Mar.21–25 2004.

[39] I. Hammerström, M. Kuhn, and A. Wittneben, "Channel adaptive scheduling for cooperative relay networks," in *Proc. IEEE Vehicular Technology Conference 2004-Fall*, Los Angeles, CA, USA, Sep.26–29 2004.

[40] A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.

[41] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[42] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. INFOCOM*, New York, NY, USA, Jun. 23–27, 2002.

[43] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006.

[44] R. U. Nabar and H. Bölcskei, "Capacity scaling laws in asynchronous relay networks," in *Proc. Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, Oct. 2004.

[45] B. Rankov and A. Wittneben, "On the capacity of relay-assisted wireless MIMO channels," in *Proc. IEEE SPAWC*, Lisbon, Portugal, Jul. 11–14, 2004.

[46] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[47] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, UK: Cambridge University Press, 2003.

[48] H. Jafarkhani, *Space-Time Coding: Theory and Practice*. New York, NY, USA: Cambridge University Press, 2005.

[49] Y. Chang and Y. Hua, "Application of space-time linear block codes to parallel wireless relays in mobile ad hoc networks," in *Proc. Asilomar Conference on Signals, Systems, and Computers 2003*, Pacific Grove, CA, USA, Nov. 2003.

[50] P. A. Anghel, G. Leus, and M. Kaveh, "Distributed space-time cooperative systems with regenerative relays," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3130–3141, Nov. 2006.

[51] S. Barbarossa and G. Scutari, "Distributed space-time coding for multihop networks," in *Proc. IEEE ICC 2004*, Paris, France, Jun.20–24 2004.

[52] ——, "Distributed space-time coding strategies for wideband multihop networks: Regenerative vs. nonregenerative relays," in *Proc. IEEE ICASSP 2004*, Montreal, Quebec, Canada, May17–21 2004.

[53] J. Mietzner, R. Thobaben, and P. A. Hoeher, "Analysis of the expected error performance of cooperative wireless networks employing distributed space-time codes," in *Proc. IEEE Globecom 2004*, Dallas, Texas, USA, Nov. 29–Dec. 3 2004.

[54] T. Wang, Y. Yao, and G. B. Giannakis, "Non-coherent distributed space-time processing for multiuser cooperative transmissions," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3339–3343, Dec. 2006.

[55] D. Gregoratti and X. Mestre, "Random DS/CDMA for the amplify and forward relay channel," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 1017–1027, Feb. 2009.

[56] ——, "Large-SNR outage analysis for the df relay channel with randomized space-time block coding," IEEE Trans. Wireless Commun., submitted.

[57] S. Verdú, *Multiuser Detection*. New York, NY, USA: Cambridge University Press, 1998.

[58] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill, 2001.

[59] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley and Sons, 1995.

[60] *Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System IS-95A*, Telecommunications Industry Association, TIA/EIA Std., 1995.

[61] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, 2004.

[62] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices, a review," *Statistica Sinica*, vol. 9, pp. 611–677, 1999.

[63] U. Haagerup and S. Thorbjønsen, "Random matrices with complex Gaussian entries," Centre for Mathematical Physics and Stochastics (Department of Mathematical Sciences), University of Aarhus, Tech. Rep. MPS-RR 1998-4, 1998. [Online]. Available: http://bib.mathematics.dk/unzip.php?filename=DMF-1998-05-002-v1.pdf.gz

[64] V. L. Girko, *Theory of Random Determinants*. The Netherlands: Kluwer Academic Publishers, 1990.

[65] E. P. Wigner, "On the distributions of the roots of certain symmetric matrices," *Annals of Mathematics*, vol. 67, pp. 325–327, 1958.

[66] T. W. Arnold, "On the asymptotic distribution of the eigenvalues of random matrices," *Journal of Mathematical Analysis and Applications*, vol. 20, pp. 262–268, 1967.

[67] V. A. Marčenko and L. A. Pastur, "The distribution of the eigenvalues in certain sets of random matrices," *Math. Sbornik*, vol. 72, pp. 507–536, 1967.

[68] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[69] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 55, pp. 331–339, Aug. 1995.

[70] P. Billingsley, *Convergence of Probability Measures*, 2nd ed. John Wiley and Sons, 1999.

[71] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2059–2078, Sep. 2000.

[72] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[73] D. N. C. Tse and V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[74] W. Fulton, "Eigenvalues of sums of Hermitian matrices (after A. Klyachko)," *Séminaire Bourbaki*, vol. 40, pp. 255–269, exposé 845, 1998.

[75] R. Speicher, "Free probability theory and random matrices," Jul. 2001, lectures at the summer school on Asymptotic Combinatorics with Applications to Mathematical Physics in St. Petersburg.

[76] P. Biane, "Free probability for probabilists," *Quantum Probability Communications*, vol. XI, pp. 55–71, 1998.

[77] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*. American Mathematical Society, 2000.

[78] D. V. Voiculescu, K. Dykema, and A. Nica, *Free Random Variables*. American Mathematical Society, 1992.

[79] K. L. Chung, *A Course in Probability Theory*, 3rd ed. Academic Press, 2001.

