MIMO TRANSMISSION FOR 4G WIRELESS COMMUNICATIONS

Pedro M. M. Marques

Dissertation for the Degree of Doctor of Philosophy in Electrical and Computer Engineering

Faculdade de Engenharia da Universidade do Porto

Porto, October 1, 2008

Dissertation submitted to
Faculdade de Engenharia da Universidade do Porto
in partial fulfillment of the requirements for
the Degree of Doctor of Philosophy in
Electrical and Computer Engineering

Author: Pedro Manuel Martins Marques
department, faculty, code: DEEC, FEUP, 030535009

Supervisor: Prof. Sílvio Almeida Abrantes Moreira
department, faculty: DEEC, FEUP

DEDICATION

To my family, and Raquel

ACKNOWLEDGMENTS

I would like to express my honest gratitude towards my thesis supervisor, Prof. Sílvio Abrantes, for his initial encouragement, his timely orientation and continued support. I thank him for his humane character and accessibility, for alleviating the pressure and for being patient when the pace of investigation became slow. I am indebted with him.

I also thank my family for being there at all times. Without their incentive and comfort this thesis would be hardly possible. I also cannot forget Raquel, whose dedication and appreciation brought new inspiration to my work.

Finally, I have to thank two institutions: 1) the Faculdade de Engenharia da Universidade do Porto (FEUP), Porto, Portugal, for approving my Ph.D program and providing electronic access to research literature; and 2) the Ministério da Ciência e Ensino Superior, Fundação para a Ciência e Tecnologia (FCT), Lisbon, Portugal, for the financial support of Grant SFRH/BD/17131/2004.

ABSTRACT

The pervasive challenge of increasingly higher spectral efficiencies in wireless communications has fomented the onset of various exceptional technologies. One of the most notorious is the concept of multiple-input multiple-output (MIMO) systems, a remarkable invention that brought the promise of extraordinary data rates of transmission. As it is gradually understood and shaped by the research community, that promise is progressively becoming reality.

The thesis herein investigates some open issues in present-day MIMO desiderata, and has the aim of filling important gaps. After a preliminary understanding of the mobile channel characteristics and the implications of wireless diversity, its primary focus is on an information-theoretic perspective of MIMO communications. Even though the capacity of ideal MIMO channels has been known for some time, that is not the case for more realistic scenarios such as the presence of cross-correlation between the antennas of MIMO systems. An in-depth study of the impact of receiver-sided correlated fading on the channel capacity is conducted, and novel, exact analytic (closed-form) formulas for the ergodic capacity are found in a unified derivation. The derivation involves the study of non-trivial multivariate statistical distributions, matrix integrals and the theories of groups and group representations. In the process, some results of the realm of mathematical statistics are derived, such as expressions for several complex Jacobians and the universal Wishart density function.

The thesis then departs from Information Theory and focuses on more practical aspects of MIMO transmission. No longer assuming that the channel is flat-fading, the goal is to study transmission overspreading in a unified framework, such that a more robust MIMO receiver can be implemented. A baseband description of the linear time-varying MIMO channel sets the ground to orthogonalization and full discretization, where a novel approach using continuous, discrete, and hybrid operators becomes indispensable foundation. A novel, optimal matched filter, the ORTHO-TS-MMF, is derived. To round out the study, a full frequency description using spectral operators is also obtained, and novel formulas for the spectral factorization of time-varying MIMO filters are recovered. Finally, after deriving three instances of MMSE based linear detectors, the optimal one is used to test the error performance of the devised discrete model using Monte Carlo simulation. The thesis finalizes with the numerical results.

RESUMO

O desafio permanente do aumento da eficiência espectral das comunicações sem fios tem fomentado o aparecimento de várias tecnologias excepcionais. Uma das mais notórias é o conceito de sistemas de múltipla-entrada múltipla-saída (conhecidos na literatura inglesa pela sigla MIMO), uma invenção notável que trouxe consigo a promessa de taxas de transmissão extraordinárias. À medida que o conceito é gradualmente compreendido e moldado pela comunidade de investigadores, a promessa tem-se tornado efectivamente realidade.

A presente tese investiga algumas questões em aberto no âmbito do desiderato MIMO actual, e almeja preencher lacunas importantes. Depois de uma compreensão preliminar das características do canal de comunicações móveis e as implicações da diversidade sem fios, o seu foco primeiro é numa perspectiva teórica da informação em comunicações MIMO. Apesar da capacidade de canais MIMO ideais ser conhecida já há algum tempo, o mesmo não se pode dizer da que concerne a cenários mais realistas tal como a presença de correlação cruzada entre as antenas dos sistemas MIMO. É realizado um estudo aprofundado do impacto de desvanecimentos correlacionados do lado do receptor, sendo obtidas novas fórmulas analíticas e exactas para a capacidade ergódica do canal, deduzidas numa derivação unificada. A derivação envolve o estudo de distribuições estatísticas multivariadas não triviais, integrais de matrizes e teorias de grupos e de representações de grupos. No processo, são derivados alguns resultados do domínio da estatística matemática tal como expressões para vários Jacobianos complexos e a função densidade de Wishart universal.

A tese afasta-se então da Teoria da Informação e centra-se em aspectos mais práticos da transmissão MIMO. Abandonando a condição de um canal não selectivo em frequência, o objecto passa a ser o estudo da sobre-dispersão da transmissão numa conjuntura unificada, viabilizando deste modo a implementação de receptores MIMO mais robustos. Uma descrição em banda base do canal MIMO linear e variável no tempo alicerça os processos de ortogonalização e discretização completa, onde uma nova abordagem baseada em operadores contínuos, discretos e híbridos, se traduz num fundamento indispensável. Adicionalmente, é derivado um novo filtro adaptado óptimo, designado por ORTHO-TS-MMF. Para alargar a abrangência do estudo, são ainda definidos operadores espectrais que conduzem a uma descrição completa no domínio das

frequências, em sequência recuperando novas fórmulas para a factorização espectral de filtros MIMO variáveis no tempo. Finalmente, depois de derivar três instâncias de detectores lineares baseados em estimação MMSE, o detector linear óptimo é usado para testar o desempenho do modelo discreto projectado, sendo, para tal, implementadas eficientes simulações Monte Carlo. A tese termina com os resultados numéricos.

RÉSUMÉ

Le défi permanent de l´augmentation de l'efficacité spectrale des communications sans fils (wireless) a fomenté l´apparition de plusieurs technologies exceptionnelles. L´une des plus notoires c´est le concept de multiple-entrée et multiple-sortie (connu dans la littérature anglaise par la sigle MIMO), une invention notable qui a porté sur soi une promesse de taux de transmission extraordinaires. À mesure que le concept est graduellement compris et modelé par la communauté de chercheurs, la promesse devient effectivement une réalité.

Cette thèse recherche quelques questions en ouvert, dans ce qui concerne l´actuel désidératum MIMO, et prétend remplir d´importantes lacunes. Aussitôt après la compréhension préliminaire des caractéristiques du canal des communications mobiles et ses implications dans la diversité sans fils, son premier cible c´est une perspective théorique de l´information dans les communications MIMO. Bien que la connaissance de la capacité des canaux MIMO idéaux soit connue il y a longtemps, on ne peut pas affirmer le même concernant des scènes plus réalistes tel que la présence de la corrélation croisée entre les antennes des systèmes MIMO. Un étude approfondi sur l´impact d´évanouissements corrélationnés du côté du récepteur a été réalisé, et on a obtenu de nouvelles formules analytiques e exactes pour la capacité ergodique du canal, déduites dans une formule unifiée. La dérivation comprend l'étude des distributions statistiques multivariées non triviaux, intégraux des matrices et des théories des groupes et des représentations de groupes. Dans ce processus, sont dérivés quelques résultats du domaine de la statistique mathématique tel que des expressions pour divers Jacobiens complexes et pour la fonction densité d'Wishart universelle.

Alors, cette thèse s'éloigne de la Théorie de l´Information et se centre dans des aspects plus pratiques de la transmission MIMO. Renonçant la condition d´un canal non sélectif en fréquence, l´objet c´est l´étude de la sur-dispersion de la transmission dans une conjunture unifiée, viabilisant ainsi l´implémentation de récepteurs MIMO plus robustes. Une description en bande-base du canal MIMO linéaire et variable dans le temps consolide les processus d´orthogonalisation et de discrétisation complète, où un nouvel abordage fondé dans des opérateurs continus, discrètes et hybrides, constitue un fondement indispensable. Additionnellement, un nouveau filtre adapté optime, désigné par ORTHO-TS-MMF, est dérivé. Pour élargir cet étude, sont encore définis des opérateurs spectraux qui

conduisent à une description complète dans le domaine des fréquences, en séquence récupérant de nouvelles formules pour la factorisation spectralle des filtres MIMO variables dans le temps. Finalement, après dériver trois instances de détecteurs linéaires fondés en estimation MMSE, le détecteur linéaire optime est utilisé pour tester l´accomplissement du modèle discret projeté, c´est pourquoi d´efficaces simulations Monte Carlo ont été réalisées. Cette thèse termine avec des résultats numériques.

PUBLICATIONS

Paper 1 - IEEE Transactions on Information Theory: CLN: 5-170
Title: "On the Derivation of the Closed-Form Capacity of MIMO Channels under Correlated Fading"
Authors: P. M. Marques, S. A. Abrantes
Status: submitted on March 10, 2005, accepted for revision.

Paper 2 - IEEE Transactions on Information Theory: CLN: 5-351
Title: "On the Exact, Closed-Form Capacity Formulas for MIMO Channels with an Arbitrary number of Inputs and Outputs, under Correlated Fading"
Authors: P. M. Marques, S. A. Abrantes
Status: submitted on April 14, 2005, accepted for revision.

Paper 3 - Unified development submitted on July 23, 2007, and published in the March 2008 issue of the IEEE Transactions on Information Theory:

P. M. Marques and S. A. Abrantes, "On the Derivation of the Exact, Closed-Form Capacity Formulas for Receiver-Sided Correlated MIMO Channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1139–1161, Mar. 2008.

Paper 4 -
Provisional Title: "Overspread Transmission over Wireless LTV MIMO Systems"
Authors: P. M. Marques, S. A. Abrantes
Status: in preparation; to be submitted to an international journal of reference.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Wireless mobile communications at the physical level is an exceedingly challenging topic in both theoretical and practical terms. The mobile propagation channel inherently suffers from severe impairments such as random small-scale fading and frequency selectivity that, without powerful mitigation techniques, deem reliable communication virtually impossible. Besides these impairments, the wireless medium is effectively shared by a lot of services to different users that are increasingly eager of high data rates, which means that either high spectral efficiencies are attained by proficient technologies or it will not be possible to accommodate all of them. Relevant sources of information regarding the modelling and characterization of the mobile channel are [1], [2], [3] and [4], and an extensive review is provided in [5].

One ineluctable tool to combat the wireless impairments is to take advantage of the random nature of the channel fading to explore *spatial diversity*. Spatial diversity is considered one of the most important instruments for wireless communications' reliability, as it introduces impressive channel amelioration by the employment of diversity combining techniques. Spatial diversity also allows for spatial multiplexing which, in essence, permits substantial data rate improvements and user separation in multiuser systems. A broad review of these facts is given in [6], [7], [8] and [9].

Multiple-input multiple-output (MIMO) systems have originated a full discipline of study which sole purpose is to formulate methods to fully exploit the enhancements of spatial diversity that are steering wireless mobile communications. Ranging from transmit beamforming to space-time processing and space-time coding, the goal is to obtain either diversity gain, spatial multiplexing gain, or interference reduction. In what concerns spatial multiplexing, Information Theory has provided results that foresee an almost linear scaling of the information-theoretic capacity with MIMO array sizes, confirming that the simultaneous use of transmit and receive diversity offers spectral efficiency improvements

[10], [11] and [12]. MIMO literature is vast, but for an introductory study see [13], [14], [15], [16] and [17], and for a comprehensive instalment [18], [19].

1.1 *Motivation*

The fundamental motivation for this thesis is the global urge for data rates in the order of hundreds or even thousands of megabits per second in a 4th generation (4G) mobile wireless communications setting. Inherent complexity makes the subject difficult, yet challenging, and the ability to grasp the mathematics that tackles the problem and moulds a solution to one's advantage is no less than stimulating. MIMO systems are an enthusiastic new technology that is becoming a landmark in our time, and will definitely drive the future of wireless communications.

All the beauty of MIMO systems can be directly drawn from theoretical principles of a multidimensional extension of classical Information Theory, and this is where the first far-reaching incursion is made. The goal is to deprive the formulas for attainable information rates in frequency non-selective MIMO systems from an open-form, computationally demanding situation, within a propagation context as general as possible. Due to the matricial nature of the MIMO model, this task is anything but easy when the degrees of freedom that MIMO is supposed to introduce are not so free after all. These are the correlated fading scenarios that the thesis investigates in pursuit of a possible solution.

Having concluded the information-theoretic study, the motivation becomes more practical. Now, in a frequency-selective and/or Doppler-spread set-up, the aim is to understand how to optimally design the MIMO input/output model when no space-time channel coding is used. MIMO channels also suffer from overspreading in multipath delay and in Doppler shift, so a basic challenge is to devise a receiver that copes with the problem in a time-variability aware manner. Instead of designing a receiver and then adapting it to the time-varying nature of the wireless MIMO channel, the approach is to assume time-variability from the beginning so as to explore Doppler diversity.

In the process of solving these questions, the thesis finds additional motivation in exciting new mathematical tools that it builds to reach its conclusions.

1.2 *Thesis Contributions*

So that the reader gets acquainted with the innovative scope of this thesis, the purpose of this section is to provide a brief summary of its most relevant contributions to the current research work. They are:

1. Derivation of the probability density function of the *Nakagami-m distribution* of envelope fading, by resorting to hyperspherical coordinates;

2. Derivation of the probability density function of the output signal-to-noise ratio of the maximal-ratio combiner under Nakagami-m distributed correlated envelope fading;

3. Derivation of closed-form formulas for the error probability and information-theoretic capacity of the maximal-ratio combiner under Nakagami-m distributed correlated envelope fading;

4. Derivation of the complex Jacobians of four matrix decompositions, namely: 1) the Cholesky factorization; 2) the QR factorization; 3) the congruence transformation; and 4) the eigenvalue decomposition; the Jacobians were derived for both nonsingular and singular complex matrices, making them universal;

5. Derivation of the density function of both nonsingular and singular, complex Wishart matrices, based on the transformations associated with the given Jacobians;

6. Derivation of the unordered eigenvalue distributions of both nonsingular and singular, complex Wishart matrices;

7. Extension of the Itzykson-Zuber integral to integration over a quotient space, using the theory of group representations;

8. Derivation of exact, closed-form formulas, for the information-theoretic capacity of wireless MIMO channels under receiver-sided correlated fading, in the configurations: 1) no more receiver antennas than transmitter antennas; and 2) more receiver antennas than transmitter antennas;

9. Full description of the input/output model of the linear time-varying MIMO channel using continuous, discrete and hybrid operators;

10. Orthogonalization of the MIMO channel response and determination of a fully discrete input/output model of confirmed optimality in the maximum likelihood sense; the latter model is obtained from the first by filtering with a new matrix matched filter, the ORTHO-TS-MMF; it is understood that the matched filter is always the hermitian adjoint operator;

11. Derivation of the time-varying noise whitening filter that applied at the output of the TS-MMF produces the same output as the ORTHO-TS-MMF;

12. Full discretization of the matched filter operators, which allows for a fully discrete receiver after sampling;

13. Derivation of constrained, unconstrained and zero-forcing linear detectors operators by the MMSE method of error energy minimization;

14. Implementation of a novel method of MIMO channel operator orthogonalization, based on backward block-wise Householder reflections; the method is numerically stable and extremely time-efficient (in certain cases, dozens or even hundreds of times faster than the conventional, unstable, backward block-wise Gram-Schmidt procedure); the speed-up made the ensuing Monte Carlo simulations possible to complete in due time.

1.3 *Thesis Outline*

The thesis is organized in logical progression. Chapter 2 is devoted to a preliminary study of the wireless mobile channel. It begins with the fundamentals of electromagnetic propagation and retrieves a model for the received signal of dynamic channels. After presenting some notions from large-scale path loss, it concentrates on the small-scale fading phenomenon. The channel is statistically characterized in terms of time and frequency autocorrelation functions and models are provided for the power delay profile and angular distribution of wave arrival. The spaced-time spaced-frequency autocorrelation function

of the wide-sense stationary uncorrelated (isotropic) scattering (WSSUS) channel model is derived, which leads to formulas for the *coherence time* and *coherence bandwidth* of the wireless channel. Simulations of the behaviour of time-selective and frequency-selective channels are performed. Also, the fading is statistically characterized in terms of the Rayleigh, Rice, and Nakagami-m distributions.

Chapter 3 evolves into the study of spatial diversity at the receiver end. A brief incursion into autocorrelation of the fading envelope sets the ground to the study of the maximal-ratio combiner as a predetection combining technique. Information-theoretic and error performance results expose first evidence of the benefit of spatial diversity even under correlation scenarios.

In Chapter 4 the information-theoretic scaffolding pertaining to MIMO systems is presented. Beginning with the model for a frequency non-selective, slowly fading channel, and the principle of maximum entropy, an all-round derivation of the information-theoretic capacity of MIMO channels is performed until exact, analytic formulas, are obtained. The complexity of the study lies in the presence of fading correlation at the receiver end of the MIMO system, which is a realistic situation in space-constrained (reduced antenna separations) mobile transceivers. The initial study is for the case of no more receiver antennas than transmitter antennas, and the more complex case of more receiver antennas is treated thereafter. Abundant plots are presented to demonstrate and validate the results.

Chapter 5 departs from the simple flat-fading assumption and aims to study overspreading in MIMO transmission. A baseband input/output model is conceived and, oriented by a maximum likelihood estimation approach, an optimal input/output discrete model is obtained. Arising from a general theory of matrix operators, time-variability aware matched filters are proposed. Then, a frequency domain description of the input/output model is made, trying to purvey new insight into the model construction and operation. To complete, a set of linear detectors for the mentioned model is obtained, and Monte Carlo simulations are used for testing the error performance of the global system.

Finally, Chapter 6 concludes the thesis and anticipates new vectors of research.

CHAPTER 2

MOBILE CHANNEL CHARACTERIZATION

2.1 *Introduction*

Due to its distinctive features, the mobile radio propagation channel places fundamental limits in the performance of wireless communication systems. Various types of phenomena affect the transmission of radio signals, the most severe of which are responsible for *multipath propagation* between the transmitter and the receiver. These are:

- *Reflection* - resulting from specular reflections at objects with dimensions larger than the wavelength of the propagating waves (e.g. ground, buildings, and walls).
- *Diffraction* - resulting from surface irregularities (edges) between the transmitter and receiver, which "bend" the propagating waves.
- *Scattering* - resulting from heaped up objects with dimensions shorter than the wavelength (e.g. rough surfaces, trees, and lamp posts).

Furthermore, due to the mobility of the transceivers (i.e. the mobile units), all these phenomena are time-variant, meaning that they become unpredictable and need to be treated stochastically. The first issue that arises when designing a communication system is what model to choose for the mobile radio environment, so that the performance can be maximized for a particular application. A thorough knowledge of the channel behaviour is needed to pursuit this goal, beginning with the study of fundamental aspects of electromagnetic theory, and the stochastic estimation of the channel output for a given transmitted signal. With this information at hand, signal processing can be employed to maximize the performance of the whole communication system. Also, the transmitted signal can be "intelligently formatted" to even improve the results. Some fundamental questions stand up: How can one exploit the knowledge of the channel characteristics to obtain an optimum design solution? And, may it be possible to put the channel impairments to one's favour?

In general, the mobile radio propagation environment is characterized by two partially

separable effects: *large-scale path loss*, which is a function of the distance between transmitter and receiver, and the presence of large obstructions along the path; and *small-scale fading*, which results from constructive and destructive interference of multipath propagation, and varies over distances proportional to the signal wavelength.

2.2 Fundamental Propagation Aspects

This discussion firstly introduces some basic antenna theory, and then proceeds to estimate the channel dependent, received signal characteristics. A good direction to take within the analysis is to begin with ideal conditions of propagation (i.e. no multipath propagation phenomena and a free-space medium) and after that, generalize the results to non-ideal conditions.

Using spherical coordinates, the *far zone* - $\beta d \gg 1$ (β is the phase constant and d is the distance from the radiating element) - electric and magnetic fields produced by a radiating element (antenna) operating in free-space can be written generically in phasor (baseband) notation as:

$$\mathbf{E} \approx E_m \frac{e^{j(\varphi_0 - \beta d)}}{d} S(\theta, \phi)\, \mathbf{a}_\theta \tag{2.1}$$

$$\mathbf{H} \approx \frac{E_m}{\eta} \frac{e^{j(\varphi_0 - \beta d)}}{d} S(\theta, \phi)\, \mathbf{a}_\varphi \tag{2.2}$$

where

E_m = amplitude function proportional to excitation current
$S(\theta, \phi)$ = E-plane *radiation pattern* or *space factor*
η = characteristic impedance of the medium
θ / ϕ = elevation/azimuth angle
φ_0 = fixed phase offset

Also, for $d \geq 2D^2 / \lambda$, where D is the antenna dimension and λ is the wavelength, the electromagnetic field is within the *Fraunhofer region*, which means that $S(\theta, \phi)$ may be considered independent of d [20]. The average power density in a given direction may then be expressed as

$$\begin{aligned} S_{av} &= \frac{1}{2} \mathrm{Re}\left\{ E \times H^* \right\} \\ &= \frac{|E|^2}{2\eta} \\ &= \frac{E_m^2}{2\eta} \frac{S^2(\theta, \phi)}{d^2} \end{aligned} \tag{2.3}$$

The directivity of an antenna is a power measure of its directional pattern, and is defined as the ratio of the radiated power density to the power density averaged over all directions (equivalent isotropic radiation), that is

$$D(\theta, \phi) = \frac{S_{av}}{P_{rad} / 4\pi d^2} \tag{2.4}$$

where P_{rad} is the total radiated power. To expand the conduction, dielectric, and mismatch losses in the antenna, (2.4) can be reformulated to

$$D(\theta,\phi) = \frac{S_{av}}{\varepsilon_T P_T \,/\, 4\pi d^2} \tag{2.5}$$

where P_T is the total input power at the transmitter side and ε_T is the total radiation efficiency of the antenna. Neglecting polarization losses, the effective directional power gain to be used in link calculations may be obtained from (2.5) as

$$G_T(\theta,\phi) = \varepsilon_T D(\theta,\phi) = \frac{S_{av}}{P_T \,/\, 4\pi d^2} \tag{2.6}$$

2.2.1 Free-Space Power Transmission

Looking at (2.6), it becomes apparent that the power delivered to a matched load connected to a receiver antenna (with effective area A_R) in line-of-sight with the transmitter is

$$P_L = P_T G_T(\theta,\phi)\frac{A_R}{4\pi d^2} \tag{2.7}$$

A general expression for the effective area of a receiving antenna can be derived by exchanging the transmission/reception roles of two communicating antennas. For a linear, passive, and isotropic medium, and when both antennas are matched to the connected loads, the following relation must be true

$$P_L = P_T G_T(\theta,\phi)\frac{A_R}{4\pi d^2} = P_T G_R(\theta,\phi)\frac{A_T}{4\pi d^2} \tag{2.8}$$

Equation (2.8) interprets the fact that a change in the input-output direction of a two-antenna system does not affect the system response and, in this particular case, the power response. It is a direct consequence of the *reciprocity and reaction* theorems of electromagnetic theory. Thus, simplifying (2.8), the relation

$$A_R = \frac{A_T}{G_T(\theta,\phi)} G_R(\theta,\phi) \tag{2.9}$$

states that the effective area of a receiving antenna is related to its directional power gain by the constant $A_T \,/\, G_T$. More specifically, since the effective area of reception should be independent of the transmitter antenna, it asserts that the quotient $A_T \,/\, G_T$ may be a constant independent of the type of antenna. It can be proved generically that it is so. In particular, the well known effective area and power gain of an infinitesimal dipole are given by [21]

$$\begin{aligned} A_T &= \frac{3\lambda^2}{8\pi}\sin^2\theta \\ G_T &= \frac{3}{2}\sin^2\theta \end{aligned} \tag{2.10}$$

and their ratio is

$$\frac{A_T}{G_T} = \frac{\lambda^2}{4\pi} \tag{2.11}$$

Substituting (2.11) into (2.9), the desired expression for the effective area of an antenna is obtained as

$$A_R = \frac{\lambda^2}{4\pi} G_R(\theta,\phi) \tag{2.12}$$

which, after direct substitution into (2.7), relates the power P_L (delivered to the receiver load) to the input power of the transmitting antenna P_T - the *Friis Transmission Formula*:

$$P_L = P_T G_T(\theta,\phi) G_R(\theta,\phi) \left(\frac{\lambda}{4\pi d}\right)^2 \tag{2.13}$$

The term $(\lambda/4\pi d)^2$ is called the *free-space loss factor*, and it takes into consideration the losses due to the spherical spreading of the energy by the antenna.

2.2.2 RECEIVED SIGNAL FOR STATIC CHANNELS

Although (2.13) assumes line-of-sight and free-space conditions, it may be used as a basis to find a general expression for the voltage V_L induced at the receiver load $Z_r = R_r + jX_r$. For a matched load, and neglecting the antenna input reactance X_r, the delivered power is also given by

$$\begin{aligned} P_L &= \frac{1}{2}\mathrm{Re}\left\{V_L I_L^*\right\} \\ &= \frac{1}{2}\mathrm{Re}\left\{\frac{V_L V_L^*}{R_r}\right\} \\ &= \frac{|V_L|^2}{2R_r} \end{aligned} \tag{2.14}$$

which equated to (2.13) gives

$$\begin{aligned} |V_L| &= \sqrt{2R_r P_L} \\ &= \sqrt{2R_r P_T G_T(\theta,\phi) G_R(\theta,\phi)} \left(\frac{\lambda}{4\pi d}\right) \end{aligned} \tag{2.15}$$

Moreover, the voltage phase should be equal to the phase of the impinging wave plus a phase shift δ due to the receiver circuits. The complete voltage expression is

$$V_L = \sqrt{2R_r P_T G_T(\theta,\phi) G_R(\theta,\phi)} \left(\frac{\lambda}{4\pi d}\right) e^{j(\varphi_0 - \beta d - \delta)} \tag{2.16}$$

To inspect how the load voltage relates to the impinging wave, (2.6) can be substituted into (2.16), giving

$$\begin{aligned} V_L &= \sqrt{2R_r (4\pi d^2 S_{av}) G_R(\theta,\phi)} \left(\frac{\lambda}{4\pi d}\right) e^{j(\varphi_0 - \beta d - \delta)} \\ &= \sqrt{2R_r (4\pi d^2 \frac{|E|^2}{2\eta}) G_R(\theta,\phi)} \left(\frac{\lambda}{4\pi d}\right) e^{j(\varphi_0 - \beta d - \delta)} \\ &= \sqrt{\frac{R_r G_R(\theta,\phi)}{4\pi\eta}}\, \lambda\, e^{-j\delta}\, |E|\, e^{j(\varphi_0 - \beta d)} \\ &= E_{rad}\, L(\theta,\phi) \end{aligned} \tag{2.17}$$

where $L(\theta,\phi) = \sqrt{R_r G_R(\theta,\phi)/4\pi\eta}\,\lambda e^{-j\delta}$ and $E_{rad} = |E|e^{j(\varphi_0-\beta d)}$. Apart from a phase shift, (2.17) shows a direct proportion relationship between incident field and load voltage. It can be used to induct the general load voltage when the propagation conditions of the mobile radio channel are not the ideal ones.

In a realistic situation, there may exist "objects" nearby the transmitting and receiving antennas. These "objects" include natural and man-made structures, such as the ground, trees, hills, buildings, walls, and towers. For frequencies above 100 MHz, these structures reflect/scatter much of the incoming waves, possibly leading to several replicas of the same signal arriving at the receiving antenna. Using the superposition principle, the total impinging field may be written as

$$E_{rad} = \sum_i E_i \Gamma_i \tag{2.18}$$

where $\Gamma_i = \rho_i e^{j\alpha_i}$ is the equivalent reflection coefficient associated with the i-th path, which includes all the reflections undergone by the path. It is assumed that the polarization mismatches between the incident waves and the receiving antenna are negligible, which is a good approximation for vertical and horizontal reflectors. Also, for high frequencies (>100 MHz) and grazing reflection angles, $\Gamma_i \approx -1$ irrespective of the polarization. The total voltage delivered to the load will be given by

$$\begin{aligned} V_L &= \sum_i E_i \Gamma_i L(\theta_i,\phi_i) \\ &= \sum_i \sqrt{2R_r P_T G_T(\theta_i,\phi_i) G_R(\theta_i,\phi_i)} \left(\frac{\lambda}{4\pi d_i}\right) \rho_i e^{j(\varphi_0+\alpha_i-\beta d_i-\delta)} \\ &= \sum_i \upsilon_i e^{j\varphi_i} \end{aligned} \tag{2.19}$$

Each signal component experiences a different multipath environment which determines the amplitude $\upsilon_i = \sqrt{2R_r P_T G_T(\theta_i,\phi_i) G_R(\theta_i,\phi_i)}\lambda\rho_i/(4\pi d_i)$, and the associated phase shift $\varphi_i = \varphi_0 + \alpha_i - \beta d_i - \delta$. Furthermore, the transmitter may linearly modulate the propagating wave with a baseband signal, say

$$s_b(t) = \sum_{n=-\infty}^{\infty} I_n\, g(t-nT) \tag{2.20}$$

where the transmission rate is $1/T$, $\{I_n\}$ represents the sequence of complex symbols mapped from k-bit blocks, and $g(t)$ is a pulse shaping function. If this is the case, the load voltage must be rewritten as

$$V_L(t) = \sum_i \upsilon_i e^{j\varphi_i} s_b(t-\tau_i) \tag{2.21}$$

where τ_i is the time delay of the i-th propagation path.

2.2.3 Received Signal for Dynamic Channels

Equation (2.21) is only valid for static radio channel conditions, that is, when neither the environment "objects" nor the transceivers are in relative motion. When the multipath environment is "dynamic", there is a continuous change in the length of each propagation

path, the reflection coefficients, and angles of wave departure and incidence. Thus, the received voltage amplitudes and phase shifts are a function of the transceivers' spatial location and, therefore, a function of time.