[80] D. V. Voiculescu, "Limit laws for random matrices and free products," *Inventiones Mathematicae*, vol. 104, pp. 201–220, 1991.

[81] K. Dykema, "On certain free product factors via an extended matrix model," *Journal of Functional Analysis*, vol. 112, pp. 31–60, 1993.

[82] R. Speicher, "Combinatorics of free probability theory," 1999, lectures at IHP (Paris).

[83] R. U. Nabar, F. W. Kneubühler, and H. Bölcskei, "Performance limits of amplify-and-forward based fading relay channels," in *Proc. IEEE ICASSP 2004*, Montreal, Canada, May 17–21, 2004.

[84] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[85] O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161–163, May 2000.

[86] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 858–871, May 1997.

[87] J. Zhang, E. K. P. Chong, and D. N. C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1128–1144, Mar. 2001.

[88] D. Gregoratti and X. Mestre, "On the low source power regime of the DS/CDMA relay channel," in *Proc. IEEE SPAWC 2008*, Recife, Brazil, Jul. 6–9 2008.

[89] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[90] B. Walke, *Mobile Radio Networks: networking, protocols and traffic performance*. John Wiley and Sons, 2001.

[91] M. Debbah, W. Hachem, P. Loubaton, and M. de Courville, "MMSE analysis of certain large isometric random precoded systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1293–1311, May 2003.

[92] M. Debbah, P. Loubaton, and M. de Courville, "Asymptotic performance of successive interference cancellation in the context of linear precoded OFDM systems," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1444–1448, Sep. 2004.

[93] J.-M. Chaufray, W. Hachem, and P. Loubaton, "Asymptotic analysis of optimum and suboptimum CDMA downlink MMSE receivers," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2620–2638, Nov. 2004.

[94] J. B. Conway, *Functions of One Complex Variable I*, 2nd ed., ser. Graduate Texts in Mathematics. Springer, 1978, vol. 11.

[95] U. Haagerup and F. Larsen, "Brown's distribution measure of R-diagonal elements in finite von Neumann algebras," *Journal of Functional Analysis*, no. 176, pp. 331–367, 2000.

[96] F. Benaych-Georges, "Rectangular random matrices, related convolution," *Probability Theory and Related Fields*, vol. 144, no. 3-4, pp. 471–515, Jul. 2009.

[97] S. Belinschi, F. Benaych-Georges, and A. Guionnet, "Regularization by free additive convolution, square and rectangular cases," *Complex Analysis and Theory*, vol. 3, no. 3, pp. 611–660, 2009.

[98] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, ser. University Mathematical Monographs. Hafner Pub. Co., 1965.

[99] G. Szegö, *Orthogonal Polynomials*, 4th ed. American Mathematical Society, 1975.

[100] P. Loubaton and W. Hachem, "Asymptotic analysis of reduced rank Wiener filters," in *Proc. IEEE Information Theory Workshop (ITW '03)*, Paris, France, spring 2003.

[101] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed. San Diego, CA, USA: Academic Press, 2000.

[102] R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* John Wiley and Sons, 2007.

[103] D. Petz and J. Réffy, "On asymptotics of large haar distributed unitary matrices," *Periodica of Mathematica Hungarica*, vol. 49, no. 1, pp. 103–117, 2004.

[104] D. Gregoratti and X. Mestre, "Asymptotic spectral efficiency analysis of the DS/CDMA amplify and forward relay channel," in *Proc. Asilomar Conference on Signals, Systems, and Computers 2007*, Pacific Grove, CA, USA, Nov. 4–7 2007.

[105] ——, "About asymptotic spectral efficiency in the DS/CDMA amplify and forward relay channel," in *Proc. IEEE ISWPC 2008*, Santorini, Greece, May 7–9 2008.

[106] ——, "The DS/CDMA amplify and forward relay channel: Asymptotic spectral efficiency," in *Proc. ICT-MobileSummit 2008*, Stockholm, Sweden, Jun. 10–12 2008.

[107] ——, "The single relay channel: Does randomized coding increase diversity?" in *Proc. ICT-MobileSummit 2009*, Santander, Spain, Jun. 10–12 2009.

[108] ——, "Diversity analysis of a randomized distributed space-time coding in an amplify and forward relay channel," in *Proc. IEEE ICC 2009*, Dresden, Germany, Jun. 14–18 2009.

[109] ——, "Decode and forward relays: Full diversity with randomized distributed space-time coding," in *Proc. IEEE ISIT 2009*, Seoul, Korea, Jun. 28–Jul. 3 2009.

[110] ——, "Diversity order for the amplify-and-forward multiple-relay channel with randomized distributed space-time coding," in *Proc. EUSIPCO 2009*, Glasgow, Scotland, Aug. 24–28 2009.

[111] D. Gregoratti, W. Hachem, and X. Mestre, "Orthogonal matrix precoding for relay networks," in *Proc. IEEE ISWPC 2010*, Modena, Italy, May 5–7 2010.

[112] ——, "Randomized isometric linear-dispersion space-time block coding for the DF relay channel," IEEE Trans. Signal Process., submitted.

Index