The received signal variations can be related to the motion of the propagation environment using the *Doppler shift* effect. When a receiver moves with a velocity ν relative to the transmitter, and with an angle ϕ towards the incoming wave, the apparent change in angular frequency is

$$\omega_i = 2\pi \frac{\nu}{\lambda} \cos \phi_i \tag{2.22}$$

Hence, to take channel related time variations (i.e. spatial motion) into account, (2.21) should be modified to

$$\begin{aligned} V_L(t) &= \sum_i v_i(t) e^{j\varphi_i(t)} s_b(t - \tau_i(t)) \\ &= \sum_i v_i(t) e^{j(\varphi_i + \omega_i t)} s_b\left(t - \tau_i + \frac{\omega_i}{\omega_c} t\right) \end{aligned} \tag{2.23}$$

where ω_i is the *Doppler shift* and ω_c is the carrier frequency.

2.3 *Large-Scale Path Loss*

The *Friis Transmission Formula* - (2.13), states that the path loss of radio waves propagating through free-space is proportional to the inverse of d^2, where d is the distance between the transmitter and receiver. However, propagation in a mobile radio channel is neither free-space nor line-of-sight. Determining the received signal power in a multipath environment is a difficult problem, nevertheless a basic understanding may be achieved by separating the essential effects of reflections.

2.3.1 PATH LOSS POWER LAW

Reflections can be roughly divided into two types: horizontal reflections and vertical reflections. Horizontal reflections occur from vertical objects, such as building walls, and are characterized by an approximate power decay as d^{-2} (i.e. the free-space decay). Vertical reflections occur from horizontal objects, such as the ground, and combine at the receiver with the line-of-sight paths. To study the vertical reflection of waves, (2.19) can be simplified by assuming constant gain patterns for the direct and reflected rays, which is a good assumption for grazing reflection angles. Moreover, for large distances between transmitter and receiver (as compared to their heights), the load voltage can be modified to

$$V_L = \sqrt{2R_r P_T G_T G_R} \left(\frac{\lambda}{4\pi d}\right) \sum_{i=1}^{2} \rho_i e^{j(\varphi_0 + \alpha_i - \beta d_i - \delta)} \tag{2.24}$$

where a simplified two-ray model is assumed, with single direct ($i = 1$) and reflected ($i = 2$) rays. Also, $\Gamma_1 = 1$ and $\Gamma_2 \approx -1$. We shall consider that the transmit antenna is lo-

cated at height h_t above a flat ground, and the receive antenna is located at height h_r. The total power delivered to the load can now be expressed as

$$P_L = P_T G_T G_R \left(\frac{\lambda}{4\pi d}\right)^2 \left|1 - e^{-j\beta(d_2 - d_1)}\right|^2 \tag{2.25}$$

where $d_1 = \sqrt{d^2 + (h_t - h_r)^2}$ and $d_2 = \sqrt{d^2 + (h_t + h_r)^2}$. For large d, (2.25) can be simplified to (see e.g. [3])

$$\begin{aligned} P_L &\approx P_T G_T G_R \left(\frac{\lambda}{4\pi d}\right)^2 \left|1 - e^{-j2\beta h_t h_r / d}\right|^2 \\ &\approx P_T G_T G_R \left(\frac{h_t h_r}{d^2}\right)^2 \end{aligned} \tag{2.26}$$

Equation (2.26) shows that, in the limiting case of very large distances, for a two-ray ground reflection model the received power falls off as d^{-4}, which is much more severe compared to that of free-space.

Figure 2.1 shows a plot of (2.25) as a function of the distance between transmitter and

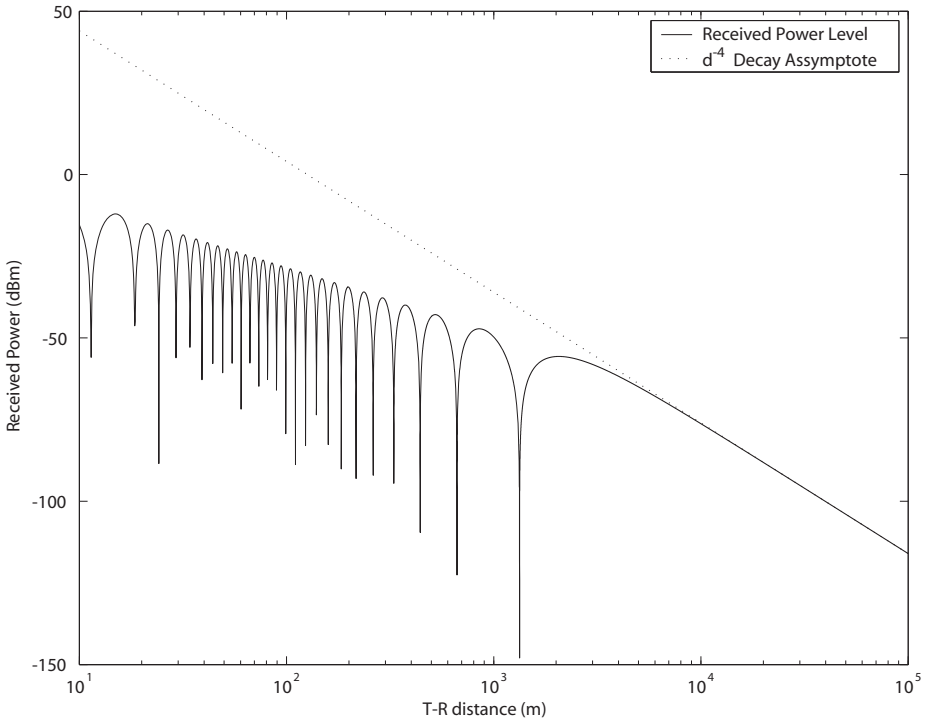

FIGURE 2.1 Received power as a function of transmitter-receiver (T-R) separation for the ground reflection model, and asymptotic behaviour for large d.

receiver for a typical WCDMA reverse link. The plot parameters are: $P_T = 24$ dBm; $G_T = 18$ dBi; $h_t = 50$ m; $G_R = 2$ dBi; $h_r = 2$ m; $f = 2$ GHz. Up to a critical distance $d_c \approx 5$ km , the wave experiences constructive and destructive interference of the two rays, and the average power fall off with distance corresponds to free space. For $d > d_c$ the average power fall off with distance is approximated by the fourth power law in (2.26).

Empirical results have shown that this is the worst case propagation path loss decay, and fundament the decay law as being inversely proportional to $d^p, 2 \leq p \leq 4$. However,

neither (2.13) nor (2.26) hold for $d = 0$. For this reason, large-scale propagation models use a reference distance d_0, known as the *close-in distance*. For instance, the free-space transmission equation is changed as follows

$$\begin{aligned} P_L &= P_T G_T(\theta,\phi) G_R(\theta,\phi) \left(\frac{\lambda}{4\pi d_0}\right)^2 \left(\frac{d_0}{d}\right)^2 \\ &= P_L(d_0) \left(\frac{d_0}{d}\right)^2 \end{aligned} \tag{2.27}$$

In general, the received power will be written as

$$P_L = P_L(d_0) \left(\frac{d_0}{d}\right)^p \tag{2.28}$$

where the value $P_L(d_0)$ may be predicted from (2.13), or may be measured by averaging the received power at many points near d_0. When multipath fading is present, the measurements should span distances between 40λ and 200λ, so that the short-term fading components are removed and the details of the local means are maintained [22]. From (2.28), the average power path loss in dB is given by

$$\overline{L}_{dB} = \overline{L}_{dB}(d_0) + 10p \log_{10}\left(\frac{d}{d_0}\right) \tag{2.29}$$

which represents a straight line with a slope equal to $10p$ dB per decade.

In practice, actual propagation paths will be more complicated than the idealized cases, due to the possibility of multiple reflections, diffraction, wave scattering, and shadowing. The following sections will study the last two.

2.3.2 WAVE SCATTERING

When the propagating waves are incident on smooth surfaces, specular reflections occur and Snell's law applies. However, if the surfaces contain protuberances with dimensions similar to the wavelength of the incident waves, a diffuse reflection occurs which may not be characterized by a simple reflection coefficient. A fraction of the incident energy may be scattered in the direction of the receiving antenna, making the received signal often stronger than what is predicted by reflection models.

The model shown in Figure 2.2 may be used to characterize surface roughness by analysing what happens to a wavefront when it impinges on a protuberance with height η.

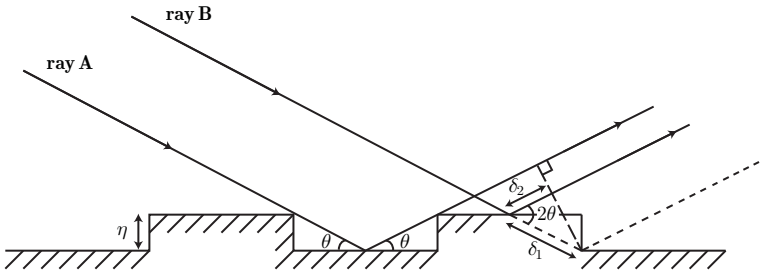

FIGURE 2.2 Surface roughness model for *Rayleigh criterion*.

If one ray of the wave front hits the protuberance (ray B) and other ray misses it (ray A), the difference in path length of the two rays when they meet again is

$$\begin{aligned}\delta &= \delta_1 - \delta_2 = \delta_1 - \delta_1 \cos 2\theta \\ &= \frac{\eta}{\sin\theta} - \frac{\eta}{\sin\theta}\cos 2\theta = \frac{\eta}{\sin\theta}(1 - \cos 2\theta) \\ &= 2\eta \sin\theta\end{aligned} \tag{2.30}$$

and the phase difference is

$$\begin{aligned}\Delta\phi &= \frac{2\pi}{\lambda}(\delta_1 - \delta_2) \\ &= \frac{4\pi\eta \sin\theta}{\lambda}\end{aligned} \tag{2.31}$$

When the protuberances are very small compared to the wavelength, then $\Delta\phi \approx 0$, and the two reflected rays will add almost in phase. When $\Delta\phi \approx \pi$, the reflected rays will be out of phase and cancel each other. A valid boundary for roughness is then $\Delta\phi = \pi/2$, which after direct substitution into (2.31) yields

$$\eta_c = \frac{\lambda}{8\sin\theta} \tag{2.32}$$

where η_c is the critical protuberance height. The *Rayleigh criterion* for roughness states that a surface is considered rough if $\eta > \eta_c$ (see e.g. [23]).

In practice, the value used as a measure of surface undulation height is σ_η, the standard deviation of the surface irregularities relative to the mean height. Additionally, for typical mobile radio scenarios, the incidence angles are grazing, and justify the approximation $\sin\theta \approx \theta$. Replacing η by σ_η in (2.31), we obtain the following roughness measure

$$\Delta = \frac{4\pi\sigma_d\theta}{\lambda} \tag{2.33}$$

The usual criterion for roughness is $\Delta > 10$, whereas the criterion for smoothness is $\Delta < 0.1$. For example, at 2 GHz and for $\theta = 1^\circ$, the value of σ_η necessary to make a surface rough is 7 m. This means that typical objects such as vehicles, trees, and houses will cause wave scattering.

Wave scattering may be accounted for in the propagation models by adjusting the reflection coefficient of surfaces

$$\Gamma_{rough} = \rho_s\Gamma \tag{2.34}$$

where ρ_s is the *scattering loss factor*, and can be modelled as [26]

$$\rho_s = \exp\left[-8\left(\frac{\pi\sigma_\eta \sin\theta}{\lambda}\right)^2\right] I_0\left[8\left(\frac{\pi\sigma_\eta \sin\theta}{\lambda}\right)^2\right] \tag{2.35}$$

where $I_0(\cdot)$ is the zeroth order modified Bessel function of the first kind.

2.3.3 Log-Normal Shadowing

When large obstructions are present between the transmitter and receiver, the path loss as given by (2.29) may be very different at two different locations with the same T-R separa-

tion. Different levels of clutter on the propagation path lead to random *shadowing* effects at the receiver, and are usually modelled by a log-normally distributed, slowly varying random process [24], [25]. The random process is usually considered to be a multiplicative process

$$r(t) = a(t)s(t) \tag{2.36}$$

where $s(t)$ is the expected signal level obtained from (2.28). As the T-R separation changes, the presence or not of an obstruction on the path will cause a change in the expected signal level given by

$$a(t) = e^{\alpha(t)} \tag{2.37}$$

where $\alpha(t)$ is the attenuation/gain in nepers. Moreover, in a multiple obstructions scenario the change in signal level will be

$$a(t) = e^{\sum_i \alpha_i(t)} \tag{2.38}$$

where the $\{\alpha_i(t)\}$ are assumed to be zero-mean, independent random variables. Thus, the power loss variation may be expressed stochastically as

$$\begin{aligned} X_\sigma &= 20\log_{10} a(t) \\ &= \sum_i 20\alpha_i(t)\log_{10} e \end{aligned} \tag{2.39}$$

For a large number of obstructions, the central limit theorem suggests that X_σ will be approximately a zero-mean, Gaussian distributed random variable with standard deviation σ (dB). In addition, $a(t)$ may be regarded as a sample function of a log-normally distributed random process. X_σ represents the path loss in dB and can be included in (2.29) to give

$$\overline{L}_{dB} = \overline{L}_{dB}(d_0) + 10p\log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \tag{2.40}$$

In practice, σ is computed from measured data. The probability distribution of X_σ may then be used to determine the probability that the received signal level will exceed or fall below a particular level.

2.4 *Small-Scale Fading*

The small-scale fading of the signal in a mobile radio environment is a phenomenon that occurs when either the transmitter or the receiver are surrounded by nearby scatterers, such as houses, buildings, walls, and trees. These scatterers will create multiple propagation paths between the transmitter and the receiver, each arriving with a particular phase delay.

Furthermore, (2.32) reveals that as the operating frequency rises the more intense the wave scattering becomes, because more and more objects of smaller dimensions are then prone to scatter the incoming waves. For instance, the light frequencies are so high (in the order of 10^{15} Hz), that even the irregularities of a small sheet of paper produce wave scattering, and that is why the paper does not perform as a specular reflective surface, appearing white.

Therefore, and because there is an increasing tendency for wideband wireless communications and higher operating frequencies, the scattering effects become of more significance and must be considered.

The equivalent baseband received signal for a multipath environment is given by (2.23) and reproduced here for convenience

$$r_b(t) = \sum_{i=0}^{N-1} v_i(t) e^{j(\varphi_i + \omega_i t)} s_b\left(t - \tau_i + \frac{\omega_i}{\omega_c} t\right) \tag{2.41}$$

where

$s_b(t) =$ baseband digital transmitted signal
$v_i(t) =$ amplitude of the i$^{\text{th}}$ propagation path
N = number of multipath replicas
$\varphi_i =$ phase shift of the i$^{\text{th}}$ propagation path
$\omega_i =$ Doppler shift of the i$^{\text{th}}$ propagation path
$\tau_i =$ time delay of the i$^{\text{th}}$ propagation path
$\omega_c =$ carrier frequency

Some considerations may be used to simplify (2.41) while keeping the small-scale fading information. First, since the *Doppler shift* is very small compared to the carrier frequency, the delay variation given by $(\omega_i / \omega_c)t$ is negligible during a symbol period. For example, at 2 GHz and for a mobile unit velocity of 200 km/h, the maximum *Doppler shift* is about 400 Hz, while the data/chip rates are around 5 Mbps (e.g. WCDMA), which means that approximately 50000 symbols are transmitted before a delay variation of 1% the symbol period is reached. These 50000 symbols correspond approximately to a 4λ displacement of the mobile unit, which is enough to preserve the fading behaviour of the signal. Second, since 4λ is small compared to the multipath distances d_i between transmitter and receiver, the voltage amplitude $v_i(t)$ of each propagation path may be considered time-invariant during the transmission. Assuming large distances between transmitter and receiver, and near-horizontal wave propagation, $v_i(t)$ may be reformulated as

$$\begin{aligned} v_i(t) &= \sqrt{2R_r P_T G_T(\theta_i, \phi_i) G_R(\theta_i, \phi_i)} \left(\frac{\lambda}{4\pi d_i}\right) \rho_i \\ &= \sqrt{2R_r P_T G_T G_R} \left(\frac{\lambda}{4\pi d_0}\right) \frac{d_0}{d_i} \rho_i \\ &= v_0 \frac{d_0}{d_i} \rho_i \end{aligned} \tag{2.42}$$

where d_0 is still the close-in distance, the received power reference point. Introducing these simplifications, (2.41) reduces to

$$r_b(t) = v_0 \sum_{i=0}^{N-1} c_i e^{j(\varphi_i + \omega_i t)} s_b(t - \tau_i) \tag{2.43}$$

where $c_i = \rho_i (d_0 / d_i)$.

Apart from the deterministic digital baseband signal $s_b(t)$, (2.43) clearly shows that the received signal is the sum of complex valued stochastic time-variant phasors, which, depending on the relative phases, will add in or cancel out, causing the multipath fading ef-

fect. It should be noticed that relatively slow motions of the medium can cause significant phase changes. For instance, a $\lambda/4$ displacement of the mobile unit causes a phase change of $\pi/2$ and $-\pi/2$ radians in the multipath components arriving at 0º and 180º relative to the velocity vector, respectively. If the components were initially in-phase, after the movement they will be out-of-phase and will cancel each other. This example agrees with experimental data, which shows that the deep fades of the signal occur in intervals of $\lambda/2$ displacements of the mobile unit.

2.4.1 STATISTICAL CHANNEL CHARACTERIZATION

The equivalent baseband time-variant response of the mobile radio channel to an input equivalent baseband signal, may be written as the convolution

$$r_b(t) = \int_{-\infty}^{\infty} h_b(\tau, t) s_b(t-\tau) d\tau \tag{2.44}$$

where $h_b(\tau, t)$ represents the channel response at time t due to an impulse applied at time $t-\tau$. Comparing (2.44) with (2.43), the equivalent baseband channel impulse response is obtained as a sequence of delayed impulses, each scaled by a multipath propagation stochastic phasor, that is

$$h_b(\tau, t) = v_0 \sum_{i=0}^{N-1} c_i e^{j(\varphi_i + \omega_i t)} \delta(\tau - \tau_i) \tag{2.45}$$

As can be seen, the phase is the only time-variant parameter, meaning that the amplitude (and power) associated with each propagation path remains constant.

To analyse how the power distributes among the multiple propagation paths, the autocorrelation function of the channel impulse response may be computed as

$$R_{hh}(\tau_1, \tau_2; t_1, t_2) = \frac{1}{2} E\left[h_b^*(\tau_1, t_1) h_b(\tau_2, t_2)\right] \tag{2.46}$$

Assuming that the propagation paths are uncorrelated with one another (independent scatterers), the autocorrelation function is only non-zero for $\tau_1 = \tau_2$, that is

$$R_{hh}(\tau_1, \tau_2; t_1, t_2) = \frac{1}{2} E\left[h_b^*(\tau_1, t_1) h_b(\tau_1, t_2)\right] \delta(\tau_1 - \tau_2) \tag{2.47}$$

Also, if the channel is wide-sense stationary, the autocorrelation function only depends on the difference $\Delta t = t_1 - t_2$ of observation times

$$\begin{aligned} R_{hh}(\tau_1, \tau_2; t_1, t_2) &= \frac{1}{2} E\left[h_b^*(\tau_1, t_1) h_b(\tau_1, t_1 + \Delta t)\right] \delta(\tau_1 - \tau_2) \\ &= R_{hh}(\tau_1, \Delta t) \delta(\tau_1 - \tau_2) \end{aligned} \tag{2.48}$$

and $R_{hh}(\tau, \Delta t)$ is called the *delay cross-power density*. This type of channel is called a *wide-sense stationary - uncorrelated scattering* (WSSUS) channel, and it has shown to be a good model for the mobile radio channel over small-scale distance or time intervals [27]. The WSSUS channel impulse response is approximated by (2.45), where the stochastic parameters c_i, φ_i, ω_i, and τ_i are considered statistically independent and identically distributed (i.i.d.) from path to path (uncorrelated scattering) and time-invariant for small-scale dis-

tances (wide-sense stationarity).

2.4.2 DELAY PROFILE MEASUREMENTS

Setting $\Delta t = 0$ in the impulse response autocorrelation function, $R_{hh}(\tau)$ results. It is called the delay power density or *power delay profile* of the channel, and is measurable by transmitting very narrow probe pulses and squaring the received signal amplitude. If the probe pulse is defined as

$$p(t) = \begin{cases} \sqrt{2/T_p} & 0 \le t \le T_p \\ 0 & T_p < t \end{cases} \tag{2.49}$$

where T_p is the pulse width, the received signal will be

$$h_b'(t) = \upsilon_0 \sum_{i=0}^{N-1} c_i e^{j(\varphi_i + \omega_i t)} p(t - \tau_i) \tag{2.50}$$

Since any practical measuring system is bandwidth-limited, it is often difficult to resolve all multipath components arriving at the receiver. Some multipath components will then vectorially combine to yield a multipath amplitude fading process, which means that the measured $R_{hh}(\tau)$ will be time-variant over small-scale distances. If the bandwidth of the measuring system is high enough to resolve every multipath component, avoiding the fading process, an approximate replica of $R_{hh}(\tau)$ may be obtained as

$$\begin{aligned} R_{hh}'(\tau) &= \frac{1}{2} h_b'(t) h_b'^{*}(t) \\ &= \frac{1}{2} \upsilon_0^2 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} c_i c_j e^{j(\varphi_i - \varphi_j + \omega_i t - \omega_j t)} p(t - \tau_i) p(t - \tau_j) \\ &= \frac{1}{2} \upsilon_0^2 \sum_{i=0}^{N-1} c_i^2 \, p^2(t - \tau_i) \\ &= \frac{1}{T_p} R_{hh}(\tau) * rect\left(\frac{t}{T_p}\right) \end{aligned} \tag{2.51}$$

where the bandwidth of the measuring system is $1/T_p$ and should be sufficient to resolve all multipaths. If the bandwidth is not high enough, then a small-scale distance average $\overline{R_{hh}'(\tau)}$ must be obtained to conceal the fading processes. A novel approach based on pseudo-random sequences for high resolution delay profile measurements was presented in [28]. A typical measure of the delay power density (relative to the time and power of the first arriving multipath component) is shown in Figure 2.3.

The delay dispersion parameters that characterize the delay power density are the first

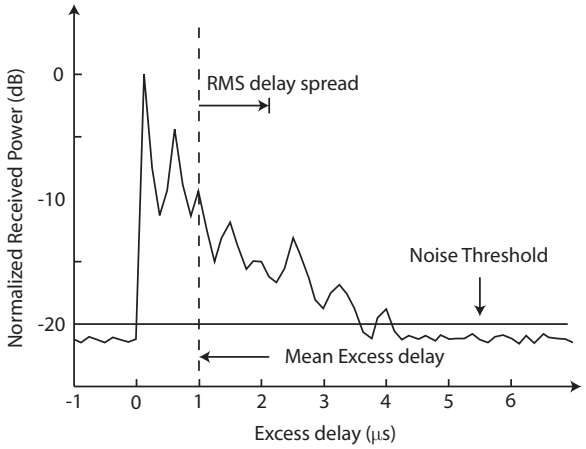

FIGURE 2.3 Typical measured excess *delay power density* for an urban environment.

moment or *mean excess delay*

$$\bar{\tau} = \int_0^\infty \tau R_{hh}(\tau)d\tau \Big/ \int_0^\infty R_{hh}(\tau)d\tau \tag{2.52}$$

and the square root of the second central moment or *RMS delay spread*

$$\sigma_\tau = \sqrt{\left(\int_0^\infty \tau^2 R_{hh}(\tau)d\tau \Big/ \int_0^\infty R_{hh}(\tau)d\tau\right) - \bar{\tau}^2} \tag{2.53}$$

2.4.3 DELAY PROFILE MODELLING

The exponentially decaying attribute of the delay power density is easily understood. In an urban propagation environment, the receiver is far from the transmitter and cluttered with nearby reflectors and scatterers, so that the distance spread ($\Delta d = \max d_i - \min d_i$) of the arising multipath components is much less than the propagation distance d and, therefore, that the difference in attenuation between the multipath components will be mainly due to the effective reflection coefficient ρ_i of each component. If the i-th component experiences n_i reflections, each of which with reflection coefficient ρ, the power it carries can be expressed as

$$P_i = P_0 \rho_i^2 = P_0 \rho^{2n_i} \tag{2.54}$$

where P_0 is a multiplicative constant and $\rho_i = \rho^{n_i}$. Moreover, the number of reflections can be estimated to be proportional to the propagation delay of the multipath component, $n_i = k\tau_i$, and (2.54) may be rewritten as

$$\begin{aligned} P_i &= P_0 \rho^{2n_i} \\ &= P_0 \rho^{2k\tau_i} \\ &= P_0 e^{-2k\tau_i \ln(1/\rho)} \\ &= P_0 e^{-\tau_i/\sigma_\tau} \end{aligned} \tag{2.55}$$

where $\sigma_\tau = [2k\ln(1/\rho)]^{-1}$ is the channel dependent *RMS delay spread*. If the multipath spreading is assumed a continuous process, the power delay profile is approximated by

$$R_{hh}(\tau) = P_0 e^{-\tau/\sigma_\tau} \tag{2.56}$$

Equation (2.56) justifies the exponential decay shown in Figure 2.3, but naturally does not take into account the dominant reflections of the waves, which may appear at the receiver as strong signals with large excess delays. Furthermore, the delay probability distribution can be expressed as

$$\begin{aligned} p(\tau) &= \frac{R_{hh}(\tau)}{\int_0^\infty R_{hh}(\tau)d\tau} \\ &= \frac{P_0 e^{-\tau/\sigma_\tau}}{\int_0^\infty P_0 e^{-\tau/\sigma_\tau}\, d\tau} = \frac{1}{\sigma_\tau} e^{-\tau/\sigma_\tau} \end{aligned} \tag{2.57}$$

In fact, if the propagation delay is partitioned into small delay bins, and since many of the multipath components arrive closely in time and few arrive spaced far apart, the power received within each delay bin should be approximately proportional to the number of multipath components within the bin. Therefore, from (2.57), the multipath arrival delay τ may be treated as an exponentially distributed random variable [24], [25].

The RMS delay spread σ_τ is used as a measure of the range of values of τ over which $R_{hh}(\tau)$ is essentially nonzero or, in other words, as a measure of the delay dispersion of the different multipath components relative to the mean delay. The typical mean RMS delay spreads in suburban and urban areas are 0.5 μs and 1.3 μs, respectively, and the maximum RMS delay spread generally used as a rule of thumb for calculations is 3 μs. This means that it is highly probable that any signalling rate in the order of megabits per second will cause intersymbol interference.

2.4.4 Time and Frequency Autocorrelation

The delay spread has a close relationship with the frequency response of the mobile radio channel, which will now be statistically characterized. Given that the mobile radio channel impulse response $h_b(\tau,t)$ must be modelled as a time-variant stochastic process, the same applies to the Fourier transform $H_b(f,t)$. Its autocorrelation function is

$$\begin{aligned} R_{HH}(f_1,f_2;\Delta t) &= \frac{1}{2}E\left[H_b^*(f_1;t)H_b(f_2;t+\Delta t)\right] \\ &= \int_{-\infty}^{\infty} R_{hh}(\tau_1;\Delta t)\delta(\tau_1-\tau_2)\, e^{j2\pi(f_1\tau_1 - f_2\tau_2)}d\tau_1 d\tau_2 \\ &= \int_{-\infty}^{\infty} R_{hh}(\tau_1;\Delta t)\, e^{-j2\pi\Delta f\tau_1}d\tau_1 = R_{HH}(\Delta f;\Delta t) \end{aligned} \tag{2.58}$$

showing a Fourier transform relationship with the delay cross-power density $R_{hh}(\tau,\Delta t)$, and implying a frequency difference dependence only, which is the result of the uncorrelated scattering assumption. The spaced-time, spaced-frequency correlation function $R_{HH}(\Delta f;\Delta t)$ outlines two important channel measures: coherence time $(\Delta t)_c$ and coherence bandwidth $(\Delta f)_c$. These characterize the time and frequency distortion (selectivity) of the channel, respectively.

For the channel model in (2.45), $R_{HH}(\Delta f;\Delta t)$ can be derived as follows. The Fourier

transform of the channel impulse response is

$$\begin{aligned} H_b(f,t) &= \int_{-\infty}^{\infty} h_b(\tau,t)e^{-j2\pi f\tau}d\tau \\ &= \upsilon_0\sum_{i=0}^{N-1} c_i e^{j(\varphi_i+2\pi f_D\cos\phi_i t-2\pi f\tau_i)} = I(f;t)+jQ(f;t) \end{aligned} \tag{2.59}$$

where $f_D = \nu/\lambda$ is the maximum *Doppler shift*. Then, the time-frequency correlation of $H_b(f,t)$is given by

$$\begin{aligned} R_{HH}(f_a,f_b;\Delta t) &= \frac{1}{2}E\left[H_b^*(f_a;t)H_b(f_b;t+\Delta t)\right] \\ &= \frac{\upsilon_0^2}{2}\sum_{i=0}^{N-1} E[\,c_i^*c_k\,]E\left[e^{j\left((\varphi_k-\varphi_i)+2\pi f_D(\cos\phi_k(t+\Delta t)-\cos\phi_i t)-2\pi f_b\tau_k+2\pi f_a\tau_i\right)}\right] \end{aligned} \tag{2.60}$$

The random variable $\varphi_i-\varphi_j$ is uniformly distributed in the interval $(0,2\pi)$, except when $i=j$, where it equals 0 with probability 1. Considering the statistical independence of all the random variables $\{c_i\}$, $\{\varphi_i\}$, $\{\phi_i\}$, and $\{\tau_i\}$, this means that taking the expectation in (2.60) over $\varphi_i-\varphi_j$ gives a non-zero value only for $i=j$. This greatly simplifies the result in (2.60) to

$$\begin{aligned} R_{HH}(f_a,f_b;\Delta t) &= \frac{\upsilon_0^2}{2}\sum_{i=0}^{N-1} E\left[c_i^2\right]E\left[e^{j\left(2\pi f_D\Delta t\cos\phi_i+2\pi(f_a-f_b)\tau_i\right)}\right] \\ &= \frac{N\upsilon_0^2}{2}E\left[c_i^2\right]E\left[e^{j\left(2\pi f_D\Delta t\cos\phi+2\pi(f_a-f_b)\tau\right)}\right] \\ &= E\left[e^{j\left(2\pi f_D\Delta t\cos\phi+2\pi\Delta f\tau\right)}\right] = R_{HH}(\Delta f;\Delta t) \end{aligned} \tag{2.61}$$

where, with no loss of generality, it is assumed that $E\left[c_i^2\right]=2/N\upsilon_0^2$. Furthermore, it is easily checked that

$$\begin{aligned} R_{HH}(f_a,f_b;\Delta t) &= E[\,I(f_a;t)I(f_b;t+\Delta t)\,]+jE[\,I(f_a;t)Q(f_b;t+\Delta t)\,] \\ &= R_{II}(f_a,f_b;\Delta t)+jR_{IQ}(f_a,f_b;\Delta t) \end{aligned} \tag{2.62}$$

and that the following relations are valid (they are the consequence of time and frequency wide-sense stationarity)

$$\begin{aligned} E[\,I(f_a;t)I(f_b;t+\Delta t)\,] &= E[\,Q(f_a;t)Q(f_b;t+\Delta t)\,] \\ E[\,Q(f_a;t)I(f_b;t+\Delta t)\,] &= -E[\,I(f_a;t)Q(f_b;t+\Delta t)\,] \end{aligned} \tag{2.63}$$

As expected from the WSSUS channel assumptions, the channel autocorrelation function $R_{HH}(f_a,f_b;t,t+\Delta t)$ only depends on time and frequency separations. To complete the expectation process, the probability density function of ϕ and τ must be specified. The delay was already inferred in (2.57) to be exponentially distributed. Also, if the multipath scatterers are numerous and distributed 360° around the receiver, the waves will arrive from all angles in the azimuth plane with equal probability [29], i.e.

$$p(\phi) = \frac{1}{2\pi} \qquad 0\le\phi\le 2\pi \tag{2.64}$$

Thus, the in-phase correlation function becomes

$$\begin{aligned} R_{II}(\Delta f;\Delta t) &= E\left[\cos\left(2\pi f_D \Delta t\cos\phi + 2\pi\Delta f\tau\right)\right] \\ &= \int_0^{2\pi}\frac{1}{2\pi}\int_0^{\infty}\cos\left(2\pi f_D\Delta t\cos\phi + 2\pi\Delta f\tau\right)\frac{1}{\sigma_\tau}e^{-\tau/\sigma_\tau}d\tau d\phi \end{aligned} \tag{2.65}$$

Performing a double integration by parts in the inner integral, (2.65) may be simplified to

$$\begin{aligned} R_{II}(\Delta f;\Delta t) &= \frac{1}{1+(2\pi\Delta f\sigma_\tau)^2}\left[\frac{1}{2\pi}\int_0^{2\pi}\cos\left(2\pi f_D\Delta t\cos\phi\right)d\phi\right] \\ &= \frac{J_0(2\pi f_D\Delta t)}{1+(2\pi\Delta f\sigma_\tau)^2} \end{aligned} \tag{2.66}$$

where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind. In much the same way, it can be shown that

$$\begin{aligned} R_{IQ}(\Delta f;\Delta t) &= \frac{1}{1+(2\pi\Delta f\sigma_\tau)^2}\left[\frac{2\pi\Delta f\sigma_\tau}{2\pi}\int_0^{2\pi}\cos\left(2\pi f_D\Delta t\cos\phi\right)d\phi\right] \\ &= 2\pi\Delta f\sigma_\tau\frac{J_0(2\pi f_D\Delta t)}{1+(2\pi\Delta f\sigma_\tau)^2} \end{aligned} \tag{2.67}$$

Hence, the spaced-time, spaced-frequency autocorrelation function may be written as

$$\begin{aligned} R_{HH}(\Delta f;\Delta t) &= R_{II}(\Delta f;\Delta t) + jR_{IQ}(\Delta f;\Delta t) \\ &= J_0(2\pi f_D\Delta t)\frac{1+j2\pi\Delta f\sigma_\tau}{1+(2\pi\Delta f\sigma_\tau)^2} \\ &= \frac{J_0(2\pi f_D\Delta t)}{\sqrt{1+(2\pi\Delta f\sigma_\tau)^2}}e^{j\tan^{-1}(2\pi\Delta f\sigma_\tau)} \end{aligned} \tag{2.68}$$

One should notice from (2.68) that for $\Delta t = \Delta f = 0$, $R_{HH}(\Delta f;\Delta t) = 1$, and also that the channel will present a phase deviation for $\Delta f > 0$, as expected. Plots of (2.68) for the two cases $\Delta f = 0$ and $\Delta t = 0$ are shown in Figures 2.4(a) and 2.4(b), respectively. These clear-

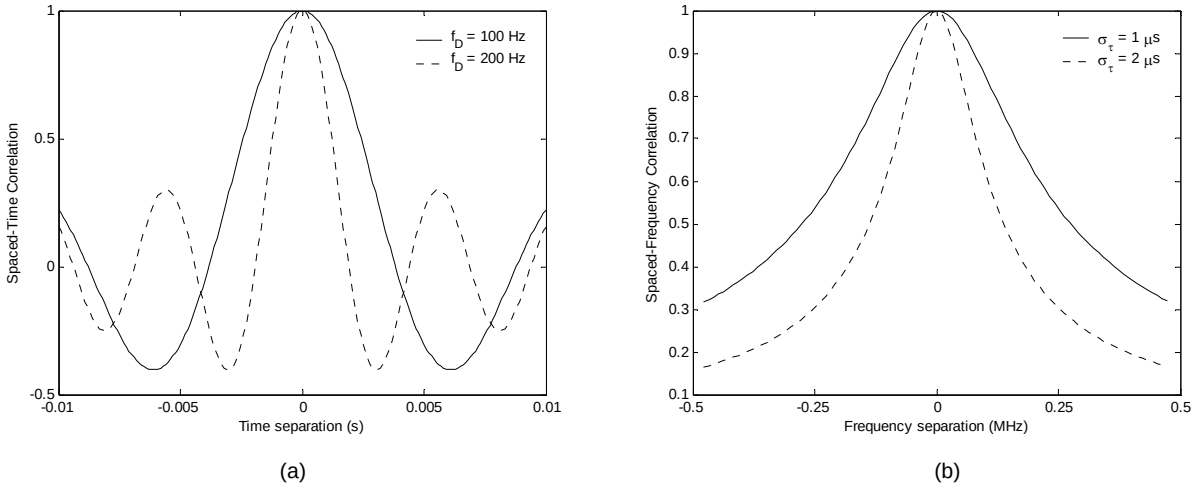

FIGURE 2.4 Spaced-Time (a) and Spaced-Frequency (b) correlation functions.

ly illustrate that, the higher are the Doppler shift and the RMS delay spread, the more selective is the channel in the time and frequency domains, respectively.

To find the *coherence bandwidth* of the channel, one may put $\Delta t = 0$ in (2.68), and use

the -3 dB cutoff criterion

$$|R_{HH}((\Delta f)_c;\Delta t=0)| = \frac{1}{\sqrt{1+(2\pi(\Delta f)_c\sigma_\tau)^2}} = \frac{1}{\sqrt{2}} \tag{2.69}$$

which gives the coherence bandwidth as

$$(\Delta f)_c = \frac{1}{2\pi\sigma_\tau} \tag{2.70}$$

The inverse proportionality between the coherence bandwidth $(\Delta f)_c$ and the RMS delay spread σ_τ was already expected because of the Fourier transform relationship between the delay power density and the spaced-frequency autocorrelation function, as given by (2.58) and rewritten here with $\Delta t = 0$:

$$R_{HH}(\Delta f) = \int_{-\infty}^{\infty} R_{hh}(\tau)\, e^{-j2\pi\Delta f\tau} d\tau \tag{2.71}$$

For the rule of thumb *RMS delay spread* $\sigma_\tau = 3$ μs, the coherence bandwidth is approximately $(\Delta f)_c = 50$ kHz, which means that, besides intersymbol interference, wideband mobile communications will experience frequency distortion of the modulated signal.

Similarly to the coherence bandwidth, the *coherence time* can be found by putting $\Delta f = 0$ in (2.68) and writing

$$|R_{HH}(\Delta f = 0;(\Delta t)_c)| = J_0(2\pi f_D(\Delta t)_c) = \frac{1}{\sqrt{2}} \tag{2.72}$$

which numerically gives

$$(\Delta t)_c \approx \frac{1.126}{2\pi f_D} \approx \frac{9}{16\pi f_D} \approx \frac{1}{2\pi f_D} \tag{2.73}$$

2.4.5 FREQUENCY SELECTIVITY SIMULATION

To confirm the frequency selectivity of the multipath propagation channel, a frequency response numerical simulation based on (2.59) is plotted in Figure 2.5. The simulation parameters are: number of propagation paths $N = 10$, RMS delay spread $\sigma_\tau = 3$ μs, maximum Doppler shift $f_D = 200$ Hz, LOS distance $d = 1$ km, frequency band $B = 5$ MHz, and angle of arrival $\phi_i = (2\pi/N)i$, $i = 0,..,N-1$. For simplicity, the multipath components were assumed equally spaced in their delay and, in addition, their amplitudes were assumed exponentially decaying. One observes that the response is periodic and also that the rate of frequency response variation is necessarily proportional to the Doppler shift of the channel. The time step in the simulation was $\Delta t = 0.167$ ms, which means that for signalling rates of 5 MHz the channel frequency response remains virtually unchanged for a symbol/chip period ($T_c = 0.2$ μs). We also gather that the spacing of two consecutive dips in the frequency response is approximately 0.5 MHz, which is almost equal to the reciprocal of the RMS delay spread.

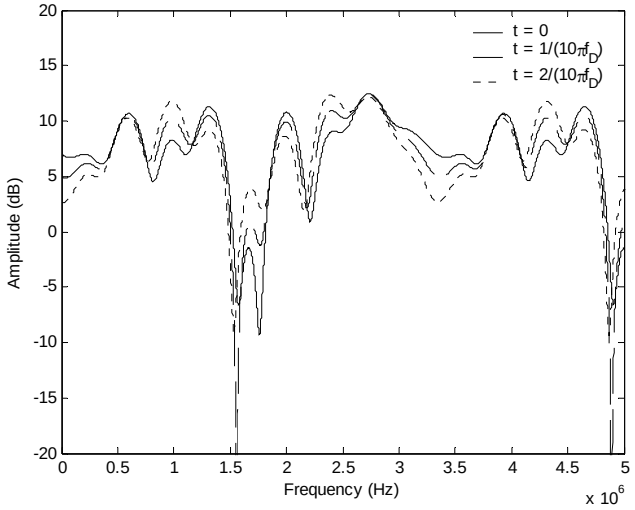

FIGURE 2.5 Simulated frequency response of multipath propagation channel for consecutive time instants, within a 5 MHz bandwidth.

2.4.6 FREQUENCY SELECTIVITY MODELLING

The noiseless received signal from a time-variant mobile propagation channel can be expressed as

$$\begin{aligned} r_b(t) &= \int_{-\infty}^{\infty} h_b(\tau;t)s_b(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} H_b(f;t)S_b(f)e^{j2\pi ft}df \end{aligned} \tag{2.74}$$

where $S_b(f)$ is the baseband Fourier transform of the input signal. If the channel is frequency-selective, i.e. its coherence bandwidth is less than the bandwidth of the baseband input signal, $H_b(f;t)$ cannot be considered constant within the input signal bandwidth, and so it cannot be removed from the integral expression. However, due to the limited bandwidth condition of $S_b(f)$, the output signal $r_b(t)$ should also be bandwidth-limited, meaning that both the input and output channel signals can be represented by their respective samples at the Nyquist rate.

$$\begin{aligned} s_{bs}(t) &= s_b(t)\sum_{n=-\infty}^{\infty}\delta(t-nT_s) \\ r_{bs}(t) &= r_b(t)\sum_{n=-\infty}^{\infty}\delta(t-nT_s) \end{aligned} \tag{2.75}$$

If the symbol rate is R, for less than 100% excess bandwidth pulses (e.g. raised cosine pulses with less than unity roll-off factors), the maximum bandwidth of the input and output signals is $W = R$ and the minimum sample period is $T_s = 1/f_s = 1/2R$. The frequency

spectrum of $s_{bs}(t)$ is given by

$$\begin{aligned} S_{bs}(f) &= S_b(f) * \Im\left\{\sum_{n=-\infty}^{\infty}\delta(t-nT_s)\right\} \\ &= f_s S_b(f) * \sum_{n=-\infty}^{\infty}\delta(f-nf_s) \\ &= f_s\sum_{n=-\infty}^{\infty} S_b(f-nf_s) \end{aligned} \tag{2.76}$$

and the frequency spectrum of $r_{bs}(t)$ is

$$R_{bs}(f) = f_s\sum_{n=-\infty}^{\infty} R_b(f-nf_s) \tag{2.77}$$

Moreover, the samples of the channel response $h_b(\tau;t)$ obtained at the same sampling rate are needed to perform the convolution with the input signal, giving the frequency domain response as

$$R_{bs}(f) = \frac{1}{f_s}H_{bs}(f;t)S_{bs}(f) \tag{2.78}$$

Now, noticing (2.76) and (2.77), it is clear that the channel output can be related to the frequency response of the discrete impulse response $h_{bs}(\tau;t)$ as

$$\begin{aligned} R_b(f) &= \begin{cases} R_{bs}(f)/f_s & |f| \le f_s/2 \\ 0 & |f| > f_s/2 \end{cases} \\ &= \frac{1}{f_s}H_{bs}(f;t)S_b(f) \end{aligned} \tag{2.79}$$

The discrete channel impulse response $h_{bs}(\tau;t)$ is derived from $h_b(\tau;t)$ by partitioning the delay axis τ into equal time delay segments called *excess delay bins* $\tau_i = i\Delta\tau$, where $\Delta\tau = 1/2R = T/2$ is the partition width or time resolution of the multipath delay profile. In effect, the finite bandwidth W of the input signal and the receiver leads to an upper bound for the resolution of structures in the impulse response due to the smearing of the pulses over a width $\Delta\tau = (2W)^{-1}$. Thus, all the multipath components received within the i^{th} partition are assumed to have the same time delay τ_i, and are considered unresolvable, which means that they approximately combine to produce a complex stochastic process. Ultimately, the discrete impulse response is obtained by separating resolvable paths from unresolvable paths, yielding

$$\begin{aligned} h_{bs}(\tau,t) &= \sum_{i=0}^{N-1}\left[\upsilon_0\sum_{k=0}^{L_i-1}c_{ik}e^{j(\varphi_{ik}+\omega_{ik}t)}\right]\delta(\tau-\tau_0-iT_s) \\ &= \sum_{i=0}^{N-1}v_i(t)\delta(\tau-\tau_0-iT_s) \end{aligned} \tag{2.80}$$

where $v_i(t)$ is a complex-valued wide-sense stationary random process. The Fourier trans-

form of (2.80) is

$$H_{bs}(f,t) = \sum_{i=0}^{N-1} v_i(t) e^{-j2\pi f(\tau_0 + iT_s)} \tag{2.81}$$

which, after direct substitution into (2.79), and taking the inverse Fourier transform, yields the baseband response

$$\begin{aligned} r_b(t) &= \int_{-\infty}^{\infty} R_b(f) e^{j2\pi ft} df \\ &= \frac{1}{f_s} \int_{-\infty}^{\infty} H_{bs}(f;t) S_b(f) e^{j2\pi ft} df \\ &= \sum_{i=0}^{N-1} \frac{1}{f_s} v_i(t) \int_{-\infty}^{\infty} S_b(f) e^{j2\pi f(t-\tau_0 - iT_s)} df \\ &= \sum_{i=0}^{N-1} h_i(t) s_b(t - \tau_0 - iT_s) \end{aligned} \tag{2.82}$$

where

$$h_i(t) = \frac{v_0}{f_s} \sum_{k=0}^{L_i - 1} c_{ik} e^{j(\varphi_{ik} + \omega_{ik} t)} \tag{2.83}$$

From (2.82) it is clear that the baseband output signal of the frequency selective mobile propagation channel is a finite convolution sum of the input signal with the discrete channel impulse response. A tapped delay line model of (2.83) is shown in Figure 2.6, which is particularly useful for channel numerical simulations, provided that the stochastic tap weights $h_i(t)$ are correctly modelled.

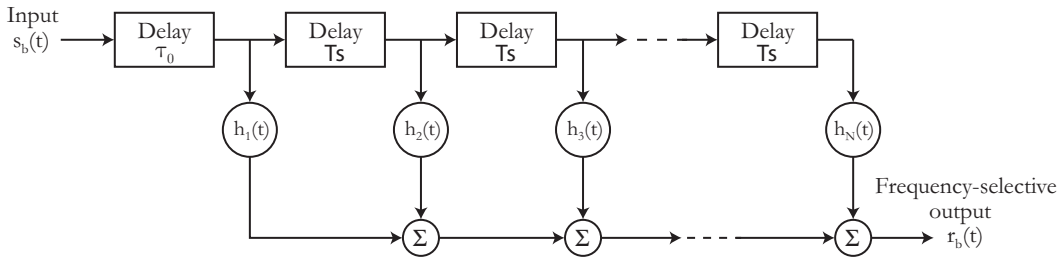

FIGURE 2.6 Tapped delay line model for a noiseless frequency-selective multipath channel.

Short-term channel variations may differ for each delay bin τ_i, and may be ideally characterized by the WSSUS model

$$R_{hh}(\tau_i, \tau_j; \Delta t) = \frac{1}{2} E\left[h_i^*(t) h_j(t + \Delta t)\right] \delta(\tau_i - \tau_j) \tag{2.84}$$

which is a modified version of (2.48) to include the tap fading processes. Also, the time autocorrelation function of each tap fading process is the same of (2.68) with $\Delta f = 0$ for uniformly distributed angles-of-arrival (AOA).

2.4.7 STATISTICAL FADING CHARACTERIZATION

The statistical characterization of the $h_i(t)$'s can be done by assuming L_i is very large, and

using the central limit theorem to estimate the probability distribution. Equation (2.83) can be rewritten as

$$\begin{aligned} h_i(t) &= \sum_{k=0}^{L_i-1} a_{ik}\cos(\varphi_{ik} + \omega_{ik}t) + j\sum_{k=0}^{L_i-1} a_{ik}\sin(\varphi_{ik} + \omega_{ik}t) \\ &= X(t) + jY(t) \\ &= \sqrt{X^2(t) + Y^2(t)}e^{j\arctan(Y(t)/X(t))} \end{aligned} \tag{2.85}$$

Now, since a_{ik}, φ_{ik} and ω_{ik} are considered statistically independent random variables and the phases φ_{ik} are uniformly distributed in the interval $(0,2\pi)$, $E[X(t)] = E[Y(t)] = 0$, and by the central limit theorem both $X(t)$ and $Y(t)$ are zero mean Gaussian random variables, then

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x}e^{-x^2/2\sigma_x^2} \tag{2.86}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_y}e^{-y^2/2\sigma_y^2} \tag{2.87}$$

Moreover, since $E[X^2(t)] = E[Y^2(t)]$, they have equal variance $\sigma^2 = \sigma_x^2 = \sigma_y^2$, and because $E[X(t)Y(t)] = 0$, they are also uncorrelated. The Gaussian assumption and the uncorrelatedness are sufficient conditions for the statistical independence of $X(t)$ and $Y(t)$, and thus their joint density is

$$\begin{aligned} f_{XY}(x,y) &= f_X(x)f_Y(y) \\ &= \frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \end{aligned} \tag{2.88}$$

If the amplitude and phase are

$$R(t) = \sqrt{X^2(t) + Y^2(t)} \qquad \theta(t) = \arctan(Y(t)/X(t)) \tag{2.89}$$

then a solution pair is

$$x(t) = r(t)\cos\theta(t) \qquad y(t) = r(t)\sin\theta(t) \tag{2.90}$$

and the joint distribution of $R(t)$ and $\theta(t)$ is

$$\begin{aligned} f_{R\theta}(r,\theta) &= f_{XY}(x,y)|J(r,\theta)| \\ &= f_{XY}(r\cos\theta, r\sin\theta)\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \frac{r}{2\pi\sigma^2}e^{-r^2/2\sigma^2} \end{aligned} \tag{2.91}$$

The envelope distribution is obtained by integrating with respect to θ, giving

$$f_R(r) = \frac{r}{\sigma^2}e^{-r^2/2\sigma^2} \tag{2.92}$$

In conclusion, the envelope of the stochastic tap values $h_i(t)$ is Rayleigh distributed for any fixed value of t and the phase is uniformly distributed in the interval $(0,2\pi)$. The envelope statistic characterizes the small-scale fading behaviour of the mobile channel signal when

a receiver moves over small distances as compared to the wavelength.

The Rayleigh distribution obtained is valid only when the multipath phenomenon occurs mainly because of local scatterers around the receiver (e.g. urban areas), meaning that the zero mean Gaussian distribution of $X(t)$ and $Y(t)$ is an acceptable assumption, as shown in Figure 2.7 a).Whenever there is a LOS multipath component reaching the re-

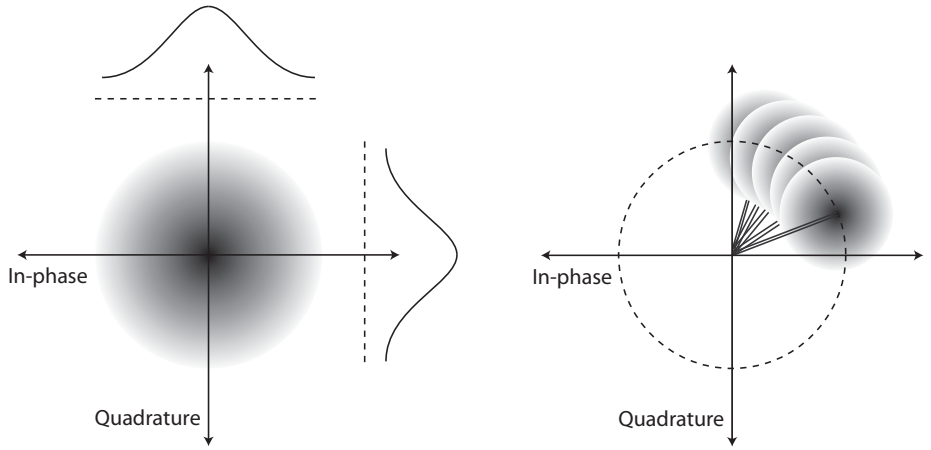

FIGURE 2.7 Possible probability density functions for local area signal fading characterization.

ceiver, or some dominant specular reflections of the signal are present, the zero mean assumption of $X(t)$ and $Y(t)$ is no longer valid because the signal strength of the scattered multipath components will be much lower than the former's. In this case, the complex multipath process can be assumed a nonzero mean Gaussian process, as depicted in Figure 2.7 b), and with joint pdf given by

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-m_x)^2+(y-m_y)^2}{2\sigma^2}} \tag{2.93}$$

where m_x and m_y are the in-phase and quadrature specular component amplitudes. Following the same procedure from (2.88) to (2.91), it follows that

$$\begin{aligned} f_{R\theta}(r,\theta) &= \frac{r}{2\pi\sigma^2} e^{-\frac{(r\cos\theta-m_x)^2+(r\sin\theta-m_y)^2}{2\sigma^2}} \\ &= \frac{r}{2\pi\sigma^2} e^{-\left(r^2+A^2\right)/2\sigma^2} e^{rA\cos(\theta-\theta_0)/\sigma^2} \end{aligned} \tag{2.94}$$

where $A = \sqrt{m_x^2 + m_y^2}$ and $\theta_0 = \arctan(m_y / m_x)$. Thus

$$\begin{aligned} f_R(r) &= \int_0^{2\pi} \frac{r}{2\pi\sigma^2} e^{-\left(r^2+A^2\right)/2\sigma^2} e^{rA\cos(\theta-\theta_0)/\sigma^2} d\theta \\ &= \frac{r}{\sigma^2} e^{-\left(r^2+A^2\right)/2\sigma^2} \left[\frac{1}{2\pi}\int_0^{2\pi} e^{rA\cos\theta/\sigma^2} d\theta\right] \\ &= \frac{r}{\sigma^2} e^{-\left(r^2+A^2\right)/2\sigma^2} I_0\left(\frac{rA}{\sigma^2}\right) \end{aligned} \tag{2.95}$$

where $I_0(\cdot)$ is the zeroth order modified Bessel function of the first kind. The pdf of the envelope given by (2.95) is called the Rice distribution [30]. When plotted, the Rician pdf confirms a less severe fading effect than the Rayleigh pdf, making it suitable for the char-

acterization of rural or sub-urban areas.

While the Rayleigh and Rice distributions can indeed be used to model the envelope of fading channels in many cases of interest, it has been found experimentally that the Nakagami-m distribution offers a better fit for a wider range of fading conditions (see [25], [31], [32], [33], and [34]. The Rice distribution can only describe better-than-Rayleigh fading conditions, whereas the Nakagami-m distribution can describe Rayleigh, Rice and worst-than-Rayleigh fading conditions.

A roughly generic form for the expression of the fading tap weights $h_i(t)$ is obtained when the scattering process at time delay τ_i is assumed to be the sum of the contributions of several clusters of scatterers, each cluster approximately producing an independent complex zero-mean Gaussian process. If this is the case, X and Y in (2.85) can be written as

$$\begin{aligned} X &= \sum_{k=0}^{m-1} X_k = \sum_{k=0}^{m-1} \alpha_k \cos\theta_k \\ Y &= \sum_{k=0}^{m-1} Y_k = \sum_{k=0}^{m-1} \alpha_k \sin\theta_k \end{aligned} \tag{2.96}$$

where for simplicity of presentation the time dependency of the random variables has been omitted. The joint probability density function of the X_k's and the Y_k's is given by

$$\begin{aligned} f_{..X_k..Y_k..}(..,x_k,..,y_k,..) &= \prod_{k=0}^{m-1} \frac{1}{2\pi\sigma^2} e^{-(x_k^2+y_k^2)/2\sigma^2} \\ &= \frac{1}{(2\pi)^m \sigma^{2m}} e^{-\sum_{k=0}^{m-1}(x_k^2+y_k^2)/2\sigma^2} \end{aligned} \tag{2.97}$$

The squared envelope of the random fading process can be written as

$$\begin{aligned} R^2 &= X^2 + Y^2 \\ &= \left(\sum_{k=0}^{m-1} X_k\right)^2 + \left(\sum_{k=0}^{m-1} Y_k\right)^2 \\ &= \sum_{k=0}^{m-1} \left(X_k^2 + Y_k^2\right) + \sum_{k=0}^{m-1} \sum_{\substack{l=0 \\ l\neq k}}^{m-1} (X_k X_l + Y_k Y_l) \end{aligned} \tag{2.98}$$

Due to the independence of the scatterers and, as a result, the uncorrelatedness of the X_k's and the Y_k's, the second term in (2.98) is negligible as compared with the first term, and so a reasonable approximation is to equate it to zero, and thus

$$R^2 = \sum_{k=0}^{m-1} \left(X_k^2 + Y_k^2\right) = \sum_{k=0}^{m-1} \alpha_k^2 \tag{2.99}$$

Furthermore, expressing the joint probability density function in 2m-dimensional hyper-

spherical coordinates R, $\theta_0..\theta_{2m-2}$, where

$$\begin{aligned} X_k &= R\cos\theta_{2k}\prod_{l=0}^{2k-1}\sin\theta_l \qquad k=0,..,m-1 \\ Y_k &= R\cos\theta_{2k+1}\prod_{l=0}^{2k}\sin\theta_l \qquad k=0,..,m-2 \\ Y_{m-1} &= R\prod_{l=0}^{2m-1}\sin\theta_l \end{aligned} \tag{2.100}$$

for $0 \le \theta_l \le \pi$, yields

$$f_{R\theta_0..\theta_{2m-2}}(r,\theta_0,..,\theta_{m-2}) = \frac{1}{(2\pi)^m\sigma^{2m}}e^{-r^2/2\sigma^2}\,|J(r,\theta_0,..,\theta_{2m-2})| \tag{2.101}$$

Now, since the vector differential length in 2m-dimensional hyperspherical coordinates is given by

$$d\boldsymbol{l} = dr\,\mathbf{a}_{\mathbf{r}} + rd\theta_0\,\mathbf{a}_{\theta_0} + r\sin\theta_0 d\theta_1\,\mathbf{a}_{\theta_1} + ... + r\left[\prod_{k=0}^{2m-3}\sin\theta_k\right]d\theta_{2m-2}\,\mathbf{a}_{\theta_{2m-2}} \tag{2.102}$$

the Jacobian in (2.101) may be written as the product of all differential lengths

$$J(r,\theta_0,..,\theta_{2m-2}) = r^{2m-1}\prod_{k=0}^{2m-3}(\sin\theta_k)^{2m-2-k} \tag{2.103}$$

which gives the envelope probability density function as

$$f_R(r) = \frac{r^{2m-1}}{(2\pi)^m\sigma^{2m}}e^{-r^2/2\sigma^2}\int_{\theta_0..\theta_{2m-2}}\prod_{k=0}^{2m-3}(\sin\theta_k)^{2m-2-k}\,d\theta_0..d\theta_{2m-2} \tag{2.104}$$

The integral in (2.104) is over all solid angles subtended by an hypersphere of unit radius in 2m-dimensional space, giving its total hyper-surface area. To evaluate it, write

$$\int_0^\infty f_R(r)dr = S_{2m}\int_0^\infty \frac{r^{2m-1}}{(2\pi)^m\sigma^{2m}}e^{-r^2/2\sigma^2}\,dr = 1 \tag{2.105}$$

and make the change of variables $t = r^2/2\sigma^2$, yielding

$$S_{2m}\int_0^\infty \frac{t^{m-1}}{2\pi^m}e^{-t}\,dt = 1 \tag{2.106}$$

Defining the gamma function as

$$\Gamma(m) = \int_0^\infty t^{m-1}e^{-t}\,dt \tag{2.107}$$

the area of the hypersphere in 2m-dimensional space is

$$S_{2m} = \frac{2\pi^m}{\Gamma(m)} \tag{2.108}$$

Substitution into (2.104) gives

$$f_R(r) = \frac{2r^{2m-1}}{(2\sigma^2)^m \Gamma(m)} e^{-r^2/2\sigma^2} \tag{2.109}$$

and defining $\Omega = E[R^2] = 2m\sigma^2$, (2.109) may be expressed as

$$f_R(r) = \frac{2r^{2m-1}}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m e^{-mr^2/\Omega} \tag{2.110}$$

The envelope probability density function in (2.110) is called the *Nakagami-m distribution*, and can be used to statistically characterize the fading processes that appear in the frequency-flat or frequency-selective mobile radio propagation channel. An alternative approach to the derivation of the above expression based on the physical properties of the radio channel and on the Hankel transformation of radially symmetric pdfs appears in [35]. Figure 2.8 illustrates the Nakagami-m pdf for different values of m. Varying the *fading figure* m, various channel fading conditions can be statistically emulated. For instance, when $m = 1$ (a single scattering cluster) the Nakagami-m distribution is identical to the Rayleigh distribution, and when $m > 1$ it approaches the Rice distribution (i.e. when strong specular components appear at the receiver). For $1/2 < m < 1$ the Nakagami-m distribution can model worst-than-Rayleigh scenarios that may arise when the X_k's and the Y_k's in (2.96) cannot be considered Gaussian uncorrelated random variables.

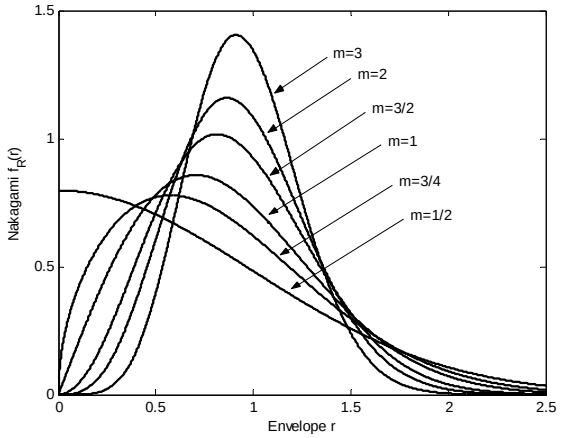

FIGURE 2.8 Nakagami-m pdf for several m parameters, shown with $\Omega = 1$.

2.4.8 TIME SELECTIVITY SIMULATION

In the absence of specular signal components, and when the delay spread of the channel is small compared to the symbol period ($\sigma_\tau \ll T$ or, i.e., $N = 1$ in (2.82)), the channel will produce a frequency-flat signal fading output. Otherwise, if $\sigma_\tau \gg T$, the channel output will be the sum of several time shifted independent stochastic processes, producing frequency distortion and small-scale fading at the same time. A simulation of (2.83) is shown in Figure 2.9 assuming that the multipath components' amplitudes c_{ik} are equal, the phases φ_{ik} are uniformly distributed in the interval $(0, 2\pi)$, $L_i = 10$, and $\omega_{ik} = 2\pi f_D \cos\phi_{ik}$, where $f_D = 200$ Hz and $\phi_{ik} = k(2\pi / L_i)$, $k = 0, .., L_i - 1$. The carrier frequency is $f_c = 2$ GHz.

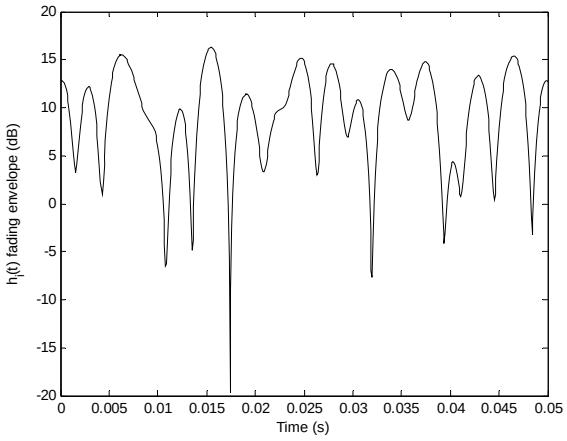

FIGURE 2.9 Rayleigh fading behaviour of the time varying stochastic tap weights $\{h_i(t)\}$ for a frequency-selective mobile propagation channel.

Since the multipath phenomenon is slowly fading for wideband communications ($B > 1$ MHz), the deep fades in Figure 2.9 occur at bursts, that is, many symbols apart, reducing the instantaneous signal-to-noise ratio at the receiver. Consequently, the error performance of a communications system is severely degraded in multipath fading environments. The frequency domain manifestation of signal fading is the time variation of the frequency response, as already depicted in Figure 2.5.

2.4.9 TIME SELECTIVITY AND DOPPLER SPECTRUM

A spectral "picture" of the time selectivity of the channel, or in other words, the time variation of channel due to the receiver motion relative to the transmitter, can be obtained by taking the Fourier transform of (2.68) with respect to the Δt variable, when $\Delta f = 0$, that is

$$\begin{aligned} S_{HH}(\kappa) &= \int_{-\infty}^{\infty} J_0(2\pi f_D \Delta t) e^{-j2\pi\kappa\Delta t} d\Delta t \\ &= \int_{-\infty}^{\infty} \frac{1}{\pi} \int_0^{\pi} e^{j2\pi f_D \Delta t \cos\phi} e^{-j2\pi\kappa\Delta t} d\phi \, d\Delta t \\ &= \frac{1}{\pi} \int_0^{\pi} \int_{-\infty}^{\infty} e^{j2\pi(f_D\cos\phi - \kappa)\Delta t} d\Delta t \, d\phi \\ &= \frac{1}{\pi} \int_0^{\pi} \delta(\kappa - f_D \cos\phi) \, d\phi \end{aligned} \tag{2.111}$$

Now, since $\cos\phi$ is injective in the interval $0 \le \phi \le \pi$, the change of variable $x = f_D \cos\phi$ is applicable, and (2.111) can thus be written

$$\begin{aligned} S_{HH}(\kappa) &= \int_{-f_D}^{f_D} \frac{1}{\pi\sqrt{f_D^2 - x^2}} \delta(\kappa - x) \, dx \\ &= \frac{1}{\pi\sqrt{f_D^2 - \kappa^2}} \end{aligned} \tag{2.112}$$

Equation (2.112) is called the wavenumber spectrum or *Doppler power spectrum* of the mobile channel, and has an U-shaped appearance as shown in Figure 2.10. It characterizes the frequency shift (spectral broadening) of the waves as perceived at the receiver due to its motion relative to the transmitter and, as expected, shows that the multipath components

that contribute the most to the frequency shift are the ones arriving at 0º and 180º relative to the receiver velocity vector.

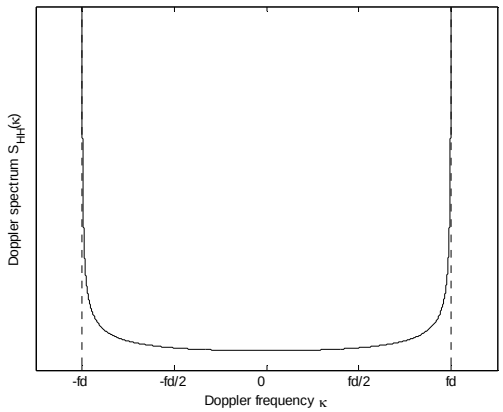

FIGURE 2.10 Typical *Doppler power spectrum* for mobile radio channels in cluttered outdoor environments.

Although the Doppler spectrum in Figure 2.10 is an experimentally confirmed, well accepted model for the mobile radio propagation channel, it is not the better model for all channel conditions. For instance, if both the transmitter and receiver are standing still, but the propagation environment is densely occupied with randomly moving scatterers, namely wind-blown leaves or random street traffic, measurements confirm that the Doppler power spectrum is concentrated heavily at $\kappa = 0$, and quickly diminishes for larger values of $|\kappa|$, approximating a Laplacian-shaped function as illustrated in Figure 2.11 [36] (cf. [37]). Thus, in practice a receiver will perceive a Doppler spectrum which is a "mixture" of Figure 2.10 with Figure 2.11.

Another time-varying channel example is the tropospheric scatter channel, which has the Doppler spectrum approximated by a Gaussian-shaped function

$$S_{HH}(\kappa) = S_0 e^{-\kappa^2/2\sigma_\kappa^2} \tag{2.113}$$

where S_0 is some arbitrary constant and σ_κ is the *RMS Doppler spread*.

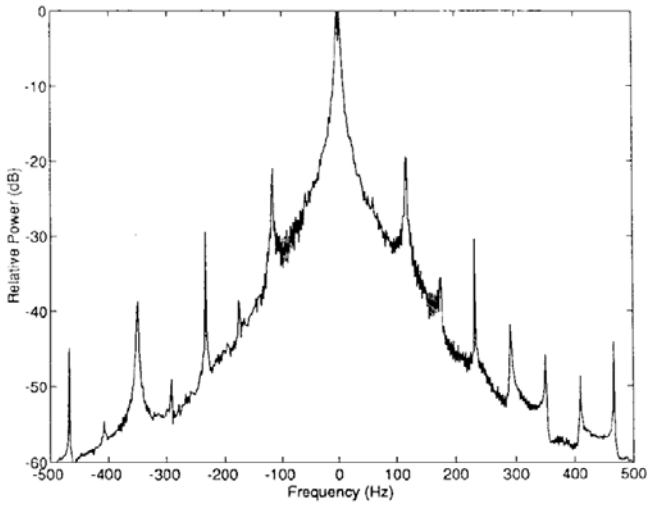

FIGURE 2.11 Average Doppler spectrum for a fixed link and dynamic environment (e.g. traffic road) at 40 GHz.

2.4.10 BELLO FUNCTIONS AND RELATIONS

Taking the Fourier transform of $R_{hh}(\tau, \Delta t)$ in the Δt variable

$$S_{HH}(\tau, \kappa) = \int_{-\infty}^{\infty} R_{hh}(\tau, \Delta t)\, e^{-j2\pi\kappa\Delta t} d\Delta t \tag{2.114}$$

gives the average power output of the channel as a function of the time delay τ and the Doppler frequency κ. $S_{HH}(\tau, \kappa)$ is called the *scattering function of the channel*, and is particularly useful in obtaining the power delay profile and Doppler spectrum functions by a direct integration

$$\begin{aligned} R_{hh}(\tau) &= \int_{-\infty}^{\infty} S_{HH}(\tau, \kappa)\, d\kappa \\ S_{HH}(\kappa) &= \int_{-\infty}^{\infty} S_{HH}(\tau, \kappa)\, d\tau \end{aligned} \tag{2.115}$$

The relations among the correlation and power spectra functions described so far constitute the *Bello relations* [38] for time-varying channels, and are shown in Figure 2.12. The functions are called the *Bello system functions* of the WSSUS channel, and are related by a loop of Fourier transform pairs in the Δt, τ, Δf and κ variables.

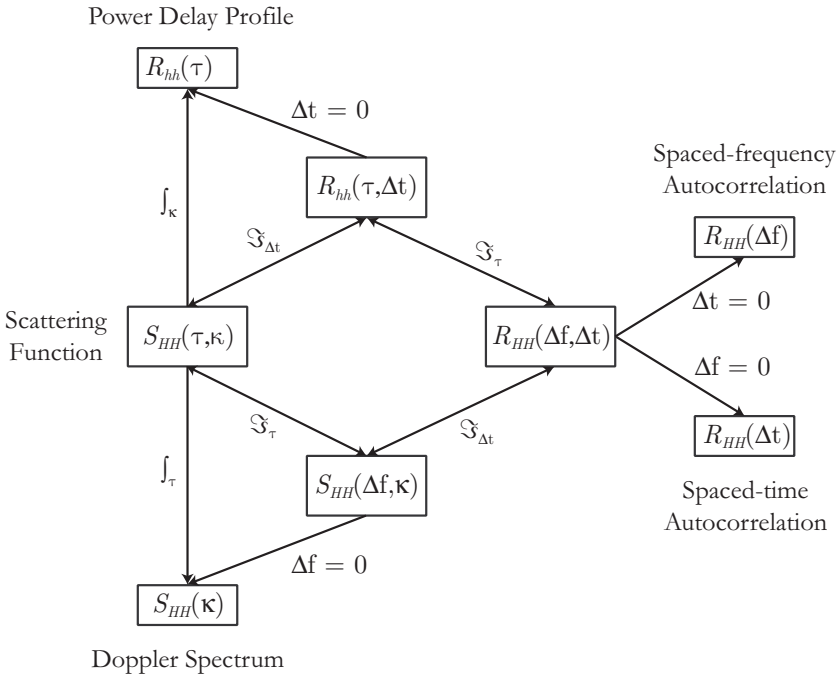

FIGURE 2.12 Relations among the correlation and power spectra functions for the WSSUS channel model.

To complete the picture, one can easily deduce from the wide-sense stationarity property of the channel in the time and frequency domains that 1)

$$\begin{aligned} R_{HH}(\Delta f, \Delta t) &= R_{HH}(f, f', t, t') = R_{HH}(f' - f, t' - t) \\ &= \int_{\tau}\int_{\kappa} S_{HH}(\tau, \kappa) e^{-j2\pi(f'-f)\tau} e^{j2\pi\kappa(t'-t)} d\nu d\tau \\ &= \int_{\tau}\int_{\tau'}\int_{\kappa}\int_{\kappa'} S_{HH}(\tau, \kappa)\delta(\tau' - \tau)\delta(\kappa' - \kappa) e^{j2\pi(-f'\tau' + \kappa' t' + f\tau - \kappa t)} d\kappa' d\kappa d\tau' d\tau \end{aligned} \tag{2.116}$$

and also 2)

$$\begin{aligned} R_{HH}(f,f',t,t') &= \frac{1}{2}E[H(f',t')H^*(f,t)] \\ &= \int_{\tau}\int_{\tau'}\int_{\kappa}\int_{\kappa'}\frac{1}{2}E[H(\tau',\kappa')H^*(\tau,\kappa)]e^{j2\pi(-f'\tau'+\kappa't'+f\tau-\kappa t)}d\kappa' d\kappa d\tau' d\tau \\ &= \int_{\tau}\int_{\tau'}\int_{\kappa}\int_{\kappa'}R_{HH}(\tau,\tau',\kappa,\kappa')e^{j2\pi(-f'\tau'+\kappa't'+f\tau-\kappa t)}d\kappa' d\kappa d\tau' d\tau \end{aligned} \tag{2.117}$$

which must imply that

$$R_{HH}(\tau,\tau',\kappa,\kappa') = R_{HH}(\tau,\kappa)\delta(\tau'-\tau)\delta(\kappa'-\kappa) \tag{2.118}$$

showing that the uncorrelated scattering in a WSSUS channel is present in both delay and Doppler domains.

CHAPTER 3

DIVERSITY RECEPTION

3.1 *Introduction*

It was shown in Section 2.4 that, in a typical urban or sub-urban propagation environment, the reflection and scattering of electromagnetic waves is responsible for signal fading at the receiver. The severity of the fading process will depend on the operating frequency, the data signalling rate, and the statistical properties of the radio channel. For broadband communications (high signalling rates) the fading is typically slow and frequency selective, that is, it results from several delay-spaced fading processes (Figure 2.6). Depending on the channel conditions, each fading process can be statistically characterized by a Rayleigh, Rice or, more generically, by a Nakagami-m probability distribution.

The error performance of wireless communication systems is harshly degraded by signal fading conditions, encouraging the employment of efficient techniques to overcome the problem. Diversity reception techniques are based on the fact that the transmitted signal has several degrees of freedom, namely time, frequency and space. In fact, the way in which the multipath components add at the receiver should be nearly independent from time to time, frequency to frequency and spatial position to spatial position, as long as enough separation is provided. Indeed, there is an intuitive fundamental premise in spatial statistics that states that nearby things are in average more alike than remote things, and it finds applications in geostatistics, ecology (e.g. biodiversity), archeology and even agriculture. As a result, it may seem reasonable to foresee that the fading statistics associated with a multipath propagation channel will also be nearly independent. For instance, in the case of spatial diversity, if $R_1 = R(t)$ and $R_2 = R(t + \Delta t)$ are the sampled envelopes of the received signal at two different receiver positions separated by $d = v\Delta t$, then the probability

that both signal envelopes will be below a certain threshold A is

$$\begin{aligned} P_A &= P(R_1 < A; R_2 < A) \\ &= P(R_2 < A \mid R_1 < A)P(R_1 < A) \end{aligned} \tag{3.1}$$

and the independence condition leads to $P_A = P^2(R_1 < A)$. For L independent fading signals the joint probability is further reduced to $P_A = P^L(R_1 < A)$. A simple method to obtain several replicas of the transmitted signal is to place several antennas at the receiver, and if they are spaced sufficiently far apart so that their received signals fade independently, then they can be used for diversity reception. Furthermore, since the uncorrelatedness of the fading signals is a manifestation of their independence, the autocorrelation function may be used to find the necessary spatial separation of the antennas.

3.2 *Envelope Autocorrelation*

Because time separation can be easily converted to space separation and it is easier to manipulate the square of the envelope, it is appropriate to start with the spaced-time autocorrelation function of the squared envelope, that is

$$\begin{aligned} \phi_{R^2}(\Delta t) &= E\left[R^2(t)R^2(t+\Delta t)\right] \\ &= 2E\left[X^2(t)X^2(t+\Delta t)\right] + 2E^2\left[X^2(t)\right] \end{aligned} \tag{3.2}$$

where $X(t)$ and $Y(t)$ are given in (2.85) and

$$E[X(t)Y(t+\Delta t)] = E[Y(t)X(t+\Delta t)] = 0 \tag{3.3}$$

for uniform distribution of wave arrival. Now, assuming for simplicity that both $X(t)$ and $Y(t)$ are zero-mean, Gaussian distributed random variables with variance σ^2, and using the equality [39]

$$E\left[X_1^2 X_2^2\right] = E\left[X_1^2\right]E\left[X_2^2\right] + 2E^2\left[X_1 X_2\right] \tag{3.4}$$

Equation (3.2) simplifies to

$$\begin{aligned} \phi_{R^2}(\Delta t) &= 2E\left[X^2(t)\right]E\left[X^2(t+\Delta t)\right] \\ &\quad + 4E^2\left[X(t)X(t+\Delta t)\right] + 2E^2\left[X^2(t)\right] \\ &= 4E^2\left[X^2(t)\right] + 4E^2\left[X(t)X(t+\Delta t)\right] \end{aligned} \tag{3.5}$$

The second term in (3.5) expands to

$$\begin{aligned} &E[X(t)X(t+\Delta t)] = \\ &= \sum_{i=0}^{L-1}\sum_{k=0}^{L-1} E[a_i a_k]E\left[\cos(\varphi_k + \omega_k t)\cos(\varphi_i + \omega_i(t+\Delta t))\right] \\ &= \frac{L}{2}E\left[a_i^2\right]E[\cos(2\pi v\cos\phi_i \Delta t/\lambda)] \\ &= E\left[X^2(t)\right]J_0(2\pi v\Delta t/\lambda) \end{aligned} \tag{3.6}$$

where $LE\left[a_i^2\right]$ is the power of the received signal's envelope. The spaced-time autocorre-

lation function of the squared envelope is finally

$$\phi_{R^2}(\Delta t) = 4E^2\left[X^2(t)\right]\left[1 + J_0^2(2\pi v\Delta t\,/\,\lambda)\right] \tag{3.7}$$

The correlation coefficient is given by

$$\rho_{R^2} = \frac{\phi_{R^2}(\Delta t) - E^2\left[R^2(t)\right]}{\phi_{R^2}(0) - E^2\left[R^2(t)\right]} \tag{3.8}$$

Substitution of (3.7) into (3.8) gives

$$\begin{aligned}\rho_{R^2} &= J_0^2(2\pi v\Delta t\,/\,\lambda)\\ &= J_0^2(\beta d)\end{aligned} \tag{3.9}$$

where β is the phase constant and d is the antenna separation distance. This simple equation gives the spatial autocovariance between the squared envelopes of the signals received by the antennas. It may also be derived that the correlation coefficient of the envelope is approximated by [22]

$$\rho_R \approx J_0^2(\beta d) \tag{3.10}$$

The signal envelopes will be uncorrelated when the correlation coefficient becomes zero, and it can be found numerically that the first zero of the Bessel function occurs for $d = 0.38\lambda$. However, due to the lack of uniform distribution of wave arrival (e.g. sub-urban areas), and the possibility of specular components arriving at the receiver, measurements show that the first null is at about $d = 0.8\lambda$. For mobile units located in urban/sub-urban areas a good rule of thumb for the antenna separation distance is usually $d = 0.5\lambda$.

The values just estimated often do not apply to base station receivers due to the reduced angular spread of the incoming waves, which introduces extra correlation between the antennas. As illustrated in Figure 3.1, on account of base stations being typically located in isolated places (above the clutter), the angles of wave arrival span only a small fraction of 360º, and hence higher antenna separations are required so that the propagation differences between the multipath components becomes significant. Moreover, the correlation also depends on the angle α and the height of the base station antennas. To achieve low correlation between base station antennas, separations as high as 10-20λ are usually required [40].

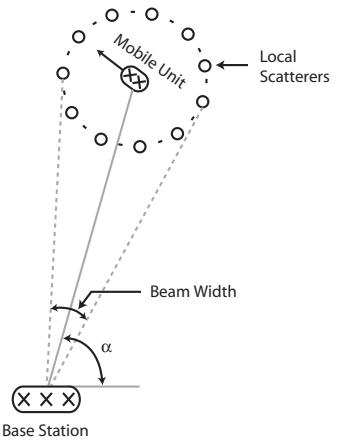

FIGURE 3.1 Illustration of the angular spread of multipath components at a base station.

3.3 *The Maximal-Ratio Combiner*

When multiple antennas are employed at the receiver side, the original single-input single-output (SISO) channel is converted to a single-input multiple-output (SIMO) channel. This type of channel makes several signal replicas available to a receiver *signal processor* which may combine them in order to strengthen the output signal. The optimum combining technique depends on the fading characteristics of the input signal replicas but, as we shall see, whenever it is possible by some procedure to convert the type of fading to flat-fading, then the *maximal-ratio combiner* (MRC) is the optimum combining scheme for maximum output SNR.

Let $h_k(t)$ represent a baseband-equivalent flat-fading process associated with the channel response between the transmitter antenna and the k-th receiver antenna, and let it be expressed as

$$\begin{aligned} h_k(t) &= \sum_{i=0}^{L_k-1} a_{ik}\, e^{j(\varphi_{ik}+\omega_{ik}t)} \\ &= \alpha_k(t) e^{j\theta_k(t)} \end{aligned} \tag{3.11}$$

Then, the output signal from the k-th antenna may be written as

$$\begin{aligned} r_{bk}(t) &= h_k(t) s_b(t-\tau_0) + n_k(t) \\ &= \alpha_k(t) e^{j\theta_k(t)} s_b(t-\tau_0) + n_k(t) \end{aligned} \tag{3.12}$$

where $n_k(t) = n_{kI}(t) + jn_{kQ}(t)$ is a complex-valued zero-mean Gaussian noise process. The diversity combiner weighs and sums the signals from all the antennas, producing the output signal

$$\begin{aligned} r_b(t) &= \sum_{k=0}^{M-1} w_k(t) r_{bk}(t) \\ &= \sum_{k=0}^{M-1} w_k(t)\alpha_k(t) e^{j\theta_k(t)} s_b(t-\tau_0) + \sum_{k=0}^{M-1} w_k(t) n_k(t) \end{aligned} \tag{3.13}$$

where M is the number of receiver antennas (or *diversity order*), and the w_k's are the combiner weights. This technique is also a type of *predetection combining*, meaning that the detection process for $s_b(t-\tau_0)$ will occur after the branch combining. The performance of *predetection combining* and *postdetection combining* is identical when coherent detection is used, but for nonlinear detectors (e.g. square-law detectors) predetection is the better combining scheme since it provides the maximum output SNR [41]. Now, since the noise processes from different antennas are uncorrelated, the average noise power at the output of the combiner is given by

$$\begin{aligned} \eta &= \sum_{k=0}^{M-1} |w_k(t)|^2 \frac{1}{2} E\left[|n_k(t)|^2\right] \\ &= \sum_{k=0}^{M-1} |w_k(t)|^2 \eta_k \end{aligned} \tag{3.14}$$

where $\eta_k = \mathcal{N}_0 B$ is the total noise power in each branch before detection and B is the

bandwidth occupied by the signal. Furthermore, the instantaneous power of the output signal may be written as

$$\begin{aligned} S_t &= \frac{1}{2}\left|\sum_{k=0}^{M-1} w_k(t)h_k(t)s_b(t-\tau_0)\right|^2 \\ &\leq \frac{1}{2}\left(\sum_{k=0}^{M-1}|w_k(t)|^2\,\eta_k\right)\left(\sum_{k=0}^{M-1}\left|\alpha_k(t)s_b(t-\tau_0)\right|^2/\eta_k\right) \end{aligned} \tag{3.15}$$

where the last inequality follows from Cauchy-Schwarz. Consequently, the instantaneous output SNR is

$$\gamma = \frac{S_t}{\eta} \leq \frac{1}{2}\sum_{k=0}^{M-1}\left|\alpha_k(t)s_b(t-\tau_0)\right|^2/\eta_k \tag{3.16}$$

The maximum output SNR will occur when $w_k(t) = K\,h_k^*(t)s_b^*(t-\tau_0)/\eta_k$, that is, when the weights are proportional to the conjugate of each fading signal and inversely proportional to the average noise power. In fact, since the factor $s_b^*(t-\tau_0)$ is identical for all the branches, all that the combiner weights have to do is to compensate for the phase shift of the different channels and assign a larger weight to stronger signals, that is $w_k(t) = K\,h_k^*(t)/\eta_k$. Rewriting (3.16) in the case of maximal-ratio combining, gives

$$\begin{aligned} \gamma_{\max} &= \frac{1}{2}\sum_{k=0}^{M-1}\left|\alpha_k(t)s_b(t-\tau_0)\right|^2/\eta_k \\ &= \frac{1}{2}\sum_{k=0}^{M-1} r_k^2/\eta_k = \sum_{k=0}^{M-1}\gamma_k \end{aligned} \tag{3.17}$$

Thus, the SNR at the combiner's output is the sum of the branch SNRs. The former will be responsible for the error performance of the receiver system and should be as high as possible. A more important SNR measure is obtained by averaging (3.17) over a symbol period $T = 1/B$, yielding

$$\gamma_s = \frac{1}{\eta_k}\sum_{k=0}^{M-1}\alpha_k^2(t)\left[\frac{1}{T}\int_{\tau_0}^{\tau_0+T}\frac{1}{2}\left|s_b(t-\tau_0)\right|^2 dt\right] = \frac{\mathcal{E}_s}{\mathcal{N}_0}\sum_{k=0}^{M-1}\alpha_k^2(t) = \sum_{k=0}^{M-1}\gamma_{sk} \tag{3.18}$$

where $\mathcal{E}_s$ is the symbol energy of the bandpass signal $s(t)$, and it is assumed that $\alpha_k^2(t)$ is invariant during the symbol period. To evaluate the error performance of MRC, the probability density function of the output SNR must be obtained and, if possible, should account for the correlation between the antennas, since in most situations it is not feasible to provide enough separations among them.

3.4 *Statistical Characterization of the MRC*

In Section 2.4.7 it was shown that the envelope of a multipath fading signal may be statistically described by the Nakagami-m distribution. This conclusion was withdrawn from the fact that the squared envelope $\alpha_k^2(t)$ of the channel response between the transmitter antenna and the k-th receiver branch can be approximated by the sum of the squares of $2m$ zero-mean independent Gaussian random variables, as given by (2.98). This means

that (3.18) can alternatively be written as

$$\gamma_s = \frac{\mathcal{E}_s}{\mathcal{N}_0}\sum_{k=0}^{M-1}\alpha_k^2(t) = \frac{1}{2}\sum_{k=0}^{M-1}\sum_{i=0}^{2m-1}\frac{2\mathcal{E}_s}{\mathcal{N}_0}\alpha_{ki}^2(t) = \frac{1}{2}\sum_{k=0}^{M-1}\sum_{i=0}^{2m-1}x_{ki}^2 \tag{3.19}$$

where x_{ki} is a zero-mean Gaussian random variable associated with the k-th branch and m is the *fading figure* of the multipath channel (it is assumed that the fading figure is the same for all antenna branches). The joint characteristic function of the γ_{sk}'s is given by

$$\Phi_{\boldsymbol{\Gamma}_s}(\boldsymbol{\Omega}) = E\left[e^{j\boldsymbol{\Omega}^T\boldsymbol{\Gamma}_s}\right] \tag{3.20}$$

where $\boldsymbol{\Omega} = (\omega_0,..,\omega_{M-1})^T$ and $\boldsymbol{\Gamma}_s = (\gamma_{s0},..,\gamma_{s(M-1)})^T$. Expanding the vectors and noting that the x_{*i}'s are statistically independent from the x_{*j}'s for $i \neq j$, (3.20) gives

$$\begin{aligned}\Phi_{\boldsymbol{\Gamma}_s}(\boldsymbol{\Omega}) &= E\left[\exp\left(j\sum_{k=0}^{M-1}\omega_k\gamma_{sk}\right)\right] \\ &= E\left[\exp\left(j\frac{1}{2}\sum_{i=0}^{2m-1}\sum_{k=0}^{M-1}\omega_k x_{ki}^2\right)\right] \\ &= E\left[\exp\left(j\frac{1}{2}\sum_{k=0}^{M-1}\omega_k x_{ki}^2\right)\right]^{2m} = E\left[e^{j\frac{1}{2}\boldsymbol{X}^T\boldsymbol{D}_{\boldsymbol{\Omega}}\boldsymbol{X}}\right]^{2m}\end{aligned} \tag{3.21}$$

where $\boldsymbol{X} = \left(x_{0i},..,x_{(M-1)i}\right)^T$ and $\boldsymbol{D}_{\boldsymbol{\Omega}} = diag\,(\omega_0,..,\omega_{M-1})$. Now, since the x_{ki}'s are jointly Gaussian random variables, the probability density function of $\boldsymbol{X}$ is given in vector notation as

$$f(\boldsymbol{X}) = (2\pi)^{-M/2}\,|\boldsymbol{C}_{\boldsymbol{X}}|^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{X}^T\boldsymbol{C}_{\boldsymbol{X}}^{-1}\boldsymbol{X}\right) \tag{3.22}$$

where $\boldsymbol{C}_{\boldsymbol{X}} = E\left[\boldsymbol{X}\boldsymbol{X}^T\right]$ is a positive-definite symmetrical covariance matrix. Taking the expectation in (3.21) it follows that

$$\begin{aligned}\Phi_{\boldsymbol{\Gamma}_s}(\boldsymbol{\Omega}) &= \left[(2\pi)^{-M/2}\,|\boldsymbol{C}_{\boldsymbol{X}}|^{-1/2}\int_{-\infty}^{\infty}..\int_{-\infty}^{\infty}e^{-\frac{1}{2}\boldsymbol{X}^T(\boldsymbol{C}_{\boldsymbol{X}}^{-1}-j\boldsymbol{D}_{\boldsymbol{\Omega}})\boldsymbol{X}}\,d\boldsymbol{X}\right]^{2m} \\ &= \left[\left|\boldsymbol{C}_{\boldsymbol{X}}^{-1} - j\boldsymbol{D}_{\boldsymbol{\Omega}}\right|^{-1/2}|\boldsymbol{C}_{\boldsymbol{X}}|^{-1/2}\right]^{2m} \\ &= |\boldsymbol{I} - j\boldsymbol{D}_{\boldsymbol{\Omega}}\boldsymbol{C}_{\boldsymbol{X}}|^{-m}\end{aligned} \tag{3.23}$$

This result easily generalizes for the case of different fading figures in each branch. Rearranging the branches such that $m_k \leq m_{k+1}$, (3.21) becomes

$$\begin{aligned}\Phi_{\boldsymbol{\Gamma}_s}(\boldsymbol{\Omega}) &= E\left[\exp\left(j\frac{1}{2}\sum_{k=0}^{M-1}\sum_{i=0}^{2m_k-1}\omega_k x_{ki}^2\right)\right] \\ &= \prod_{l=0}^{M-1}E\left[\exp\left(j\frac{1}{2}\sum_{k=l}^{M-1}\omega_k x_{ki}^2\right)\right]^{2n_l} = \prod_{l=0}^{M-1}E\left[e^{j\frac{1}{2}\boldsymbol{X}_l^T\boldsymbol{D}_{\boldsymbol{\Omega}_l}\boldsymbol{X}_l}\right]^{n_l} \\ &= \prod_{l=0}^{M-1}\left|\boldsymbol{I}_l - j\boldsymbol{D}_{\boldsymbol{\Omega}_l}\boldsymbol{C}_{\boldsymbol{X}_l}\right|^{-n_l}\end{aligned} \tag{3.24}$$

where $\boldsymbol{X}_l = \left(x_{li},..,x_{(M-1)i}\right)^T$, $\boldsymbol{\Omega}_l = \left(\omega_{li},..,\omega_{(M-1)i}\right)^T$ and the difference of the fading param-

eters is

$$n_l = \begin{cases} m_0, & l = 0 \\ m_l - m_{l-1}, & 1 \le l \le M-1 \end{cases} \tag{3.25}$$

Since the antennas are usually close to one another, it is often reasonable to assume equal fading figures for all the branches, and therefore (3.23) will serve as the starting point for the ensuing analysis.

From (3.23), the characteristic function (CHF) of the maximum achievable output SNR in (3.17) is obtained by replacing $\boldsymbol{D_\Omega}$ by $\omega \boldsymbol{I}$, giving

$$\Phi_{\gamma_s}(\omega) = |\boldsymbol{I} - j\omega \boldsymbol{C_X}|^{-m} \tag{3.26}$$

Furthermore, since $\boldsymbol{C_X}$ is a real-symmetric matrix, it is orthogonally similar to a real-diagonal matrix $\boldsymbol{D_\lambda}$, that is

$$\boldsymbol{C_X} = \boldsymbol{Q}\boldsymbol{D_\lambda}\boldsymbol{Q}^T \tag{3.27}$$

where $\boldsymbol{\lambda} = (\lambda_0,..,\lambda_{M-1})$ is a vector of real and positive eigenvalues of $\boldsymbol{C_X}$ (assumed distinct without loss of generality), and $\boldsymbol{Q}$ is a unitary matrix with a complete set of M orthogonal eigenvectors forming its columns. Equation (3.26) simplifies to

$$\begin{aligned} \Phi_{\gamma_s}(\omega) &= \left|\boldsymbol{Q}(\boldsymbol{I} - j\omega \boldsymbol{D_\lambda})\boldsymbol{Q}^T\right|^{-m} \\ &= \prod_{k=0}^{M-1} (1 - j\omega\lambda_k)^{-m} \end{aligned} \tag{3.28}$$

An alternative interpretation of (3.28) stems from the fact that $\Phi_{\gamma_s}(\omega)$ is independent of the coordinate system chosen to express $\boldsymbol{C_X}$, and therefore (3.22) may be expressed as the product of M independent Gaussian pdfs

$$\begin{aligned} f(\boldsymbol{Y}) &= (2\pi)^{-M/2} |\boldsymbol{D_\lambda}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{Y}^T \boldsymbol{D_\lambda}^{-1}\boldsymbol{Y}\right) \\ &= \prod_{k=0}^{M-1} (2\pi\lambda_k)^{-1/2} \exp\left(\frac{y_k^2}{2\lambda_k}\right) \end{aligned} \tag{3.29}$$

in which $\boldsymbol{Y} = \boldsymbol{Q}^T\boldsymbol{X}$ is an isometry in $\Re^M$ that diagonalizes the quadratic form $\boldsymbol{X}^T\boldsymbol{C_X}^{-1}\boldsymbol{X}$. Consequently, from (3.19), γ_s can be regarded as the sum of mM independent chi-squared random variables with two degrees of freedom, and hence (3.28) results. Furthermore, it is a rational function in $j\omega$, and therefore can be expanded into partial fractions

$$\Phi_{\gamma_s}(\omega) = \sum_{k=0}^{M-1} \sum_{l=1}^{m} A_{kl} (1 - j\omega\lambda_k)^{-l} \tag{3.30}$$

where the coefficients A_{kl} are computed from the equation

$$A_{kl} = \frac{1}{(m-l)!(-\lambda_k)^{m-l}} \left[\frac{d^{m-l}}{d(j\omega)^{m-l}} \left[(1 - j\omega\lambda_k)^m \Phi_{\gamma_s}(\omega) \right] \right]_{j\omega = 1/\lambda_k} \tag{3.31}$$

Taking the inverse Fourier transform of (3.30), the probability density function of the

output SNR γ_s is

$$\begin{aligned} f(\gamma_s) &= \sum_{k=0}^{M-1}\sum_{l=1}^{m} A_{kl}\,\frac{1}{2\pi}\int_{-\infty}^{\infty}(1-j\omega\lambda_k)^{-l}\,e^{-j\omega\gamma_s}\,d\omega \\ &= \sum_{k=0}^{M-1}\sum_{l=1}^{m} A_{kl}\,\frac{\gamma_s^{l-1}}{(l-1)!\,\lambda_k^{l}}\,e^{-\gamma_s/\lambda_k} = \sum_{k=0}^{M-1}\sum_{l=1}^{m} A_{kl} f(\gamma_s,k,l) \end{aligned} \tag{3.32}$$

In words, the pdf of γ_s is the linear combination of mM independent gamma variables with pdfs $f(\gamma_s,k,l)$. Of particular interest in (3.32) are two special cases:

1. **Rayleigh fading** - for $m=1$ the SNR pdf in (3.32) reduces to

$$f(\gamma_s) = \sum_{k=0}^{M-1} A_{k0}\,\frac{1}{\lambda_k}\,e^{-\gamma_s/\lambda_k} \tag{3.33}$$

2. **Independent Fading** - the branches are uncorrelated and $\boldsymbol{C_X}$ is a diagonal matrix with eigenvalues

$$\lambda_k = E\left[x_{ki}^2\right] = \frac{1}{m}E[\gamma_{sk}] = \frac{1}{m}\Gamma_{sk} \tag{3.34}$$

These two cases are noteworthy mainly because they are simple, easier to manipulate mathematically, and often lead to reasonable approximations to the performance of a particular diversity system.

If the average SNR Γ_{sk} in each branch is admitted to be equal for all branches, then (3.32) simplifies to

$$f(\gamma_s) = \frac{\gamma_s^{mM-1}}{(mM-1)!}\left(\frac{mM}{\Gamma}\right)^{mM} e^{-mM\gamma_s/\Gamma_s} \tag{3.35}$$

where $\Gamma_s = E[\gamma_s] = ME[\gamma_{sk}]$ for $k = 0,1,..,M-1$. It is clear from (3.35) that $f(\gamma_s)$ is a function of mM, which is the product of the fading figure of each branch and the number of branches. This is equivalent to say that the effect of having several uncorrelated branches is an increase in the overall fading figure of the single-input multiple-output (SIMO) channel, resulting in a less severe signal fading condition.

Traditional communication systems were optimized for *Additive White Gaussian Noise* (AWGN) and *Intersymbol Interference* (ISI) as restraints to error probability and maximum achievable data rates of communication channels. Using the output SNR pdf $f(\gamma_s)$, it is possible to extend the error performance and capacity estimates to fading channels and diversity combining systems. This is done in the mean sense by statistical averaging any SNR dependent function $E(\gamma)$ over $f(\gamma)$, that is

$$E_{avg} = \int_0^{\infty} E(\gamma) f(\gamma)\,d\gamma \tag{3.36}$$

3.5 *Error Performance of the MRC*

This section will aim to illustrate the effects of fading, diversity combining and branch correlation in the error performance of a wireless communication system. It will be assumed that the channel is frequency non-selective, or otherwise can be converted to frequency

non-selective using a special type of receiver (e.g. a receiver capable of combining the multiple flat-fading processes of a frequency selective channel). Maximal-ratio combining and binary PSK modulation will be used throughout the analysis.

When all the receiver branches have the same noise power, the combiner weights may be written as $w_k(t) = \alpha_k(t)e^{-j\theta_k(t)}$, and from (3.13) the combiner output reduces to

$$r_b(t) = s_b(t-\tau_0)\sum_{k=0}^{M-1}\alpha_k^2(t) + \sum_{k=0}^{M-1}\alpha_k(t)n_k(t)e^{-j\theta_k(t)} \tag{3.37}$$

Also, for BPSK modulation a single matched filter is sufficient, and since $\alpha_k(t)$ is virtually constant during a symbol period (slow fading), the decision variable is given by

$$\begin{aligned} d &= \int_{\tau_0}^{\tau_0+T} r(t)s(t-\tau_0)dt = \frac{1}{2}\Re\mathrm{e}\left\{\int_{\tau_0}^{\tau_0+T} r_b(t)s_b^*(t-\tau_0)\,dt\right\} \\ &= \left(\frac{1}{2}\int_{\tau_0}^{\tau_0+T}|s_b(t-\tau_0)|^2\,dt\right)\sum_{k=0}^{M-1}\alpha_k^2(t) + \sum_{k=0}^{M-1}\frac{\alpha_k(t)}{2}\Re\mathrm{e}\left\{e^{-j\theta_k(t)}\int_{\tau_0}^{\tau_0+T} n_k(t)s_b^*(t-\tau_0)\,dt\right\} \\ &= \mathcal{E}_s\sum_{k=0}^{M-1}\alpha_k^2(t) + \sum_{k=0}^{M-1}\alpha_k(t)\mathrm{w}_{kr} \end{aligned} \tag{3.38}$$

where $\mathcal{E}_s$ is still the symbol energy of the bandpass signal $s(t)$ and w_{kr} is given by

$$\mathrm{w}_{kr} = \frac{1}{2}\Re\mathrm{e}\left\{e^{-j\theta_k(t)}\int_{\tau_0}^{\tau_0+T} n_k(t)s_b^*(t-\tau_0)\,dt\right\} \tag{3.39}$$

For fixed α_k's, the variable d is Gaussian distributed with mean

$$\mu_d = \mathcal{E}_s\sum_{k=0}^{M-1}\alpha_k^2(t) \tag{3.40}$$

and, using the equality $E[\,n_k^*(t)n_k(u)\,] = 2\mathcal{N}_0\delta(t-u)$, with variance

$$\sigma_d^2 = E\big[\,\mathrm{w}_{kr}^2\,\big]\sum_{k=0}^{M-1}\alpha_k^2(t) = \frac{\mathcal{N}_0\mathcal{E}_s}{2}\sum_{k=0}^{M-1}\alpha_k^2(t) \tag{3.41}$$

When a "positive" $s_b(t)$ is sent, an error will occur when $d < 0$, for which the receiver incorrectly decides in favour of $-s_b(t)$, and the error probability is

$$\begin{aligned} P_e(\gamma_s) &= Q\left(\frac{\mu_d}{\sigma_d}\right) \\ &= Q\big(\sqrt{2\gamma_s}\,\big) \end{aligned} \tag{3.42}$$

where γ_s is the SNR at the combiner's output, averaged over a symbol period, as given by (3.18). Hence, the average error probability may be written as

$$P_e = \int_0^\infty Q\big(\sqrt{2\gamma_s}\,\big)\,f(\gamma_s)d\gamma_s \tag{3.43}$$

3.5.1 INDEPENDENT FADING

Inserting (3.35) into (3.43) and expanding the $Q(\cdot)$ function, the average error probability

is written in integral form as

$$P_e = \frac{(\Lambda/\Gamma_s)^{\Lambda}}{\sqrt{2\pi}(\Lambda-1)!}\int_0^{\infty}\left[\int_{\sqrt{2\gamma_s}}^{\infty} e^{-x^2/2}dx\right]\left[\gamma_s^{\Lambda-1}e^{-\Lambda\gamma_s/\Gamma_s}\right]d\gamma_s \tag{3.44}$$

where $\Lambda = mM$ is the MRC *effective diversity order* and $\Gamma_s = E[\gamma_s]$ is the average SNR at the output of the combiner. Using the equality

$$\int x^n e^{-ax}dx = -e^{-ax}\sum_{k=0}^{n}\frac{n!}{(n-k)!}\frac{x^{n-k}}{a^{k+1}} \tag{3.45}$$

and integrating (3.44) by parts yields

$$\begin{aligned} P_e &= \frac{1}{\sqrt{2\pi}}\left[\int_{\sqrt{2\gamma_s}}^{\infty} e^{-x^2/2}dx\right]_0^{\infty} - \\ &\quad -\frac{1}{2\sqrt{\pi}}\sum_{k=0}^{\Lambda-1}\frac{(\Gamma_s/\Lambda)^{k+1-\Lambda}}{(\Lambda-k-1)!}\int_0^{\infty} e^{-\gamma_s(1+\Lambda/\Gamma_s)}\gamma_s^{\Lambda-k-3/2}d\gamma_s \\ &= \frac{1}{2} + \sqrt{\frac{\Gamma_s}{\Lambda}}\sum_{k=0}^{\Lambda-1}\frac{(1+\Gamma_s/\Lambda)^{1/2+k-\Lambda}}{(\Lambda-k-1)!}\prod_{l=0}^{\Lambda-k-1}(l-1/2) \end{aligned} \tag{3.46}$$

A plot of Equation (3.46) as a function of the average output SNR Γ_s, for several effective

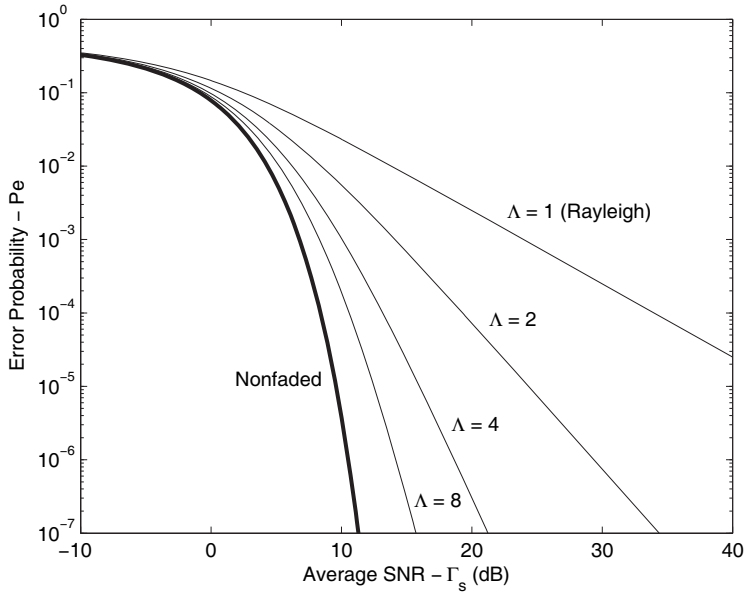

FIGURE 3.2 The effect of MRC diversity combining on the error performance of coherent BPSK.

diversity orders Λ, is shown in Figure 3.2. It is noticeable the reduction in error probability for the same output SNR when diversity is utilized. Figure 3.2 also emphasizes how the use of diversity makes the error probability curves bend towards the nonfaded signal curve when the number of receiver branches increases. For instance, achieving a 10^{-4} error probability target with no diversity and a Rayleigh fading scenario would require approximately 34 dB of SNR. Using eight antennas at the receiver (or four antennas in the case of $m = 2$ in each branch) the SNR requirement drops to more or less 11 dB, which is a remarkable SNR gain of 23 dB.

3.5.2 CORRELATED FADING

Figure 3.2 reveals a clear error performance advantage of MRC diversity combining for generic fading channels. Nevertheless, in practical situations there is always some level of correlation between the signals of different receiver branches that will diminish the diversity gain of any system. Obtaining the correlation between the γ_{sk}'s as a function of the correlation between the x_{ki}'s defined in (3.19) is a straightforward process. Putting $j\boldsymbol{D}_{\boldsymbol{\Omega}} = \boldsymbol{D}_{\boldsymbol{S}} = diag(s_1,..,s_M)$ in (3.23), the joint moment-generating function (MGF) of the γ_{sk}'s is given by

$$\boldsymbol{M}_{\boldsymbol{\Gamma}_s}(\boldsymbol{S}) = |\boldsymbol{I} - \boldsymbol{D}_{\boldsymbol{S}}\boldsymbol{C}_{\boldsymbol{X}}|^{-m} \tag{3.47}$$

Letting $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{D}_{\boldsymbol{S}}\boldsymbol{C}_{\boldsymbol{X}}$, differentiation of (3.47) gives

$$\begin{aligned} E[\,\gamma_{sk}\,] &= \left[\frac{\partial \boldsymbol{M}_{\boldsymbol{\Gamma}_s}}{\partial s_k}\right]_{\boldsymbol{S}=\boldsymbol{0}} \\ &= \left[-m|\boldsymbol{A}|^{-m-1}|\boldsymbol{A}_k|\right]_{\boldsymbol{S}=\boldsymbol{0}} \\ &= mE\left[x_{ki}^2\right] \end{aligned} \tag{3.48}$$

and also

$$\begin{aligned} E[\,\gamma_{sk}\gamma_{sl}\,] &= \left[\frac{\partial^2 \boldsymbol{M}_{\boldsymbol{\Gamma}_s}}{\partial s_k \partial s_l}\right]_{\boldsymbol{S}=\boldsymbol{0}} \\ &= \left[m(m+1)|\boldsymbol{A}|^{-m-2}|\boldsymbol{A}_k||\boldsymbol{A}_l| - m|\boldsymbol{A}|^{-m-1}\frac{\partial|\boldsymbol{A}_k|}{\partial s_l}\right]_{\boldsymbol{S}=\boldsymbol{0}} \\ &= m^2E^2\left[x_{ki}^2\right] + mE^2\left[x_{ki}x_{li}\right] \end{aligned} \tag{3.49}$$

where $\boldsymbol{A}_k$ is the matrix $\boldsymbol{A}$ with the k-th row differentiated with respect to s_k. To normalize (3.49) one may use the correlation coefficient

$$\begin{aligned} \rho_{\gamma_{sk}\gamma_{sl}} &= \frac{E[\,\gamma_{sk}\gamma_{sl}\,] - E^2[\,\gamma_{sk}\,]}{E\left[\gamma_{sk}^2\right] - E^2[\,\gamma_{sk}\,]} \\ &= \frac{E^2[\,x_{ki}x_{li}\,]}{E^2\left[x_{ki}^2\right]} = \rho_{x_{ki}x_{li}}^2 \end{aligned} \tag{3.50}$$

where $E\left[x_{ki}^2\right] = E[\,\gamma_s\,]/\,mM = \Gamma_s\,/\,\Lambda$. As expected, (3.50) is independent of the fading parameter m. Now, invoking the result of (3.6), rewritten as

$$E[\,x_{ki}x_{li}\,] = E\left[x_{ki}^2\right]J_0(\beta d_{kl}) = \frac{\Gamma_s}{\Lambda}J_0(\beta d_{kl}) \tag{3.51}$$

the correlation coefficient $\rho_{\gamma_{sk}\gamma_{sl}}$ between the SNR of two different branches can finally be expressed as

$$\rho_{\gamma_{sk}\gamma_{sl}} = J_0^2(\beta d_{kl}) \tag{3.52}$$

where β is the phase constant and d_{kl} is the separation distance between the antennas of branches k and l. This equation generalizes (3.9), which was only applicable in Rayleigh fading conditions. It has, however, two significant drawbacks: 1) it is not a good approximation for irregular distributions of wave arrival (e.g. within sub-urban areas); 2) for low angular spreads, it depends upon the direction of wave arrival (e.g. near base stations). The statistical consequence of such scenarios will be the emergence of some cross-correlation

between the in-phase and quadrature components of different γ_k's, that is

$$\begin{aligned} E\left[x_{k(2i)}x_{l(2i)}\right] &\neq 0 \\ E\left[x_{k(2i)}x_{l(2i+1)}\right] &\neq 0 \end{aligned} \tag{3.53}$$

which, for the Rayleigh case, is equivalent to say that (3.3) is not valid. Also, the in-phase and quadrature components are given generically by

$$x_{k(2i)} + jx_{k(2i+1)} = \sum_{l=0}^{L_{ki}-1} a_{lki} e^{j\left(\varphi_{lki} + 2\pi\frac{\nu}{\lambda}\cos\phi_{lki}t\right)} \tag{3.54}$$

for a receiver (branch) that moves with a velocity ν relative to the transmitter, and with an angle ϕ_{lki} towards the incoming wave of the lki-th scatterer. Carrying out the same steps in (3.19)-(3.23) and (3.47)-(3.50), but this time abiding by (3.53), it easily follows that the SNR correlation coefficient is given by

$$\rho_{\gamma_{sk}\gamma_{sl}} = \frac{E^2\left[x_{k(2i)}x_{l(2i)}\right] + E^2\left[x_{k(2i)}x_{l(2i+1)}\right]}{E^2\left[x_{k(2i)}^2\right]} \tag{3.55}$$

for any $i = 0,..,m-1$.

For the sake of simplicity, the following analysis will be performed assuming a cluttered urban environment, so that uniform distribution of wave arrival is a good approximation, and hence (3.52) can be used. Suppose that an evenly spaced linear antenna array is used to receive signals that suffer from Nakagami fading. The antennas are connected to a MRC combiner and the SNR of the detector's output is modelled by the probability density function in (3.32). The covariance matrix of the x_{ki}'s may be written as

$$\begin{aligned} \boldsymbol{C_X} &= \frac{\Gamma_s}{\Lambda}\begin{pmatrix} 1 & J_0(\beta d_{01}) & \dots & J_0\left(\beta d_{0(M-1)}\right) \\ J_0(\beta d_{10}) & 1 & \dots & J_0\left(\beta d_{1(M-1)}\right) \\ \vdots & \vdots & \ddots & \vdots \\ J_0\left(\beta d_{(M-1)0}\right) & J_0\left(\beta d_{(M-1)1}\right) & \dots & 1 \end{pmatrix} \\ &= \frac{\Gamma_s}{\Lambda}\widehat{\boldsymbol{C}}_{\boldsymbol{X}} \end{aligned} \tag{3.56}$$

where $\widehat{\boldsymbol{C}}_{\boldsymbol{X}}$ is the normalized covariance matrix. Denoting an eigenvalue of $\widehat{\boldsymbol{C}}_{\boldsymbol{X}}$ by $\hat{\lambda}_k$, and given that (3.32) is simply a linear combination of (3.35), the average error probability at the detector's output is similarly a linear combination of (3.46) with the substitutions $\Gamma_s/\Lambda \to \lambda_k = \hat{\lambda}_k\Gamma_s/\Lambda$ and $\Lambda \to l$, that is

$$P_e = \sum_{k=0}^{M-1}\sum_{l=1}^{m}\widehat{A}_{kl}\left[\frac{1}{2} + \sqrt{\hat{\lambda}_k\frac{\Gamma_s}{\Lambda}}\sum_{i=0}^{l-1}\left[\frac{\left(1+\hat{\lambda}_k\Gamma_s/\Lambda\right)^{1/2+i-l}}{(l-i-1)!}\prod_{j=0}^{l-i-1}(j-1/2)\right]\right] \tag{3.57}$$

The $\widehat{A}_{kl}$'s can be found from (3.31) using the $\hat{\lambda}_k$'s and the help of a recursive algorithm that automates the calculation for any diversity order M and fading figure m. Figures 3.3 and 3.4 show plots of (3.57) for several situations: two antenna spacings, $\lambda/4$ and $\lambda/8$, and four m/M conditions. These figures illustrate the following:

- the presence of correlation between the signals of different antennas reduces the SNR gain obtained by employing diversity at the receiver;

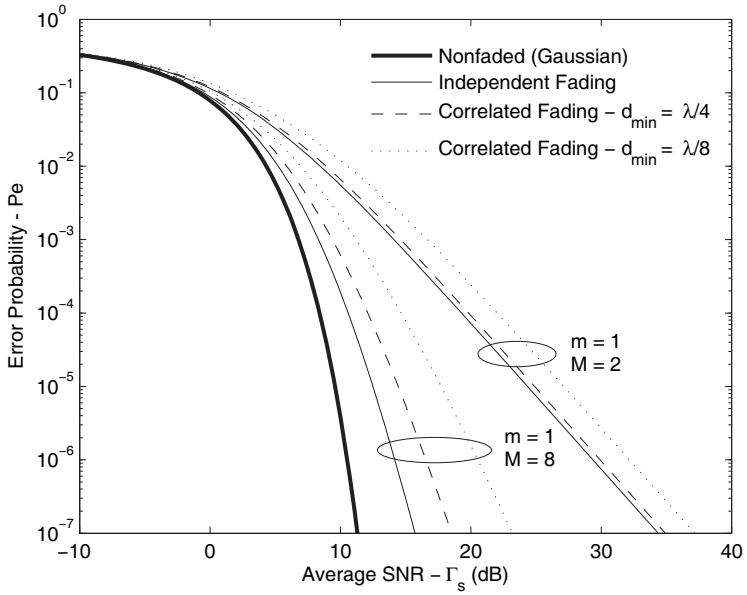

FIGURE 3.3 Error performance of MRC combiner for a linear array of M antennas, correlated fading with fading figure $m = 1$, and uniform distribution of wave arrival.

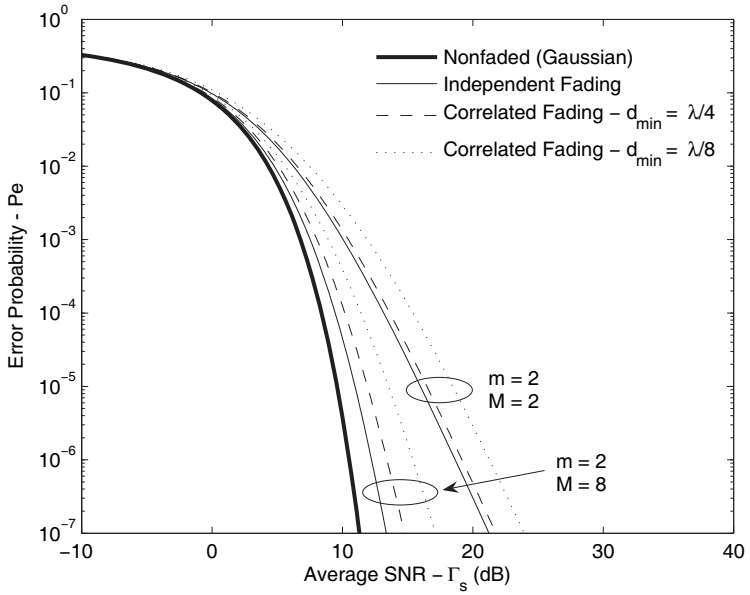

FIGURE 3.4 Error performance of MRC combiner for a linear array of M antennas, correlated fading with fading figure $m = 2$, and uniform distribution of wave arrival.

- the reduction in SNR gain is more significant for higher diversity orders and for smaller antenna separation;
- as the separation between the antennas increases the error performance approaches that of independent fading;

Thus, improving the error performance in a diversity system with correlated fading is equivalent to increasing the diversity order M and/or increasing the antenna separation. One may also deduce that sometimes it is preferable to increase the order of the array than to increase the *aperture* (i.e. the sum the antenna separations in a linear array) of a lower order array. Either way, the cost of diversity is always higher system complexity and size, which is often an expensive price to pay in mobile units that require low power consump-

tion and reduced sizes.

3.6 *Information-theoretic Capacity of the MRC*

In Section 3.5 it was possible to analyse from a general viewpoint the error performance of a slowly-fading multipath propagation channel with spatial diversity reception and maximal-ratio combining. It is well known that the error performance of a communication system is directly related to the data rate of transmission, in the sense that increasing the latter for a given SNR γ_s has the consequence of degrading the former. Here, the common denominator is γ_s. From Shannon's *information capacity theorem*, a communication channel perturbed by AWGN noise has a capacity given by

$$C = \log_2(1+\gamma_s) \quad \text{bits/cycle} \tag{3.58}$$

where a "cycle" (or seconds × Hz) is the time of two channel transmissions (i.e. two samples) if the sampling rate of the input signal equals the Nyquist rate. It means that it is theoretically possible to employ some channel coding technique so that the maximum achievable signalling rate with zero error probability is C, and also that it is theoretically impossible to have error-free communication when the signalling rate is above C. It is therefore important to know to what extent is the channel capacity reduced in the case of a fading channel, and how much can diversity reception compete against this reduction.

In a generic fading channel γ_s must be interpreted as a random variable, which means that (3.58) has to be obtained in the average sense, that is

$$C_{avg} = \int_0^\infty f(\gamma_s)\log_2(1+\gamma_s)\,d\gamma_s \quad \text{bits/cycle} \tag{3.59}$$

When the fading severity is characterized by an integral fading parameter m, $f(\gamma_s)$ is given by (3.32). Thus, integrating (3.59) by parts,

$$\begin{aligned} C_{avg} &= \sum_{k=0}^{M-1}\sum_{l=1}^{m} A_{kl}\int_0^\infty \frac{\gamma_s^{l-1}}{(l-1)!\,\lambda_k^l} e^{-\gamma_s/\lambda_k}\log_2(1+\gamma_s)\,d\gamma_s \\ &= \log_2 e \sum_{k=0}^{M-1}\sum_{l=1}^{m} A_{kl}\sum_{i=0}^{l-1}\frac{1}{i!\,\lambda_k^i}\int_0^\infty \frac{1}{1+\gamma_s}\gamma_s^i e^{-\gamma_s/\lambda_k}d\gamma_s \end{aligned} \tag{3.60}$$

performing the change of variables $\gamma_s = \eta - 1$,

$$C_{avg} = \log_2 e \sum_{k=0}^{M-1}\sum_{l=1}^{m} A_{kl}\sum_{i=0}^{l-1}\frac{1}{i!\,\lambda_k^i} e^{1/\lambda_k}\int_1^\infty \frac{1}{\eta}(\eta-1)^i e^{-\eta/\lambda_k}d\eta \tag{3.61}$$

and using the binomial theorem,

$$(\eta-1)^i = \sum_{j=0}^{i}\binom{i}{j}\eta^j(-1)^{i-j} \tag{3.62}$$

the average channel capacity is rewritten as

$$C_{avg} = \log_2 e \sum_{k=0}^{M-1}\sum_{l=1}^{m} A_{kl}\, e^{1/\lambda_k}\sum_{i=0}^{l-1}\frac{1}{i!}\sum_{j=0}^{i}\binom{i}{j}\left(-\frac{1}{\lambda_k}\right)^{i-j}\Gamma(j,1/\lambda_k) \tag{3.63}$$

where

$$\Gamma(j,1/\lambda_k) = \int_{1/\lambda_k}^{\infty} \eta^{j-1} e^{-\eta} d\eta \tag{3.64}$$

is the incomplete gamma function. If the branches fade independently, $f(\gamma_s)$ is given by (3.35) and the expression for the average channel capacity is

$$C_{avg} = \frac{e^{\Lambda/\Gamma_s}}{\ln 2} \sum_{i=0}^{\Lambda-1} \frac{1}{i!} \sum_{j=0}^{i} \binom{i}{j} \left(-\frac{\Lambda}{\Gamma_s}\right)^{i-j} \Gamma(j,\Lambda/\Gamma_s) \tag{3.65}$$

Equation (3.65) is plotted in Figure 3.5 as a function of the average SNR Γ_s, for effective diversity orders $\Lambda = 1,2,4,8$. The Gaussian channel capacity is also represented for com-

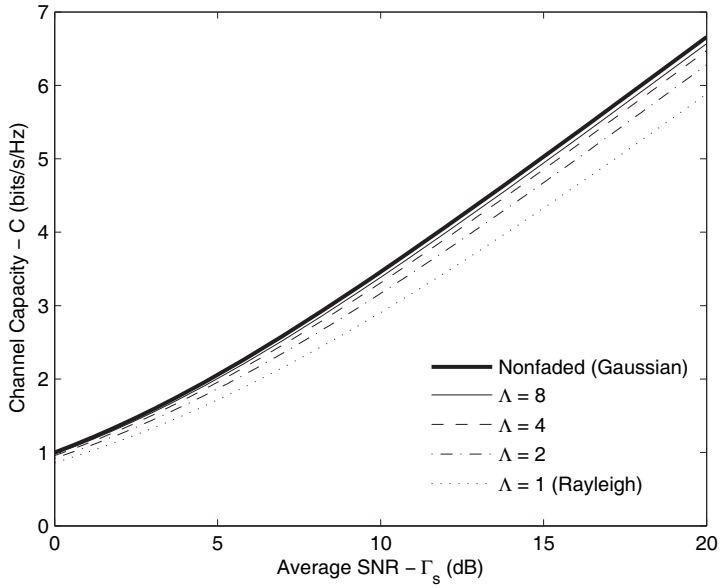

FIGURE 3.5 Information-theoretic capacity of MRC combiner for several effective diversity orders, in the case of independent fading, and uniform distribution of wave arrival.

parison (solid curve), and it is equivalent to the limiting case $\Lambda = \infty$. Therefore, the capacity of a multipath fading channel with diversity at the receiver converges to Shannon's result for an AWGN channel. However, Figure 3.5 surprisingly shows that in the range of Γ_d between 0 and 20 dB, the difference in capacity is always less than 1 bit/s/Hz, meaning that in the case of independent fading the loss of information-theoretic capacity is negligible. This conclusion applies without modification to the case of correlated fading. Computing the covariance matrix in (3.56) for two antenna separations, $\lambda/4$ and $\lambda/8$, in a linear array of four antennas, finding the eigenvalues λ_k and the coefficients A_{kl}, and plotting (3.63), Figure 3.6 results. For comparison, the capacity curves for the Rayleigh and Gaussian channels are also depicted. As long as the channel is a fading channel, the worst scenario is always Rayleigh fading (no diversity) and the best scenario is always independent fading, since the capacity difference from Shannon's result is maximum and minimum, respectively. When the fading is correlated, the higher is the antenna separation, the is higher the achievable capacity for a given diversity order M, and for a given antenna separation, the higher is the diversity order M, the higher is the achievable capacity. The par-

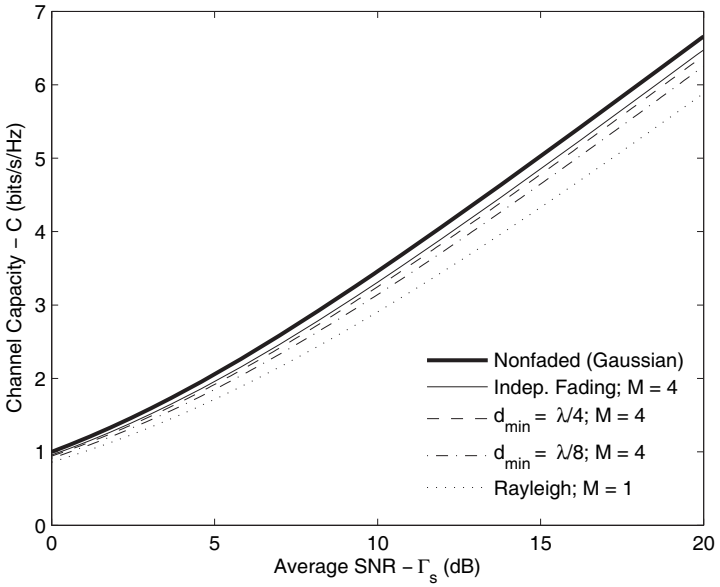

FIGURE 3.6 Information-theoretic capacity of MRC combiner for different antenna separations, in the case of correlated fading, and uniform distribution of wave arrival. $m = 1$.

ticular results shown in Figure 3.6 are fully generalizable for any diversity order M, branch fading figure m, and antenna separation. Nevertheless, this generalization assumes integral fading figures m. One should remember that worst-than-Rayleigh scenarios are possible for $1/2 \leq m < 1$, but since a mathematical treatment of nonintegral fading figures is required, these scenarios will be omitted here

CHAPTER 4

INFORMATION-THEORETIC ASPECTS OF MIMO CHANNELS

4.1 *Introduction*

The ever-growing need for high data-rates of communication triggered an "explosion" of research activity with the aim of finding revolutionary techniques to combat fading, and to increase the spectral efficiency of wireless communication systems. Based on the fundamental premise of geostatistics and biodiversity that nearby things are on average more alike than distant things, the invention of spatial diversity techniques naturally occurred.

Clearly, the way in which the received phasors add up at different locations should be nearly independent provided that the respective scatterers are also independent. To validate this argument, one could build an array of spaced antennas at the receiver and, hence, obtain a set of nearly independent linear equations with a single unknown: the transmitted signal. At this point we should be able to compute a better estimate of the transmitted signal. To enhance the system, we could add several unknowns to it or, equivalently, set up an array of spaced antennas at the transmitter side and send out independent signals from each antenna. If we could quantify the channel responses between each pair of transmitting/receiving antennas, we could also estimate a solution to the system of equations. Still, because of channel noise, the solution had to be necessarily an approximation. This is the basic concept behind single-user multiple-input multiple-output (MIMO) wireless systems, that is, to take advantage of multipath and diversity to improve communications.

There is an ongoing study within the re-search community on the actual possibilities of such concept. A review of the principles and recent advances of MIMO wireless systems can be found in [16], [17]. The primary aim of this chapter is to unveil a simple, compact, and generic analytic solution to the problem of finding the limit of the spectral efficiency (i.e. the capacity) that can be offered by single-user MIMO systems. The initial assumptions are: 1) the receiver has perfect channel state information (CSI); 2) the signals at different receiver antennas are correlated; 3) the signals from different transmitter an-

tennas are independent; and 4) there is no CSI at the transmitter.

4.2 *MIMO Channel Modelling*

A multiple-input multiple-output wireless channel, as its name indicates, is a channel composed of several antennas at the transmitter (input) and receiver (output) sides. To understand the advantages of such a channel, let the channel response between the ***j*-th transmitter antenna** and the ***i*-th receiver antenna** be modelled by a baseband-equivalent flat-fading process like the one in (3.11), rewritten as

$$h_{ij}(t) = \sum_{k=0}^{L_{ij}-1} a_{ijk} e^{j(\varphi_{ijk}+\omega_{ijk}t)} \tag{4.1}$$

This expression can be simplified by assuming that:

1. each wave that impinges on one antenna, impinges on all the others by a parallel path; essentially this means that the Doppler shift is identical for all the receiver antennas, i.e.

$$\omega_{ijk} \sim \omega_{jk} = \beta\nu\cos\phi_{jk} \tag{4.2}$$

where β is the phase constant, ν is the magnitude of the receiver's velocity relative to the transmitter, and ϕ_{jk} is the angle between the k-th impinging wave (from the j-th transmitter antenna) and the receiver's velocity vector relative to the transmitter; we also have that the number of multipaths is simply $L_{ij} \sim L_j$;

2. each wave that impinges on one antenna, impinges on all the others with the same amplitude, i.e. $a_{ijk} \sim a_{jk}$;

Introducing these simplifications, it follows that

$$h_{ij}(t) = \sum_{k=0}^{L_j-1} a_{jk} e^{j(\varphi_{ijk}+\omega_{jk}t)} = \alpha_{ij}(t)e^{j\theta_{ij}(t)} = u_{ij}(t) + jv_{ij}(t) \tag{4.3}$$

The output signal from the i-th receiver antenna is now obtained as the sum of the contributions of the input signals of the n_T transmitter antennas, weighted by the individual channel responses, that is

$$r_{bi}(t) = \sum_{j=0}^{n_T-1} h_{ij}(t)s_{bj}(t-\tau_0) + z_{bi}(t) \tag{4.4}$$

where $s_{bj}(t-\tau_0) = x_j\, g(t-\tau_0)$ is the delayed, baseband version of the input signal, x_j is the complex information symbol, $g(t-\tau_0)$ is a delayed unit energy shaping function, and $z_{bi}(t) = z_{bi}^I(t) + jz_{bi}^Q(t)$ is a complex valued baseband AWGN noise process. If the channel is slowly fading, it may be considered time-invariant during a symbol period, and consequently the time dependence in (4.4) can be disregarded. To recover the information symbol, the receiver must correlate the output signals with the pulse shaping function, yielding

$$y_i = \int_{\tau_0}^{\tau_0+T} r_{bi}(t)g(t-\tau_0)dt = \sum_{j=0}^{n_T-1} h_{ij}x_j + n_i \tag{4.5}$$

where n_i is a complex zero-mean gaussian random variable given by

$$n_i = \int_{\tau_0}^{\tau_0+T} z_{bi}(t)g(t-\tau_0)dt \tag{4.6}$$

The global channel response is now naturally written in vector/matrix form

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{4.7}$$

where $\mathbf{x}$ is the $n_T \times 1$ transmit vector, $\mathbf{y}$ is the $n_R \times 1$ receive vector, $\mathbf{n}$ is the $n_R \times 1$ noise vector, and $\mathbf{H}$ is the $n_R \times n_T$ channel matrix. In expanded form, the MIMO channel response is described by the equation

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n_R} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n_T} \\ h_{21} & h_{22} & \cdots & h_{2n_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_R1} & h_{n_R2} & \cdots & h_{n_Rn_T} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_T} \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_{n_R} \end{pmatrix} \tag{4.8}$$

4.3 *Maximum Entropy*

A measure of the uncertainty about a continuous random vector $\mathbf{x}$ is its *differential entropy* $\mathcal{H}_d$, given by

$$\mathcal{H}_d(f) = -E_f\left[\log_2 f(\mathbf{x})\right] \quad \text{(bits)} \tag{4.9}$$

where $f(\mathbf{x})$ is the probability density function of $\mathbf{x}$, and $E_f(\cdot)$ denotes the expectation over f. Let $\mathbf{x}$ be the complex random vector $\mathbf{x} = \mathbf{x}_I + j\mathbf{x}_Q$, and its covariance matrix be expressed as

$$\begin{aligned} \mathbf{C}_{\mathbf{xx}} &= E_f\left[\mathbf{x}\mathbf{x}^H\right] \\ &= \mathbf{C}_{\mathbf{x}_I\mathbf{x}_I} + \mathbf{C}_{\mathbf{x}_Q\mathbf{x}_Q} - j\left(\mathbf{C}_{\mathbf{x}_I\mathbf{x}_Q} - \mathbf{C}_{\mathbf{x}_Q\mathbf{x}_I}\right) \end{aligned} \tag{4.10}$$

The following relations apply

$$\mathbf{C}^T_{\mathbf{x}_I\mathbf{x}_I} = \mathbf{C}_{\mathbf{x}_I\mathbf{x}_I} \qquad \mathbf{C}^T_{\mathbf{x}_I\mathbf{x}_Q} = \mathbf{C}_{\mathbf{x}_Q\mathbf{x}_I} \tag{4.11}$$

If $f(\mathbf{x})$ is the joint Gaussian pdf, then it equals

$$f(\mathbf{x}_I, \mathbf{x}_Q) = \left[\det(2\pi\mathbf{D})\right]^{-1/2} \exp\left(-\frac{1}{2}\begin{pmatrix}\mathbf{x}_I^T & \mathbf{x}_Q^T\end{pmatrix}\mathbf{D}^{-1}\begin{pmatrix}\mathbf{x}_I \\ \mathbf{x}_Q\end{pmatrix}\right) \tag{4.12}$$

where n is the length of $\mathbf{x}$ and $\mathbf{D}$ is the covariance matrix

$$\mathbf{D} = \begin{pmatrix} \mathbf{C}_{\mathbf{x}_I\mathbf{x}_I} & \mathbf{C}_{\mathbf{x}_I\mathbf{x}_Q} \\ \mathbf{C}_{\mathbf{x}_Q\mathbf{x}_I} & \mathbf{C}_{\mathbf{x}_Q\mathbf{x}_Q} \end{pmatrix} \tag{4.13}$$

The corresponding characteristic function is

$$\Phi(\mathbf{w}) = \exp\left(-\frac{1}{2}\begin{pmatrix}\mathbf{w}_I^T & \mathbf{w}_Q^T\end{pmatrix}\mathbf{D}\begin{pmatrix}\mathbf{w}_I \\ \mathbf{w}_Q\end{pmatrix}\right) \tag{4.14}$$

Now, notice that if $\mathbf{x}$ is a circularly symmetric vector such that

$$\mathbf{C}_{\mathbf{x}_I\mathbf{x}_I} = \mathbf{C}_{\mathbf{x}_Q\mathbf{x}_Q} \qquad \mathbf{C}_{\mathbf{x}_I\mathbf{x}_Q} = -\mathbf{C}_{\mathbf{x}_Q\mathbf{x}_I} \tag{4.15}$$

then the following equality holds

$$\begin{aligned}\Phi(\mathbf{w}) &= \exp\left(-\frac{1}{2}\begin{pmatrix}\mathbf{w}_I^T & \mathbf{w}_Q^T\end{pmatrix}\mathbf{D}\begin{pmatrix}\mathbf{w}_I\\ \mathbf{w}_Q\end{pmatrix}\right)\\ &= \exp\left(-\frac{1}{4}\mathbf{w}^H\mathbf{C}_{\mathbf{xx}}\mathbf{w}\right)\end{aligned} \tag{4.16}$$

Therefore, taking the inverse Fourier transform of (4.16), it is found that, under the stated assumptions, the joint Gaussian pdf has the form

$$f(\mathbf{x}) = |\det(\pi\mathbf{C}_{\mathbf{xx}})|^{-1}\exp\left(-\mathbf{x}^H\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{x}\right) \tag{4.17}$$

which is a well known result from probability theory [39]. This form will be important in the following explanation. To find the pdf that maximizes the entropy $\mathcal{H}_d$, let $g(\mathbf{x})$ be another pdf such that $E_g[\mathbf{x}\mathbf{x}^H] = E_f[\mathbf{x}\mathbf{x}^H]$, and compute the difference

$$\mathcal{H}_d(g) - \mathcal{H}_d(f) = -E_f[\log_2 f(\mathbf{x})] + E_g[\log_2 g(\mathbf{x})] \tag{4.18}$$

But if $f(\mathbf{x})$ follows (4.17), then the expectations over f and g are equal

$$\begin{aligned}E_f[\log_2 f(\mathbf{x})] &= \log_2 e\left[-\ln\det(\pi\mathbf{C}_{\mathbf{xx}}) - E_f\left[\mathbf{x}^H\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{x}\right]\right]\\ &= \log_2 e\left[-\ln\det(\pi\mathbf{C}_{\mathbf{xx}}) - tr(E_f[\mathbf{x}\mathbf{x}^H]\mathbf{C}_{\mathbf{xx}}^{-1})\right]\\ &= E_g[\log_2 f(\mathbf{x})]\end{aligned} \tag{4.19}$$

where $tr(\cdot)$ denotes the trace of a matrix. As a result, the entropy difference is

$$\begin{aligned}\mathcal{H}_d(g) - \mathcal{H}_d(f) &= -E_g[\log_2 g(\mathbf{x})] + E_g[\log_2 f(\mathbf{x})]\\ &= E_g\left[\log_2\frac{f(\mathbf{x})}{g(\mathbf{x})}\right]\end{aligned} \tag{4.20}$$

To transform (4.20) into an inequality one uses a property of the logarithm:

$$\ln x \le x - 1 \tag{4.21}$$

Therefore

$$\mathcal{H}_d(g) - \mathcal{H}_d(f) \le \log_2 e\, E_g\left[\frac{f(\mathbf{x})}{g(\mathbf{x})} - 1\right] = 0 \tag{4.22}$$

with equality only if $f = g$. This equation proves that the pdf of a circularly symmetric complex gaussian random vector $\mathbf{x}$, given in (4.17), achieves the maximum possible entropy. Finding this maximum is as simple as taking the expectation in (4.9), yielding

$$\begin{aligned}\mathcal{H}_d(f) &= -E_f[\log_2 f(\mathbf{x})]\\ &= \log_2\det(\pi\mathbf{C}_{\mathbf{xx}}) + (\log_2 e)tr(E_f[\mathbf{x}\mathbf{x}^H]\mathbf{C}_{\mathbf{xx}}^{-1})\\ &= \log_2\det(\pi e\mathbf{C}_{\mathbf{xx}})\end{aligned} \tag{4.23}$$

Entropy maximization is an important result that can be used in the determination of the information-theoretic capacity of a MIMO channel.

4.4 *Information-theoretic Capacity*

The maximum theoretically achievable signalling rate with error-free detection is called the channel capacity, and it has been recently determined for the case of a MIMO channel (see [10], [11]). Here, and for the sake of completeness, the process of finding the capacity is outlined. If $\mathcal{H}_d^{\mathbf{x}}(f)$ denotes the uncertainty about the input random vector $\mathbf{x}$ of the MIMO channel, and $\mathcal{H}_d^{\mathbf{x}|\mathbf{y}}(f)$ denotes the uncertainty about $\mathbf{x}$ after observing the output vector $\mathbf{y}$, then the reduction in uncertainty is given by

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \mathcal{H}_d^{\mathbf{x}}(f) - \mathcal{H}_d^{\mathbf{x}|\mathbf{y}}(f) \tag{4.24}$$

and it is called the *mutual information* between the input and the output. For a power limited input, the information-theoretic capacity of a channel is defined as the maximum reduction in uncertainty, or maximum mutual information, over all statistical distributions of the input

$$C = \max_{f(\mathbf{x})} \mathcal{I}(\mathbf{x};\mathbf{y}) \tag{4.25}$$

Denoting the joint pdf of the input and output vectors by $h(\mathbf{x},\mathbf{y})$, substitution of (4.9) into (4.23) leads to the symmetry property of $\mathcal{I}(\mathbf{x};\mathbf{y})$

$$\begin{aligned}
\mathcal{I}(\mathbf{x};\mathbf{y}) &= \mathcal{H}_d^{\mathbf{x}}(h) - \mathcal{H}_d^{\mathbf{x}|\mathbf{y}}(h) \\
&= -E_h\left[\log_2 h(\mathbf{x})\right] + E_h\left[\log_2 h(\mathbf{x} \mid \mathbf{y})\right] \\
&= E_h\left[\log_2 \frac{h(\mathbf{x} \mid \mathbf{y})}{h(\mathbf{x})}\right] = E_h\left[\log_2 \frac{h(\mathbf{y} \mid \mathbf{x})}{h(\mathbf{y})}\right] \\
&= \mathcal{I}(\mathbf{y};\mathbf{x}) = \mathcal{H}_d^{\mathbf{y}}(h) - \mathcal{H}_d^{\mathbf{y}|\mathbf{x}}(h)
\end{aligned} \tag{4.26}$$

Now, since the input $\mathbf{x}$ and the output $\mathbf{y}$ of the MIMO channel are related by (4.7), the distribution of $\mathbf{y}$ conditioned on $\mathbf{x}$ is

$$h(\mathbf{y} \mid \mathbf{x}) = f_{\mathbf{n}}(\mathbf{y} - \mathbf{H}\mathbf{x} \mid \mathbf{H};\mathbf{x}) = f_{\mathbf{n}}(\mathbf{n}) \tag{4.27}$$

provided that $\mathbf{n}$ is independent of $\mathbf{x}$. Assuming that the noise vector $\mathbf{n}$ is a normal random variable with covariance matrix $\mathbf{C}_{\mathbf{nn}} = E[\mathbf{n}\mathbf{n}^H] = 2\sigma_n^2\mathbf{I}_{n_R}$, where σ_n^2 is the noise power in the system's bandwidth, the conditional entropy in (4.26) evaluates to

$$\begin{aligned}
\mathcal{H}_d^{\mathbf{y}|\mathbf{x}}(h) &= \mathcal{H}_d^{\mathbf{n}}(f_{\mathbf{n}}) \\
&= \log_2 \det(\pi e \mathbf{C}_{\mathbf{nn}}) = \log_2 \det(2\pi e \sigma_n^2 \mathbf{I}_{n_R})
\end{aligned} \tag{4.28}$$

Therefore, the mutual information in (4.26) will be maximum when the differential entropy of the output $\mathcal{H}_d^{\mathbf{y}}(h)$ is maximized. Hence, from Section 4.3, $\mathbf{y}$ must be a circularly symmetric complex Gaussian vector, and its entropy is given by (4.23), that is

$$\mathcal{H}_d^{\mathbf{y}}(h) = \log_2 \det(\pi e \mathbf{C}_{\mathbf{yy}}) \tag{4.29}$$

Furthermore, the output covariance matrix is

$$\begin{aligned}
\mathbf{C}_{\mathbf{yy}} &= E\left[\mathbf{y}\mathbf{y}^H\right] \\
&= E\left[(\mathbf{H}\mathbf{x} + \mathbf{n})(\mathbf{H}\mathbf{x} + \mathbf{n})^H\right] \\
&= \mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}}
\end{aligned} \tag{4.30}$$

Finally, substituting (4.28) and (4.29) into (4.26), the information-theoretic capacity of the MIMO channel is expressed as

$$\begin{aligned} C &= \log_2 \det(\pi e \mathbf{C_{yy}}) - \log_2 \det(\pi e \mathbf{C_{nn}}) \\ &= \log_2 \det[\pi e(\mathbf{H}\mathbf{C_{xx}}\mathbf{H}^H + \mathbf{C_{nn}})] - \log_2(\pi e \mathbf{C_{nn}}) \\ &= \log_2 \det\left(\mathbf{I}_{n_R} + \frac{1}{2\sigma_n^2}\mathbf{H}\mathbf{C_{xx}}\mathbf{H}^H\right) \text{ bits/vector sample} \end{aligned} \tag{4.31}$$

At first glance, the channel capacity equation does not say much about how large the capacity can really be. However, a closer inspection can unveil some facts. First, if the transmitter is unable to estimate the channel matrix $\mathbf{H}$, the best it can do is to evenly distribute the available power $P_T = n_T E\left[|x_j|^2\right]/2$ among all the antennas, i.e. $\mathbf{C_{xx}} = (2P_T/n_T)\mathbf{I}_{n_T}$. Assuming Nyquist sampling and pulse shaping the capacity is then

$$C = \log_2 \det\left(\mathbf{I}_{n_R} + \frac{P_T}{n_T\sigma_n^2}\mathbf{H}\mathbf{H}^H\right) \quad \text{(bits/s/Hz)} \tag{4.32}$$

Moreover if $\mathbf{H}$ is normalized so that

$$E\left[|h_{ij}|^2\right] = 1 \tag{4.33}$$

then the available power incorporates the average path loss and hence

$$C = \log_2 \det\left(\mathbf{I}_{n_R} + \frac{P}{n_T\sigma_n^2}\mathbf{H}\mathbf{H}^H\right) = \log_2 \det\left(\mathbf{I}_{n_R} + \frac{\gamma}{n_T}\mathbf{H}\mathbf{H}^H\right) \tag{4.34}$$

where P and γ are the average power and average SNR at the input of each receiver branch, respectively. In the limit of very large n_T it follows that

$$C_{n_T\uparrow} \approx \log_2 \det(\mathbf{I}_{n_R} + \gamma\mathbf{\Sigma}) \tag{4.35}$$

where $\mathbf{\Sigma} = E[\mathbf{h}_j\mathbf{h}_j^H]$.

Some notes about the capacity:

- it is directly proportional to the SNR of the input signal in each transmitter antenna, and it depends on the stochastic properties of the channel transfer function;
- interesting enough, it looks very similar to Shannon's capacity formula for the one-dimensional case, which is

$$C = \log_2\left(1 + \frac{P}{\sigma_n^2}|H|^2\right) \tag{4.36}$$

  where $|H|^2$ is the normalized channel power transfer characteristic ($|H|^2 = 1$ in Gaussian channels);
- it is the true channel capacity if and only if the receiver has perfect knowledge of the channel matrix $\mathbf{H}$;
- if the coherence time of the channel (given in (2.73)) is greater than the duration of each symbol transmitted through the channel, it is sufficient to have a single exact estimate of $\mathbf{H}$ for the entire coherence time;
- if the coherence time of the channel is smaller than the duration of each symbol transmitted through the channel, the capacity expression (4.34) is valid provided that

the stochastic process that generates **H** is locally ergodic, i.e. provided that the time-averaged statistics (moments) that can be obtained for the duration of the symbol are unbiased estimates of the statistics of the stochastic process; in this case the effective capacity is the time-average of (4.34), or symbolically $\langle C \rangle_T$;

- the capacity in (4.34) is formally called the *fast-fading capacity* of the MIMO channel, and its expectation the *ergodic capacity*;

Another close look to (4.34), and one distinguishes the factor

$$\mathbf{T} = \frac{1}{n_T}\mathbf{H}\mathbf{H}^H \tag{4.37}$$

Recalling that the coefficients of the channel matrix **H** are stochastic and given by (4.1), it is possible to determine the coefficients of the matrix **T**

$$t_{ij} = \frac{1}{n_T}\sum_{k=1}^{n_T}\alpha_{ik}\alpha_{jk}e^{j(\theta_{ik}-\theta_{jk})} \tag{4.38}$$

Also, invoking the *strong law of large numbers*, when the number of transmitter antennas is large and the h_{ij} are **independent** random variables, the coefficients are

$$\begin{aligned} t_{ij} &= \lim_{n_T\uparrow}\left(\frac{1}{n_T}\sum_{k=1}^{n_T}\alpha_{ik}\alpha_{jk}e^{j(\theta_{ik}-\theta_{jk})}\right) \\ &\approx E\left[\alpha_{ik}\alpha_{jk}e^{j(\theta_{ik}-\theta_{jk})}\right] = \begin{cases} E\left[\alpha_{ik}^2\right] & i = j \\ 0 & i \neq j \end{cases} \end{aligned} \tag{4.39}$$

where $E\left[\alpha_{ik}^2\right] = 1$. Therefore, in the limit of large transmitter antenna arrays, the capacity is upper-bounded by

$$C_{\lim} = n_R \log_2(1+\gamma) \quad n_T \to \infty \tag{4.40}$$

which is n_R times larger than that of the Gaussian SISO channel. When γ is considerably large, say $\gamma \gg 5$ dB, the capacity in (4.40) may be approximated as

$$C_{\lim}^{\gamma\uparrow} = n_R \log_2(1+\gamma) \approx 0.33 n_R \cdot \gamma_{\text{dB}} \qquad \gamma \gg 5 \text{ dB} \; ; \; n_T \to \infty \tag{4.41}$$

showing that the capacity grows linearly with the branch SNR in decibels, and that for each 3 dB increase in the SNR the capacity scales approximately as more n_R bits/cycle. Even though it is obtained under ideal conditions, this is incredibly higher than the Shannon limit for a Gaussian SISO channel.

The result in (4.40) comes as no surprise if one examines the channel response in (4.8). Supposing that the channel matrix entries are independent, the rank of the channel matrix for $n_T \geq n_R$ will be n_R, meaning that if the receiver knows **H**, it will be able to solve the system of equations and find n_R out of n_T input variables, i.e. n_R flows of information. The only setback is the presence of the noise term, which could produce an error when solving the system. More formally, writing the MIMO channel equation as

$$\mathbf{y} = \begin{pmatrix} \mathbf{h}_1 & \dots & \mathbf{h}_T \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_T \end{pmatrix} + \mathbf{n} \tag{4.42}$$

and applying the standard Gauss-Jordan method, one arrives at the equivalent system

$$\begin{pmatrix} \varphi_1(\mathbf{y}) \\ \vdots \\ \varphi_{n_R}(\mathbf{y}) \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 & \varphi_1(\mathbf{h}_{n_R+1}) & \cdots & \varphi_1(\mathbf{h}_{n_T}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \varphi_{n_R}(\mathbf{h}_{n_R+1}) & \cdots & \varphi_{n_R}(\mathbf{h}_{n_T}) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_T \end{pmatrix} + \begin{pmatrix} \varphi_1(\mathbf{n}) \\ \vdots \\ \varphi_{n_R}(\mathbf{n}) \end{pmatrix} \tag{4.43}$$

where $\varphi_i(\mathbf{a} = (a_1,..,a_{n_R}))$ is a linear function of $\mathbf{a}$. This is a consistent system of n_R independent equations

$$\varphi_i(\mathbf{y}) = x_i + \sum_{k=n_R+1}^{n_T} x_k \varphi_i(\mathbf{h}_k) + \varphi_i(\mathbf{n}) \qquad i = 1,...,n_R \tag{4.44}$$

where the middle term can be forced to zero if desired, yielding

$$\varphi_i(\mathbf{y}) = x_i + \varphi_i(\mathbf{n}) \qquad i = 1,...,n_R \tag{4.45}$$

As a result, the receiver is capable of separating n_R noise-limited, identical and independent flows of information, and hence (4.40) follows intuitively.

Further analysis of the MIMO channel capacity is possible by diagonalizing the matrix product $\mathbf{H}\mathbf{H}^H$, which is self-adjoint

$$(\mathbf{H}\mathbf{H}^H)^H = \mathbf{H}\mathbf{H}^H \tag{4.46}$$

Consequently, it has a complete set of orthonormal eigenvectors that can perform the *eigenvalue decomposition*

$$\mathbf{H}\mathbf{H}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \tag{4.47}$$

where $\mathbf{\Lambda} = diag(\lambda_1, \lambda_2,.., \lambda_{n_R})$, with real and non-negative eigenvalues. The capacity expression in (4.34) simplifies to

$$\begin{aligned} C &= \log_2 \det\left(\mathbf{I}_{n_R} + \frac{\gamma}{n_T} \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \right) \\ &= \log_2 \det\left(\mathbf{I}_{n_R} + \frac{\gamma}{n_T} \mathbf{\Lambda} \right) = \log_2 \left[\prod_{k=1}^{n_R} \left(1 + \frac{\gamma}{n_T} \lambda_k \right) \right] \\ &= \sum_{k=1}^{n_R} \log_2 \left(1 + \frac{\gamma}{n_T} \lambda_k \right) \end{aligned} \tag{4.48}$$

When $n_T < n_R$ the factorization $\mathbf{H}\mathbf{H}^H$ is not full-rank, which may bring some difficulties, and therefore the capacity expression in (4.34) should be reexpressed as

$$C = \log_2 \det\left(\mathbf{I}_{n_R} + \frac{\gamma}{n_T} \mathbf{H}\mathbf{H}^H \right) = \log_2 \det\left(\mathbf{I}_{n_T} + \frac{\gamma}{n_T} \mathbf{H}^H\mathbf{H} \right) \tag{4.49}$$

The eigenvalue decomposition is in this case given by

$$\mathbf{H}^H\mathbf{H} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \tag{4.50}$$

leaving the capacity expression in (4.48) in the form

$$C = \sum_{k=1}^{n_T} \log_2 \left(1 + \frac{\gamma}{n_T} \lambda_k \right) \tag{4.51}$$

A capacity expression that covers both cases, $n_T \geq n_R$ and $n_T < n_R$, is

$$C = \sum_{k=1}^{\min(n_R,n_T)} \log_2\left(1 + \frac{\gamma}{n_T}\lambda_k\right) \tag{4.52}$$

From this equation one acknowledges that a MIMO system may be described in terms of $\min(n_R, n_T)$ SISO sub-channels, or *eigenmodes*, and thus the total capacity is the sum of the capacities of these sub-channels. Applying the law of large numbers once more, but this time to (4.49), it easily follows that when $\max(n_R, n_T)$ is very large and the entries of $\mathbf{H}$ are independent, the capacity can be approximated as

$$C \approx \min(n_R, n_T)\log_2\left(1 + \frac{\max(n_R, n_T)}{n_T}\gamma\right) = \min(n_R, n_T)\log_2\left(1 + \frac{n_R}{\min(n_R, n_T)}\gamma\right) \tag{4.53}$$

4.5 *Capacity under Rayleigh Fading*

The initial assumption in the derivation of the information-theoretic capacity of the MIMO channel was that it was modelled by the channel response in (4.7), but nothing was said about the stochastic properties of the channel matrix $\mathbf{H}$. In a mobile wireless environment, matrix $\mathbf{H}$ will be time-variant and its entries will suffer from the fading phenomenon. It is therefore imperative that the channel capacity be given in the average sense

$$C_{avg} = \sum_{k=1}^{\min(n_R,n_T)} E\left[\log_2\left(1 + \frac{\gamma}{n_T}\lambda_k\right)\right] \tag{4.54}$$

where the expectation is over the pdf of each eigenvalue λ_k.

In [45], C_{avg} has been computed at the expense of restricting the quadratic form $\mathbf{H}\mathbf{H}^H$ to a positive definite matrix, and hence $\lambda_k > 0$. It then appeals to some results from mathematical statistics to find the required pdfs, but at the same time circumvents the eigenvalue approach and bases its results on hypergeometric functions of matrix arguments, characteristic functions and infinite sums. Moreover, [45] assumes that the transmit covariance matrix is a multiple of the identity, which unfolds in a mutual information that is not the true capacity when there is transmitter-sided correlation. More recently, in [46], doubly-correlated channels have been studied, but not without the same limitations. Other techniques that circumvent the eigenvalue approach are found in [47], [48], and resort to moment generating functions and direct integration. [49] does not circumvent the eigenvalue approach but fails to give a closed-form expression and only treats the case of no more receiver antennas than transmitter antennas. No such limitations and "workarounds" exist within the author's results.

When the joint pdf of the eigenvalues is a symmetric function (i.e. a function that is unchanged by any permutation of its variables), then (4.54) simplifies to

$$C_{avg} = \min(n_R, n_T)E\left[\log_2\left(1 + \frac{\gamma}{n_T}\lambda\right)\right] \tag{4.55}$$

The order of magnitude of (4.55) can be assessed as follows:

- since the logarithm is a concave function, (4.55) can be upper-bounded as

$$C_{avg} \leq \min(n_R, n_T)\log_2\left(1 + \frac{\gamma}{n_T} E[\lambda]\right) \tag{4.56}$$

- and using the properties of the trace

$$tr(\mathbf{H}^H\mathbf{H}) = tr(\mathbf{H}\mathbf{H}^H) = tr(\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H) = \sum_{k=1}^{\min(n_R,n_T)} \lambda_k \tag{4.57}$$

- yields

$$E[\lambda] = \frac{E[tr(\mathbf{H}^H\mathbf{H})]}{\min(n_R, n_T)} = n_R n_T \,/\, \min(n_R, n_T) \tag{4.58}$$

- therefore the upper bound on the average MIMO channel capacity is

$$C_{avg} \leq \min(n_R, n_T)\log_2\left(1 + \frac{n_R}{\min(n_R, n_T)}\gamma\right) \tag{4.59}$$

which is identical to (4.53). Besides defining a bound to the capacity, it also tells us that when $n = n_R = n_T$ the average capacity is always less than n times the Shannon capacity. However, it does not yet unveil the essence of the true capacities in Rayleigh fading environments. The calculation of the exact capacities will be presented in the following sections.

4.5.1 CHANNEL MATRIX DISTRIBUTION

For a better understanding of the subsequent analysis the author recommends the readings [50], [51], [52] and [53]. Finding the distribution of the eigenvalues of $\mathbf{H}$ is a somewhat intricate process which starts by outlining some particular assumptions about the channel matrix:

1. the number of multipath components between the j-th transmitter antenna and the i-th receiver antenna is very large and there is a single cluster of scatterers between the transmitter and receiver; therefore, by the central limit theorem, the entries h_{ij} of matrix $\mathbf{H}$ can be statistically characterized as zero-mean complex Gaussian random variables, with jointly Gaussian in-phase and quadrature components;

2. the h_{ij}'s corresponding to different transmitter antennas are statistically independent, thus uncorrelated, and the h_{ij}'s corresponding to different receiver antennas are correlated random variables;

In view of the two conditions above, the columns $(\mathbf{h}_1 \mid \mathbf{h}_2 \mid .. \mid \mathbf{h}_{n_T})$ of the channel matrix $\mathbf{H}$ can be regarded as independent samples from a complex multivariate zero-mean Gaussian distribution, that is, for every $1 \leq j \leq n_T$, $\mathbf{h}_j \sim \widetilde{\mathcal{N}}_{n_R}(0, \mathbf{\Sigma})$, where $\mathbf{\Sigma} = E[\mathbf{h}_j\mathbf{h}_j^H]$ is the covariance matrix of the column entries.

Determining the pdf of $\mathbf{H}$ is equivalent to finding the joint pdf of its columns, and so

it is useful to introduce the notation

$$\mathbf{h} = vec(\mathbf{H}) = \begin{pmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_{n_T} \end{pmatrix} \tag{4.60}$$

The covariance matrix of the entries of $\mathbf{H}$ is then

$$\mathbf{C_{hh}} = E[\mathbf{h}\mathbf{h}^H] = \mathbf{I}_{n_T} \otimes \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{\Sigma} \end{pmatrix} \tag{4.61}$$

where $\otimes$ denotes the Kronecker product of matrices. The entries of $\mathbf{C_{hh}}$ are therefore the entries of $\mathbf{\Sigma}$. They are given by

$$[\Sigma]_{ir} = E\left[h_{ij}(t)h_{rj}^*(t)\right] \tag{4.62}$$

It is not difficult to evaluate these since the $h_{rj}(t)$'s can be obtained from the $h_{ij}(t)$'s. Remember that the indexes i and r are identifiers for the different receiver antennas (branches). From (4.3) we have

$$h_{ij}(t) = \sum_{k=0}^{L_j-1} a_{jk}\, e^{j(\varphi_{ijk}+\omega_{jk}t)} \tag{4.63}$$

and therefore

$$E\left[h_{ij}(t)h_{rj}^*(t)\right] = \sum_{k,p=0}^{L_j-1} E\left[a_{jk}a_{jp}\right] E\left[e^{j(\varphi_{ijk}-\varphi_{rjp}+(\omega_{jk}-\omega_{jp})t)}\right] \tag{4.64}$$

Assuming an environment with uncorrelated scattering, (4.64) is clearly non-zero only for $k = p$, which reduces it to

$$\begin{aligned} E\left[h_{ij}(t)h_{rj}^*(t)\right] &= \sum_{k=0}^{L_j-1} E\left[a_{jk}^2\right] E\left[e^{j(\varphi_{ijk}-\varphi_{rjk})}\right] = L_j E\left[a_{jk}^2\right] E\left[e^{j(\varphi_{ijk}-\varphi_{rjk})}\right] \\ &= \overbrace{E\left[\left|h_{ij}(t)\right|^2\right]}^{=1} E\left[e^{j(\varphi_{ijk}-\varphi_{rjk})}\right] = E\left[e^{j(\varphi_{ijk}-\varphi_{rjk})}\right] \end{aligned} \tag{4.65}$$

where the normalization of (4.33) is used as a simplification. Now refer to the situation depicted in Figure 4.1, where the wavefront of a single propagation path impinges on two of the receiver antennas, say, antenna i and antenna r. Both antennas travel in the same direction, but within parallel paths. ϕ_{jk} is the angle between the direction of motion and the k-th incoming wave, θ_{ir} is the angle between the direction of motion and the line joining the two antennas, d_{ir} is the distance between the antennas, and $\varphi_{ijk} - \varphi_{rjk}$ is the k-th

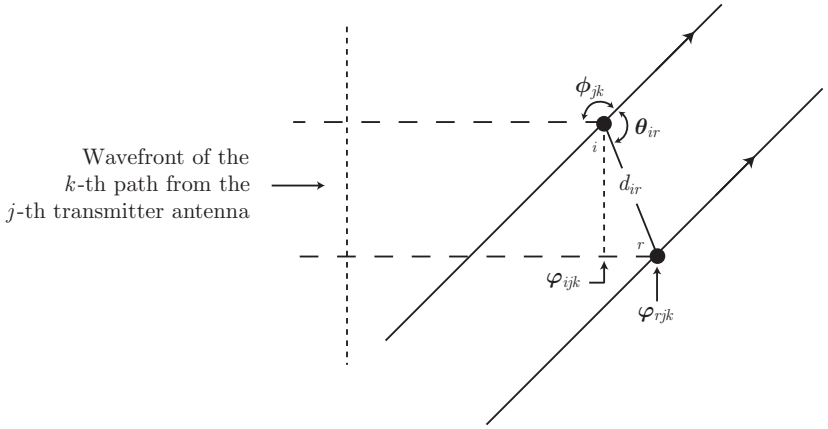

FIGURE 4.1 Illustrating the phase difference between different antennas of a moving receiver.

path phase difference between the antennas. The phase difference can be written as

$$\varphi_{ijk} - \varphi_{rjk} = -\beta d_{ir} \cos(\phi_{jk} + \theta_{ir}) \tag{4.66}$$

and a direct substitution into (4.65) results in

$$E\big[h_{ij}(t)h_{rj}^{*}(t)\big] = E\big[\cos\big(\beta d_{ir}\cos(\phi_{jk} + \theta_{ir})\big)\big] - jE\big[\sin\big(\beta d_{ir}\cos(\phi_{jk} + \theta_{ir})\big)\big] \tag{4.67}$$

From this expression, it is easy to see that the following relations are true

$$\begin{aligned} E\big[h_{ijI}(t)h_{rjI}(t)\big] &= E\big[h_{ijQ}(t)h_{rjQ}(t)\big] = \frac{1}{2}E\big[\cos\big(\beta d_{ir}\cos(\phi_{jk} + \theta_{ir})\big)\big] \\ E\big[h_{ijI}(t)h_{rjQ}(t)\big] &= -E\big[h_{ijQ}(t)h_{rjI}(t)\big] = -\frac{1}{2}E\big[\sin\big(\beta d_{ir}\cos(\phi_{jk} + \theta_{ir})\big)\big] \end{aligned} \tag{4.68}$$

where $h_{ij} = h_{ijI} + jh_{ijQ}$. The elements at the left are merely the entries of the covariance matrices of the real $\mathbf{h}_{jI}$ and imaginary $\mathbf{h}_{jQ}$ parts of $\mathbf{h}_j$. It is now a simple assertion to write the identities between the covariance matrices

$$\begin{aligned} \mathbf{C}_{\mathbf{h}_{jI}\mathbf{h}_{jI}} &= \mathbf{C}_{\mathbf{h}_{jQ}\mathbf{h}_{jQ}} = \mathbf{C}^{T}_{\mathbf{h}_{jI}\mathbf{h}_{jI}} \\ \mathbf{C}_{\mathbf{h}_{jI}\mathbf{h}_{jQ}} &= -\mathbf{C}_{\mathbf{h}_{jQ}\mathbf{h}_{jI}} = -\mathbf{C}^{T}_{\mathbf{h}_{jI}\mathbf{h}_{jQ}} \end{aligned} \tag{4.69}$$

Then again, these relations constitute the definition (4.15) of a circularly symmetric vector, and therefore the pdf of $\mathbf{h}$ is identical to (4.17), rewritten as

$$\begin{aligned} f_{\mathbf{h}}(\mathbf{h}) &= \prod_{j=1}^{n_T} |\det(\pi\boldsymbol{\Sigma})|^{-1} \exp\big(-\mathbf{h}_j^{H}\boldsymbol{\Sigma}^{-1}\mathbf{h}_j\big) \\ &= |\det(\pi\mathbf{C}_{\mathbf{hh}})|^{-1} \exp\big(-\mathbf{h}^{H}\mathbf{C}_{\mathbf{hh}}^{-1}\mathbf{h}\big) \end{aligned} \tag{4.70}$$

Equation (4.70) can be put to another form by noticing that

$$\begin{aligned} \mathbf{h}^{H}\mathbf{C}_{\mathbf{hh}}^{-1}\mathbf{h} &= \sum_{j=1}^{n_T} \mathbf{h}_j^{H}\boldsymbol{\Sigma}^{-1}\mathbf{h}_j = \sum_{j=1}^{n_T} \big[\mathbf{H}^{H}\boldsymbol{\Sigma}^{-1}\mathbf{H}\big]_{jj} \\ &= tr\big(\mathbf{H}^{H}\boldsymbol{\Sigma}^{-1}\mathbf{H}\big) = tr\big(\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{H}^{H}\big) \end{aligned} \tag{4.71}$$

where $tr(\cdot)$ is the trace function. Hence, the pdf of $\mathbf{H}$ is

$$f_{\mathbf{H}}(\mathbf{H}) = |\det(\pi\boldsymbol{\Sigma})|^{-n_T} \exp\big[-tr\big(\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{H}^{H}\big)\big] \tag{4.72}$$

4.5.2 JACOBIANS AND EXTERIOR PRODUCTS

Let $\mathbf{x}$ and $\mathbf{y}$ be two n-dimensional **real** vectors such that the transformation (or equivalently, change of coordinates) $\mathbf{y} = \mathbf{T}(\mathbf{x})$ is valid. Then, there exists an associated linear transformation of differentials

$$d\mathbf{y} = \mathbf{J}(\mathbf{y} \rightarrow \mathbf{x})d\mathbf{x} \tag{4.73}$$

associated with it that conveys the relation between the differential lengths in the two coordinate systems. $\mathbf{J}(\mathbf{y} \rightarrow \mathbf{x})$ is the Jacobian matrix

$$\mathbf{J}(\mathbf{y} \rightarrow \mathbf{x}) = \begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y_n}{\partial x_1} & \cdots & \dfrac{\partial y_n}{\partial x_n} \end{pmatrix} \tag{4.74}$$

of the transformation, and it equals $\mathbf{T}$ in the case of linear transformations $\mathbf{y} = \mathbf{T}\mathbf{x}$. The volume element (ignoring sign) in the coordinate system of vector $\mathbf{x}$ equals the exterior product of all the distinct elements in $d\mathbf{x}$

$$(d\mathbf{x})^{\wedge} = dx_1 \wedge dx_2 \wedge .. \wedge dx_n = \bigwedge_{i=1}^{n} dx_i \tag{4.75}$$

where it is understood that the differential lengths are implicitly associated with a certain direction in the coordinate system, and thus can be regarded as differential vectors. The exterior product has the following basic rules

$$dx \wedge dy = -dy \wedge dx;\ dx \wedge dx = -dx \wedge dx = 0 \tag{4.76}$$

which means that products of repeated differentials disappear. Denoting the elements of the Jacobian matrix by t_{ij}, the volume elements in the two coordinate systems are related by

$$\begin{aligned} (d\mathbf{y})^{\wedge} &= \bigwedge_{i=1}^{m} dy_i = \bigwedge_{i=1}^{m} \sum_{j=1}^{n} t_{ij} dx_j \\ &= \sum_{p=(p_1,p_2,..,p_n)} t_{1p_1} t_{2p_2} .. t_{np_n} dx_{p_1} \wedge dx_{p_2} \wedge .. \wedge dx_{p_n} \end{aligned} \tag{4.77}$$

where the summation is taken over the $n!$ permutations $(p_1, p_2, .., p_n)$ of $(1, 2, .., n)$. Using the exterior product rules in (4.76) it is possible to write

$$\begin{aligned} (d\mathbf{y})^{\wedge} &= \sum_{p=(p_1,p_2,..,p_n)} \sigma(p)\, t_{1p_1} t_{2p_2} .. t_{np_n} dx_1 \wedge dx_2 \wedge .. \wedge dx_n \\ &= \left[\sum_{p=(p_1,p_2,..,p_n)} \sigma(p)\, t_{1p_1} t_{2p_2} .. t_{np_n} \right] \bigwedge_{j=1}^{n} dx_j \end{aligned} \tag{4.78}$$

where

$$\sigma(p) = \begin{cases} +1 & \text{if } p \text{ is an } even \text{ permutation of } (1, 2, .., n) \\ -1 & \text{if } p \text{ is an } odd \text{ permutation of } (1, 2, .., n) \end{cases} \tag{4.79}$$

The first factor of (4.78) is simply the determinant of the Jacobian matrix $\mathbf{J}(\mathbf{y} \rightarrow \mathbf{x})$, and

so it follows that the equation

$$\begin{aligned}(d\mathbf{y})^{\wedge} &= \det[\mathbf{J}(\mathbf{y}\to\mathbf{x})](d\mathbf{x})^{\wedge}\\ &= J(\mathbf{y}\to\mathbf{x})(d\mathbf{x})^{\wedge}\end{aligned} \tag{4.80}$$

establishes the relation between the volume elements of the two coordinate systems. The Jacobian determinant of the transformation $\mathbf{y}=\mathbf{T}(\mathbf{x})$ is denoted here by $J(\mathbf{y}\to\mathbf{x})$.

The Jacobian determinant in the **complex** case (complex vectors and transformation matrices) is slightly different from (4.80). For simplicity, only **linear** transformations will be discussed. Writing

$$\begin{aligned}d\tilde{\mathbf{y}} &= \tilde{\mathbf{T}}d\tilde{\mathbf{x}}\\ d\mathbf{y}_r + jd\mathbf{y}_i &= (\mathbf{T}_r + j\mathbf{T}_i)(d\mathbf{x}_r + jd\mathbf{x}_i)\\ &= (\mathbf{T}_r d\mathbf{x}_r - \mathbf{T}_i d\mathbf{x}_i) + j(\mathbf{T}_i d\mathbf{x}_r + \mathbf{T}_r d\mathbf{x}_i)\end{aligned} \tag{4.81}$$

and separating the real and imaginary parts, gives

$$\begin{pmatrix} d\mathbf{y}_r \\ d\mathbf{y}_i \end{pmatrix} = \begin{pmatrix} \mathbf{T}_r & -\mathbf{T}_i \\ \mathbf{T}_i & \mathbf{T}_r \end{pmatrix}\begin{pmatrix} d\mathbf{x}_r \\ d\mathbf{x}_i \end{pmatrix} \tag{4.82}$$

Therefore, using (4.80) and observing that $(d\tilde{\mathbf{y}}) = (d\mathbf{y}_r)^{\wedge}\wedge(d\mathbf{y}_i)^{\wedge}$, the relation between the right and left volume elements is

$$(d\tilde{\mathbf{y}})^{\wedge} = J(\tilde{\mathbf{y}}\to\tilde{\mathbf{x}})(d\tilde{\mathbf{x}})^{\wedge} = \det\begin{pmatrix} \mathbf{T}_r & -\mathbf{T}_i \\ \mathbf{T}_i & \mathbf{T}_r \end{pmatrix}(d\tilde{\mathbf{x}})^{\wedge} \tag{4.83}$$

Finally, invoking the invariance of the determinant under elementary row and column operations, we may multiply the first row by j and add it to the second row,

$$(d\tilde{\mathbf{y}})^{\wedge} = \det\begin{pmatrix} \mathbf{T}_r & -\mathbf{T}_i \\ j\mathbf{T}_r + \mathbf{T}_i & \mathbf{T}_r - j\mathbf{T}_i \end{pmatrix}(d\tilde{\mathbf{x}})^{\wedge} \tag{4.84}$$

and then multiply the second column by $-j$ and add it to the first column, yielding

$$\begin{aligned}(d\tilde{\mathbf{y}})^{\wedge} &= \det\begin{pmatrix} \mathbf{T}_r + j\mathbf{T}_i & -\mathbf{T}_i \\ \mathbf{0} & \mathbf{T}_r - j\mathbf{T}_i \end{pmatrix}(d\tilde{\mathbf{x}})^{\wedge}\\ &= \det(\mathbf{T}_r + j\mathbf{T}_i)\det(\mathbf{T}_r - j\mathbf{T}_i)(d\tilde{\mathbf{x}})^{\wedge}\\ &= \left|\det(\mathbf{T}_r + j\mathbf{T}_i)\right|^2\langle d\tilde{\mathbf{x}}\rangle = \left|\det\tilde{\mathbf{T}}\right|^2(d\tilde{\mathbf{x}})^{\wedge}\end{aligned} \tag{4.85}$$

This result for complex vector transformations easily generalizes to complex linear transformations of complex matrices $\tilde{\mathbf{Y}}=\tilde{\mathbf{T}}\tilde{\mathbf{X}}$. Letting n denote the number of columns of the matrices $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$, it follows that

$$\begin{aligned}\left(d\tilde{\mathbf{Y}}\right)^{\wedge} &= \bigwedge_{i=1}^{n}(d\tilde{\mathbf{y}}_i)^{\wedge} = \bigwedge_{i=1}^{n}\left|\det\tilde{\mathbf{T}}\right|^2(d\tilde{\mathbf{x}}_i)^{\wedge}\\ &= \left|\det\tilde{\mathbf{T}}\right|^{2n}\bigwedge_{i=1}^{n}(d\tilde{\mathbf{x}}_i)^{\wedge} = \left|\det\tilde{\mathbf{T}}\right|^{2n}\left(d\tilde{\mathbf{X}}\right)^{\wedge}\end{aligned} \tag{4.86}$$

and accordingly $J\left(\tilde{\mathbf{Y}}\to\tilde{\mathbf{X}}\right) = \left|\det\tilde{\mathbf{T}}\right|^{2n}$.

Now, let $\mathbf{X}$ and $\mathbf{Y}$ be generic random matrices (real or complex) and let $f_{\mathbf{X}}(\mathbf{X})$ and $f_{\mathbf{Y}}(\mathbf{Y})$ be the respective joint pdfs of their random entries. Define a *one-to-one* generic trans-

formation $\mathbf{Y} = \mathbf{T}(\mathbf{X})$ between the random matrices, and write the relationship between the volume elements, i.e.

$$(d\mathbf{Y})^{\wedge} = J(\mathbf{Y} \to \mathbf{X})(d\mathbf{X})^{\wedge} \tag{4.87}$$

where $(d\mathbf{X})^{\wedge}$ and $(d\mathbf{Y})^{\wedge}$ are the **positive** volume elements of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Then it must be the case that

$$f_{\mathbf{Y}}(\mathbf{Y})(d\mathbf{Y})^{\wedge} = f_{\mathbf{X}}(\mathbf{X})(d\mathbf{X})^{\wedge} \tag{4.88}$$

and in the same manner

$$f_{\mathbf{Y}}(\mathbf{Y}) = f_{\mathbf{X}}(\mathbf{X})\frac{(d\mathbf{X})^{\wedge}}{(d\mathbf{Y})^{\wedge}} = f_{\mathbf{X}}(\mathbf{T}^{-1}(\mathbf{Y}))J(\mathbf{X} \to \mathbf{Y}) \tag{4.89}$$

This means that the statistics of $\mathbf{Y}$ can be determined in terms of the statistics of $\mathbf{X}$ by the expression (4.89). Together with (4.85), this important result leads to a fundamental property of complex Gaussian distributions such as (4.72): they are invariant under right-sided unitary transformations $\mathbf{Y} = \mathbf{XQ}$

$$f_{\mathbf{Y}}(\mathbf{Y}) = f_{\mathbf{X}}(\mathbf{YQ}^{H})\frac{(d\mathbf{X})^{\wedge}}{(d\mathbf{Y})^{\wedge}} = f_{\mathbf{X}}(\mathbf{Y})|\det\mathbf{Q}|^{2n} = f_{\mathbf{X}}(\mathbf{Y}) \tag{4.90}$$

The Jacobians of a few particular cases of linear matrix transformations will be useful in the subsequent sections, and so the details of their derivation will be given here.

1. Jacobian of the Cholesky factorization - let $\mathbf{S} = \mathbf{R}^{H}\mathbf{R}$ be the Cholesky factorization of the $n \times n$ complex positive definite matrix $\mathbf{S}$, where $\mathbf{R}$ is an $n \times n$ upper-triangular matrix with positive and real diagonal entries. In matrix form, $\mathbf{S}$ is given by

$$\mathbf{S} = \begin{pmatrix} r_{11}^2 & r_{11}^* r_{12} & r_{11}^* r_{13} & \cdots \\ \times & r_{12}^* r_{12} + r_{22}^2 & r_{12}^* r_{13} + r_{22}^* r_{23} & \cdots \\ \times & \times & r_{13}^* r_{13} + r_{23}^* r_{23} + r_{33}^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.91}$$

Differentiating each element of $\mathbf{S}$

$$d\mathbf{S} = \begin{pmatrix} 2r_{11}dr_{11} & dr_{11}r_{12} + r_{11}dr_{12} & dr_{11}r_{13} + r_{11}dr_{13} & \cdots \\ \times & dr_{12}^* r_{12} + r_{12}^* dr_{12} + 2r_{22}dr_{22} & dr_{12}^* r_{13} + r_{12}^* dr_{13} + dr_{22}r_{23} + r_{22}dr_{23} & \cdots \\ \times & \times & dr_{13}^* r_{13} + r_{13}^* dr_{13} + dr_{23}^* r_{23} + r_{23}^* dr_{23} + 2r_{33}dr_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.92}$$

it is easy to assess that, since only the differentials elements dr_{ij} that do not appear in a previous column or row prevail in the exterior product, solely the last term in each matrix entry must be accounted for

$$d\mathbf{S} = \begin{pmatrix} 2r_{11}dr_{11} & r_{11}dr_{12} + * & r_{11}dr_{13} + * & \cdots \\ \times & 2r_{22}dr_{22} + * & r_{22}dr_{23} + * & \cdots \\ \times & \times & 2r_{33}dr_{33} + * & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.93}$$

Finally, taking exterior products and noticing that $dr_{ij} = [dr_{ij}]_r + j[dr_{ij}]_c$ is a complex differential for $i < j$, and thus $(r_{ii}dr_{ij})^{\wedge} = r_{ii}^2[dr_{ij}]_r \wedge [dr_{ij}]_c = r_{ii}^2(dr_{ij})^{\wedge}$, the volume element of matrix $\mathbf{S}$ is

$$\begin{aligned}(d\mathbf{S})^{\wedge} &= \bigwedge_{i\le j}^{n} ds_{ij} = 2^n \bigwedge_{i\le j}^{n} (r_{ii}dr_{ij})^{\wedge} \\ &= 2^n r_{11}^{2n-1} r_{22}^{2n-3} \ldots r_{n_R n_R} \left(\bigwedge_{i=1}^{n} dr_{ii}\right) \wedge \left(\bigwedge_{i<j}^{n} (dr_{ij})^{\wedge}\right) \\ &= 2^n \prod_{k=1}^{n} r_{kk}^{(2n-2k+1)} (d\mathbf{R})^{\wedge}\end{aligned} \tag{4.94}$$

The Jacobian determinant is

$$J(\mathbf{S} \to \mathbf{R}) = \frac{(d\mathbf{S})^{\wedge}}{(d\mathbf{R})^{\wedge}} = 2^n \prod_{k=1}^{n} r_{kk}^{(2n-2k+1)} \tag{4.95}$$

2. Jacobian of the QR factorization - let $\mathbf{S} = \mathbf{QR}$ be the QR factorization of the complex $m \times n$ $(m \ge n)$ matrix $\mathbf{S}$ of rank n, where $\mathbf{Q}$ is an $m \times n$ unitary matrix and $\mathbf{R}$ is an $n \times n$ upper-triangular matrix. The identity $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$ defines n^2 independent equations. There are two options to make the factorization unique:

- the entries of matrix $\mathbf{Q}$ are all complex, which means that $\mathbf{Q}$ has $2mn - n^2$ independent parameters; the diagonal elements of matrix $\mathbf{R}$ are real, and hence $\mathbf{R}$ has n^2 independent parameters, making a total of $2mn$ independent parameters with $\mathbf{Q}$;
- the diagonal elements of $\mathbf{Q}$ are real, which means that $\mathbf{Q}$ has $2mn - n^2 - n$ independent parameters; the entries of matrix $\mathbf{R}$ are all complex, and hence $\mathbf{R}$ has $n^2 + n$ independent parameters, making a total of $2mn$ independent parameters with $\mathbf{Q}$;

Only the Jacobian of the first case will be derived, since the second case follows by the same lines. Differentiating $\mathbf{S}$ and introducing the volume element notation, yields

$$(d\mathbf{S})^{\wedge} = (d\mathbf{Q}\mathbf{R} + \mathbf{Q}d\mathbf{R})^{\wedge} \tag{4.96}$$

In the present form, this expression makes the desired calculation exceedingly complicated. Nevertheless, a worthy simplification is introduced by the following fact:

- matrix $\mathbf{Q}$ possesses n orthogonal columns and can be extended with $m - n$ columns (i.e. the orthogonal complement) to form an $m \times m$ unitary matrix $\overline{\mathbf{Q}} = \begin{bmatrix}\mathbf{Q} & \mathbf{Q}_{\perp}\end{bmatrix}$, such that using (4.86) it is possible to write

$$\begin{aligned}\left(\overline{\mathbf{Q}}d\mathbf{S}\right)^{\wedge} &= \left|\det \overline{\mathbf{Q}}\right|^{2n} (d\mathbf{S})^{\wedge} \\ &= (d\mathbf{S})^{\wedge}\end{aligned} \tag{4.97}$$

and therefore

$$\begin{aligned}(d\mathbf{S})^{\wedge} &= \left(\overline{\mathbf{Q}}^H d\mathbf{S}\right)^{\wedge} = \begin{pmatrix}\mathbf{Q}^H d\mathbf{S} \\ \mathbf{Q}_{\perp}^H d\mathbf{S}\end{pmatrix}^{\wedge} \\ &= (\mathbf{Q}^H d\mathbf{Q}\mathbf{R} + d\mathbf{R})^{\wedge} \wedge (\mathbf{Q}_{\perp}^H d\mathbf{Q}\mathbf{R})^{\wedge}\end{aligned} \tag{4.98}$$

Moreover, because $\mathbf{Q}^H d\mathbf{Q}$ is skew-hermitian ($\mathbf{Q}^H d\mathbf{Q} = -d\mathbf{Q}^H \mathbf{Q}$), matrix $\mathbf{Q}^H d\mathbf{Q}\mathbf{R}$ is given by

$$\mathbf{Q}^H d\mathbf{Q}\mathbf{R} = \begin{pmatrix} r_{11}\mathbf{q}_1^H d\mathbf{q}_1 & r_{12}\mathbf{q}_1^H d\mathbf{q}_1 - r_{22}(\mathbf{q}_2^H d\mathbf{q}_1)^* & r_{13}\mathbf{q}_1^H d\mathbf{q}_1 - r_{23}(\mathbf{q}_2^H d\mathbf{q}_1)^* - r_{33}(\mathbf{q}_3^H d\mathbf{q}_1)^* & \cdots \\ r_{11}\mathbf{q}_2^H d\mathbf{q}_1 & r_{12}\mathbf{q}_2^H d\mathbf{q}_1 + r_{22}\mathbf{q}_2^H d\mathbf{q}_2 & r_{13}\mathbf{q}_2^H d\mathbf{q}_1 + r_{23}\mathbf{q}_2^H d\mathbf{q}_2 - r_{33}(\mathbf{q}_3^H d\mathbf{q}_2)^* & \cdots \\ r_{11}\mathbf{q}_3^H d\mathbf{q}_1 & r_{12}\mathbf{q}_3^H d\mathbf{q}_1 + r_{22}\mathbf{q}_3^H d\mathbf{q}_2 & r_{13}\mathbf{q}_3^H d\mathbf{q}_1 + r_{23}\mathbf{q}_3^H d\mathbf{q}_2 + r_{33}\mathbf{q}_3^H d\mathbf{q}_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.99}$$

It is obvious looking at (4.99) that only the last term of the diagonal and subdiagonal entries is not redundant, in the sense that the $\mathbf{q}_i^H d\mathbf{q}_j$ factor does not appear in a previous column and products of repeated differentials are zero

$$\left(\mathbf{q}_i^H d\mathbf{q}_j\right)^\wedge \wedge \left(\mathbf{q}_i^H d\mathbf{q}_j\right)^\wedge = \left(\mathbf{q}_i^H d\mathbf{q}_j\right)^\wedge \wedge \left((\mathbf{q}_i^H d\mathbf{q}_j)^*\right)^\wedge = 0 \tag{4.100}$$

In addition, the (not redundant) diagonal terms are purely imaginary, and thus independent of the real diagonal elements of $d\mathbf{R}$. In other words, $\mathbf{Q}^H d\mathbf{Q}\mathbf{R}$ and $d\mathbf{R}$ make a perfect fit in (4.98), thereby reducing it to

$$(d\mathbf{S})^\wedge = (\mathbf{Q}^H d\mathbf{Q}\mathbf{R})^\wedge \wedge (\mathbf{Q}_\perp^H d\mathbf{Q}\mathbf{R})^\wedge \wedge (d\mathbf{R})^\wedge \tag{4.101}$$

Hence, denoting the columns of $\overline{\mathbf{Q}}$ by $\overline{\mathbf{q}}_i$, the first factor is given by

$$\begin{aligned} (\mathbf{Q}^H d\mathbf{Q}\mathbf{R})^\wedge &= \bigwedge_{i\geq j}^{n} \left(\mathbf{q}_i^H d\mathbf{q}_j r_{jj}\right)^\wedge = \bigwedge_{i\geq j}^{n} \left(\overline{\mathbf{q}}_i^H d\mathbf{q}_j r_{jj}\right)^\wedge \\ &= \bigwedge_{i>j}^{n} \left(\overline{\mathbf{q}}_i^H d\mathbf{q}_j\right)^\wedge r_{jj}^2 \bigwedge_{i=j}^{n} \left(\overline{\mathbf{q}}_i^H d\mathbf{q}_j\right)^\wedge r_{jj} \\ &= \prod_{k=1}^{n} r_{kk}^{(2n-2k+1)} \bigwedge_{i\geq j}^{n} \left(\overline{\mathbf{q}}_i^H d\mathbf{q}_j\right)^\wedge \end{aligned} \tag{4.102}$$

The second factor in (4.101) follows from (4.86)

$$\begin{aligned} (\mathbf{Q}_\perp^H d\mathbf{Q}\mathbf{R})^\wedge &= |\det(\mathbf{R})|^{2m-2n} (\mathbf{Q}_\perp^H d\mathbf{Q})^\wedge \\ &= \prod_{k=1}^{n} r_{kk}^{(2m-2n)} \bigwedge_{j=1}^{n} \bigwedge_{i=n+1}^{m} \left(\overline{\mathbf{q}}_i^H d\mathbf{q}_j\right)^\wedge \end{aligned} \tag{4.103}$$

Substituting (4.102) and (4.103) into (4.101), and evaluating the product, gives

$$\begin{aligned} (d\mathbf{S})^\wedge &= \prod_{k=1}^{n} r_{kk}^{(2n-2k+1)} \bigwedge_{j=1}^{n} \bigwedge_{i=j}^{m} \left(\overline{\mathbf{q}}_i^H d\mathbf{q}_j\right)^\wedge \wedge (d\mathbf{R})^\wedge \\ &\equiv \prod_{k=1}^{n} r_{kk}^{(2m-2k+1)} \left(\overline{\mathbf{Q}}^H d\mathbf{Q}\right)^\wedge \wedge (d\mathbf{R})^\wedge = \prod_{k=1}^{n} r_{kk}^{(2m-2k+1)} (d\mathbf{Q})^\wedge \wedge (d\mathbf{R})^\wedge \end{aligned} \tag{4.104}$$

In words, the Jacobian determinant of the one-to-one QR factorization with strictly complex $\mathbf{Q}$ is

$$J(\mathbf{S} \to \mathbf{Q}, \mathbf{R}) = \frac{(d\mathbf{S})^\wedge}{(d\mathbf{Q})^\wedge \wedge (d\mathbf{R})^\wedge} = \prod_{k=1}^{n} r_{kk}^{(2m-2k+1)} \tag{4.105}$$

Going by the same steps but now with a strictly complex $\mathbf{R}$, and verifying that the factors $\mathbf{q}_i^H d\mathbf{q}_i$ are redundant and so may be ignored in the exterior products, it follows that the

Jacobian is

$$J(\mathbf{S} \rightarrow \mathbf{Q},\mathbf{R}) = \frac{(d\mathbf{S})^{\wedge}}{(d\mathbf{Q})^{\wedge} \wedge (d\mathbf{R})^{\wedge}} = \prod_{k=1}^{n} |r_{kk}|^{2(m-k)} \tag{4.106}$$

3. Jacobian of the transformation $\mathbf{S} = \mathbf{B}^H\mathbf{A}\mathbf{B}$ - let $\mathbf{B}$ be a fixed complex nonsingular $n \times n$ matrix and $\mathbf{A}$ and $\mathbf{S}$ be two hermitian positive definite $n \times n$ matrices, such that $\mathbf{S} = \mathbf{B}^H\mathbf{A}\mathbf{B}$. Since $\mathbf{A}$ and $\mathbf{S}$ are positive definite they have a Cholesky factorization, namely $\mathbf{S} = \mathbf{R}_1^H\mathbf{R}_1$ and $\mathbf{A} = \mathbf{R}_2^H\mathbf{R}_2$, and for that reason the transformation can be rewritten as

$$\begin{aligned} \mathbf{S} &= \mathbf{B}^H\mathbf{A}\mathbf{B} \\ \mathbf{R}_1^H\mathbf{R}_1 &= \mathbf{B}^H\mathbf{R}_2^H\mathbf{R}_2\mathbf{B} \end{aligned} \tag{4.107}$$

Introduce two $n \times n$ unitary matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ and write (4.107) as

$$\begin{aligned} \mathbf{S} &= \mathbf{R}_1^H\mathbf{Q}_1^H\mathbf{Q}_1\mathbf{R}_1 = \mathbf{B}^H\mathbf{R}_2^H\mathbf{Q}_2^H\mathbf{Q}_2\mathbf{R}_2\mathbf{B} \\ &= (\mathbf{Q}_1\mathbf{R}_1)^H\mathbf{Q}_1\mathbf{R}_1 = (\mathbf{Q}_2\mathbf{R}_2\mathbf{B})^H\mathbf{Q}_2\mathbf{R}_2\mathbf{B} \end{aligned} \tag{4.108}$$

Matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ can be properly chosen to make the relation between the left and write members an obvious one, that is $\mathbf{Q}_1\mathbf{R}_1 = \mathbf{Q}_2\mathbf{R}_2\mathbf{B}$. Thus, one may write the following relations

$$\mathbf{S} = \mathbf{R}_1^H\mathbf{R}_1 \;;\; \mathbf{X}_1 = \mathbf{Q}_1\mathbf{R}_1 \;;\; \mathbf{Y}_1 = \mathbf{Q}_2\mathbf{R}_2 \;;\; \mathbf{A} = \mathbf{R}_2^H\mathbf{R}_2 \tag{4.109}$$

The expression for the Jacobian determinant follows immediately

$$J(\mathbf{S} \rightarrow \mathbf{A}) = \frac{(d\mathbf{S})^{\wedge}}{(d\mathbf{A})^{\wedge}} = \frac{(d\mathbf{S})^{\wedge}}{(d\mathbf{R}_1)^{\wedge}}\frac{(d\mathbf{R}_1)^{\wedge}}{(d\mathbf{X}_1)^{\wedge}}\frac{(d\mathbf{X}_1)^{\wedge}}{(d\mathbf{Y}_1)^{\wedge}}\frac{(d\mathbf{Y}_1)^{\wedge}}{(d\mathbf{R}_2)^{\wedge}}\frac{(d\mathbf{R}_2)^{\wedge}}{(d\mathbf{A})^{\wedge}} \tag{4.110}$$

Clearly $\mathbf{X}_1 = \mathbf{Y}_1\mathbf{B}$, so from (4.86) it follows that

$$\frac{(d\mathbf{X}_1)^{\wedge}}{(d\mathbf{Y}_1)^{\wedge}} = |\det\mathbf{B}|^{2n} \tag{4.111}$$

Also, using the Jacobians of the Cholesky and QR factorizations, (4.95) and (4.105) respectively with $n = m$, leads to the following equalities

$$\frac{(d\mathbf{S})^{\wedge}}{(d\mathbf{R}_1)^{\wedge}}\frac{(d\mathbf{R}_1)^{\wedge}}{(d\mathbf{X}_1)^{\wedge}} = \frac{2^n}{(d\mathbf{Q}_1)^{\wedge}} \;;\; \frac{(d\mathbf{Y}_1)^{\wedge}}{(d\mathbf{R}_2)^{\wedge}}\frac{(d\mathbf{R}_2)^{\wedge}}{(d\mathbf{A})^{\wedge}} = \frac{(d\mathbf{Q}_2)^{\wedge}}{2^n} \tag{4.112}$$

Therefore, the Jacobian determinant in (4.110) is

$$J(\mathbf{S} \rightarrow \mathbf{A}) = |\det\mathbf{B}|^{2n}\frac{(d\mathbf{Q}_2)^{\wedge}}{(d\mathbf{Q}_1)^{\wedge}} \tag{4.113}$$

However, there is a unitary matrix $\mathbf{Q}$ (regarded as a member of the space of cosets spanning the unitary group) such that $\mathbf{Q}_2 = \mathbf{Q}\mathbf{Q}_1$, yielding

$$\frac{(d\mathbf{Q}_2)^{\wedge}}{(d\mathbf{Q}_1)^{\wedge}} = |\det\mathbf{Q}|^{2n} = 1 \tag{4.114}$$

and the final result for the Jacobian determinant is then[1](see next page)

$$J(\mathbf{S} \rightarrow \mathbf{A}) = \frac{(d\mathbf{S})^{\wedge}}{(d\mathbf{A})^{\wedge}} = |\det\mathbf{B}|^{2n} \tag{4.115}$$

4. Jacobian of the Eigenvalue decomposition - let $\mathbf{S}$ be an hermitian positive definite $n \times n$ matrix with the eigenvalue decomposition given by

$$\mathbf{S} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \tag{4.116}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with real and distinct elements. It is important to make a brief pause here and analyse the above decomposition. Matrix $\mathbf{Q}$ has n^2 independent parameters, the same number as $\mathbf{S}$. Therefore the decomposition is not one-to-one because the right side of (4.150) has a total of $n^2 + n$ independent parameters. Let the columns of the $n \times n$ unitary matrix $\mathbf{Q}$ be given by $\mathbf{q}_i$. Clearly, introducing a random phase in each of its columns will always determine the same matrix $\mathbf{S}$

$$\begin{aligned}\mathbf{Q}_\phi\mathbf{\Lambda}\mathbf{Q}_\phi^H &= (\mathbf{q}_1 e^{j\phi_1},..,\mathbf{q}_{n_R} e^{j\phi_{n_R}})\,\mathrm{diag}(\lambda_1,...,\lambda_{n_R})(\mathbf{q}_1 e^{j\phi_1},..,\mathbf{q}_{n_R} e^{j\phi_{n_R}})^H \\ &= (\mathbf{q}_1,..,\mathbf{q}_{n_R})\,\mathrm{diag}(e^{j\phi_1}\lambda_1 e^{-j\phi_1},...,e^{j\phi_{n_R}}\lambda_{n_R} e^{-j\phi_{n_R}})(\mathbf{q}_1,..,\mathbf{q}_{n_R})^H \\ &= \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H = \mathbf{S}\end{aligned} \tag{4.117}$$

which confirms that the transformation in (4.150) is not one-to-one or, in other words, does not have a unique solution. Since a one-to-one mapping is required, at least one entry in each column must have its phase fixed, for example, by requiring that the diagonal elements of matrix $\mathbf{Q}$ be real.

Differentiating (4.116), the volume element of $\mathbf{S}$ is written as

$$(\mathbf{dS})^\wedge = (\mathbf{Q}d\mathbf{\Lambda}\mathbf{Q}^H + d\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H + \mathbf{Q}\mathbf{\Lambda}d\mathbf{Q}^H)^\wedge \tag{4.118}$$

To find a closed expression for (4.118), it is helpful to use the same reasoning already employed in the case of the QR factorization (4.96)-(4.105). Multiplying the left and right sides of the differentials by $\mathbf{Q}^H$ and $\mathbf{Q}$, respectively, (4.118) reassembles to

$$\begin{aligned}(\mathbf{Q}^H\mathbf{dS}\mathbf{Q})^\wedge &= (d\mathbf{\Lambda} + \mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda} + \mathbf{\Lambda}d\mathbf{Q}^H\mathbf{Q})^\wedge \\ &= \left(d\mathbf{\Lambda} + \mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda} + (\mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda})^H\right)^\wedge\end{aligned} \tag{4.119}$$

But using (4.115), the left member of (4.119) becomes

$$(\mathbf{Q}^H\mathbf{dS}\mathbf{Q})^\wedge = |\det\mathbf{Q}|^{2n}(\mathbf{dS})^\wedge = (\mathbf{dS})^\wedge \tag{4.120}$$

which is the required volume element in (4.118). In addition, using the fact that $\mathbf{Q}^H d\mathbf{Q}$ is skew-hermitian, it follows that only the diagonal (purely imaginary) and subdiagonal elements of $\mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda}$ are relevant. These are given by

$$[\mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda}]_{ij} = \mathbf{q}_i^H d\mathbf{q}_j\lambda_j \qquad i \geq j \tag{4.121}$$

Furthermore, since the sum $\mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda} + (\mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda})^H$ in (4.119) is hermitian, its diagonal elements are all zero, which makes this sum a perfect fit with $d\mathbf{\Lambda}$

$$\begin{aligned}(\mathbf{dS})^\wedge &= (d\mathbf{\Lambda} + \mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda} + \mathbf{\Lambda}d\mathbf{Q}^H\mathbf{Q})^\wedge \\ &= (d\mathbf{\Lambda})^\wedge \wedge \left(\mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda} + (\mathbf{Q}^H d\mathbf{Q}\mathbf{\Lambda})^H\right)^\wedge\end{aligned} \tag{4.122}$$

---

1. Compare this result to the transformation $\mathbf{S} = \mathbf{B}^H\mathbf{A}\mathbf{B}$ with regular matrices $\mathbf{A}$ and $\mathbf{S}$: in this case $J(\mathbf{S} \rightarrow \mathbf{A}) = |\det\mathbf{B}|^{4n}$.

The matrix elements of the hermitian sum are

$$\left[\mathbf{Q}^H d\mathbf{Q}\boldsymbol{\Lambda} + (\mathbf{Q}^H d\mathbf{Q}\boldsymbol{\Lambda})^H\right]_{ij} = \mathbf{q}_i^H d\mathbf{q}_j(\lambda_j - \lambda_i) \qquad i \geq j \tag{4.123}$$

and therefore the volume element for this factor is

$$\begin{aligned}\left(\mathbf{Q}^H d\mathbf{Q}\boldsymbol{\Lambda} + (\mathbf{Q}^H d\mathbf{Q}\boldsymbol{\Lambda})^H\right)^\wedge &= \bigwedge_{i\geq j}^{n}\left(\mathbf{q}_i^H d\mathbf{q}_j(\lambda_j - \lambda_i)\right)^\wedge \\ &= \prod_{i>j}^{n}(\lambda_j - \lambda_i)^2\,(d\mathbf{Q})^\wedge\end{aligned} \tag{4.124}$$

Inserting this into (4.122) gives the final expression for the volume element of matrix $\mathbf{S}$

$$\begin{aligned}(\mathbf{dS})^\wedge &= (d\boldsymbol{\Lambda})^\wedge \wedge \left(\mathbf{Q}^H d\mathbf{Q}\boldsymbol{\Lambda} + (\mathbf{Q}^H d\mathbf{Q}\boldsymbol{\Lambda})^H\right)^\wedge \\ &= \prod_{i>j}^{n}(\lambda_j - \lambda_i)^2\,(d\mathbf{Q})^\wedge \wedge (d\boldsymbol{\Lambda})^\wedge\end{aligned} \tag{4.125}$$

The Jacobian determinant follows directly

$$J(\mathbf{S} \to \mathbf{Q}, \boldsymbol{\Lambda}) = \frac{(d\mathbf{S})^\wedge}{(d\mathbf{Q})^\wedge \wedge (d\boldsymbol{\Lambda})^\wedge} = \prod_{i>j}^{n}(\lambda_j - \lambda_i)^2 \tag{4.126}$$

4.5.3 THE STIEFEL MANIFOLD AND THE UNITARY GROUP

Consider the set of all $m \times n$ $(m \geq n)$ unitary matrices $\mathbf{Q}$, that is, the set of all complex matrices with n orthonormal columns (frames) in $\mathbb{C}^m$. All such matrices satisfy the equation $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$ (i.e. n^2 independent conditions), and thus define a $(2mn - n^2)$-dimensional algebraic variety in $\mathbb{R}^{2mn}$ usually called the *complex Stiefel manifold* $\widetilde{V}_{n,m}$. The set of all $n \times n$ unitary matrices $\mathbf{Q}$ with the operation of matrix multiplication is called the *unitary group* $\widetilde{U}(n)$.

Choosing a matrix $\overline{\mathbf{Q}} = \begin{bmatrix}\mathbf{Q} & \mathbf{Q}_\perp\end{bmatrix}$ which spans $\mathbf{Q}$ and its orthogonal complement, the volume element on the Stiefel manifold may be written as

$$(d\mathbf{Q})^\wedge = \left(\overline{\mathbf{Q}}^H d\mathbf{Q}\right)^\wedge \tag{4.127}$$

which defines a differential form of maximum degree, namely $(2mn - n^2)$. Thus, it represents a measure, and hence the total volume of the Stiefel manifold can be obtained by direct integration

$$\text{Volume}\left(\widetilde{V}_{n,m}\right) = \int_{\widetilde{V}_{n,m}} (d\mathbf{Q})^\wedge \tag{4.128}$$

To evaluate this integral, set $\boldsymbol{\Sigma} = \mathbf{I}$ in (4.72) and define $\mathbf{H}^H$ as an $m \times n$ $(m \geq n)$ complex matrix with independent zero-mean Gaussian entries, then integrate it and equate the result to unity, that is

$$\int_{\widetilde{V}_\mathbf{H}} \exp[-tr(\mathbf{H}\mathbf{H}^H)](d\mathbf{H})^\wedge = \pi^{mn} \tag{4.129}$$

Now, assume $\mathbf{H}^H$ has a unique QR factorization with strictly complex $\mathbf{Q}$, and with the help

of (4.105) write

$$\begin{aligned}\pi^{mn} &= \int_{\widetilde{V}_{\mathbf{R}}} \exp[-tr(\mathbf{R}^H\mathbf{R})]\prod_{k=1}^{n} r_{kk}^{(2m-2k+1)}\,(d\mathbf{R})^{\wedge}\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}\\ &= \int_{\widetilde{V}_{\mathbf{R}}} \exp\left(-\sum_{i\le j}^{n}|r_{ij}|^2\right)\prod_{k=1}^{n} r_{kk}^{(2m-2k+1)}\,\underset{i\le j}{\overset{n}{\mathbf{\Lambda}}}\big(dr_{ij}\big)^{\wedge}\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}\end{aligned} \tag{4.130}$$

The first integral can be further simplified by separating the real and imaginary parts of r_{ij}, and grouping factors

$$\begin{aligned}\pi^{mn} &= \int_{\widetilde{V}_{\mathbf{R}}} \exp\left(-\sum_{i\le j}^{n}[r_{ij}]_I^2+[r_{ij}]_Q^2\right)\prod_{k=1}^{n} r_{kk}^{(2m-2k+1)}\,\underset{i<j}{\overset{n}{\mathbf{\Lambda}}}[dr_{ij}]_I\wedge[dr_{ij}]_Q\,\underset{i=1}{\overset{n}{\mathbf{\Lambda}}}[dr_{ii}]_I\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}\\ &= \prod_{i<j}^{n}\Big[\int_{-\infty}^{\infty}\exp\big(-[r_{ij}]_I^2\big)[dr_{ij}]_I\Big]^2\prod_{i=1}^{n}\Big[\int_{-\infty}^{\infty}\exp\big(-[r_{ii}]_I^2\big)r_{ii}^{(2m-2i+1)}[dr_{ii}]_I\Big]\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}\end{aligned} \tag{4.131}$$

The first integral in (4.131) is the ordinary Gaussian integral, and the second integral can be transformed to the Gamma function by the transformation $x=[r_{ii}]_r^2$. Therefore

$$\begin{aligned}\pi^{mn} &= \pi^{n(n-1)/2}\prod_{i=1}^{n}\Big[\frac{1}{2}\int_{-\infty}^{\infty}\exp(-x)\,x^{(m-i)}dx\Big]\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}\\ &= \frac{\pi^{n(n-1)/2}}{2^n}\prod_{i=1}^{n}\Gamma(m-i+1)\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}=\frac{1}{2^n}\widetilde{\Gamma}_n(m)\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}\end{aligned} \tag{4.132}$$

where the complex multivariate gamma function is defined as

$$\widetilde{\Gamma}_n(m)=\pi^{n(n-1)/2}\prod_{i=1}^{n}\Gamma(m-i+1) \tag{4.133}$$

It now follows that the total volume of the complex Stiefel manifold is

$$\text{Volume}\big[\widetilde{V}_{n,m}\big]=\int_{\widetilde{V}_{n,m}}(d\mathbf{Q})^{\wedge}=\prod_{i=1}^{n}\frac{2\pi^{m-i+1}}{\Gamma(m-i+1)}=\frac{2^n\pi^{mn}}{\widetilde{\Gamma}_n(m)} \tag{4.134}$$

Recall that this volume was computed assuming a QR factorization with strictly complex unitary matrix $\mathbf{Q}$ and an upper-triangular matrix $\mathbf{R}$ with real diagonal. Nevertheless, it is sometimes desirable to put an additional restriction in matrix $\mathbf{Q}$. As mentioned before, this is the case of the Eigenvalue decomposition, where it is necessary to force the imaginary parts of the diagonal elements of $\mathbf{Q}$ to zero, in order to make the decomposition unique. But this restricts the range of $\mathbf{Q}$, and thus introduces an additional constraint to the Stiefel manifold that changes the total volume. The constrained Stiefel manifold will be denoted by $\widetilde{V}_{n,m}^{\subset}$. To find this volume one must follow the same steps as above, but this time use the alternative QR factorization for which the Jacobian determinant is the one given in (4.106). Equation (4.105) is now

$$\begin{aligned}\pi^{mn} &= \int_{\widetilde{V}_{\mathbf{R}}} \exp[-tr(\mathbf{R}^H\mathbf{R})]\prod_{k=1}^{n}|r_{kk}|^{2(m-k)}\,(d\mathbf{R})^{\wedge}\int_{\widetilde{V}_{n,m}^{\subset}}(d\mathbf{Q})^{\wedge}\\ &= \int_{\widetilde{V}_{\mathbf{R}}} \exp\left(-\sum_{i\le j}^{n}|r_{ij}|^2\right)\prod_{k=1}^{n}|r_{kk}|^{2(m-k)}\,\underset{i\le j}{\overset{n}{\mathbf{\Lambda}}}\big(dr_{ij}\big)^{\wedge}\int_{\widetilde{V}_{n,m}^{\subset}}(d\mathbf{Q})^{\wedge}\end{aligned} \tag{4.135}$$

Once again, separating the real and imaginary parts of r_{ij}, and grouping factors leads to

$$\pi^{mn} = \pi^{n(n-1)/2}\prod_{i=1}^{n}\left[\int_{-\infty}^{\infty}\exp\left(-[r_{ii}]_I^2-[r_{ii}]_Q^2\right)|r_{ii}|^{2(m-i)}[dr_{ii}]_I\wedge[dr_{ii}]_Q\right]\int_{\widetilde{V}_{n,m}^{\subset}}(d\mathbf{Q})^{\wedge} \tag{4.136}$$

Introducing polar coordinates

$$\begin{gathered} r_i = |r_{ii}|\ ;\ [r_{ii}]_I = r_i\cos\theta\ ;\ [r_{ii}]_Q = r_i\sin\theta; \\ [dr_{ii}]_I\wedge[dr_{ii}]_Q = r_i\,dr_i\wedge d\theta \end{gathered} \tag{4.137}$$

(4.136) simplifies to

$$\begin{aligned} \pi^{mn} &= \pi^{n(n-1)/2}\prod_{i=1}^{n}\left[\int_{-\infty}^{\infty}\exp\left(-r_i^2\right)r_i^{(2m-2i+1)}dr_i\int_0^{2\pi}d\theta\right]\int_{\widetilde{V}_{n,m}^{\subset}}(d\mathbf{Q})^{\wedge} \\ &= \frac{(2\pi)^n\pi^{n(n-1)/2}}{2^n}\prod_{i=1}^{n}\Gamma(m-i+1)\int_{\widetilde{V}_{n,m}^{\subset}}(d\mathbf{Q})^{\wedge} = \pi^n\widetilde{\Gamma}_n(m)\int_{\widetilde{V}_{n,m}^{\subset}}(d\mathbf{Q})^{\wedge} \end{aligned} \tag{4.138}$$

Therefore, from (4.138) we find that the volume subtended by the set of all unitary $m\times n$ $(m\geq n)$ matrices $\mathbf{Q}$ with real diagonal elements is

$$\text{Volume}\left[\widetilde{V}_{n,m}^{\subset}\right] = \int_{\widetilde{V}_{n,m}^{\subset}}(d\mathbf{Q})^{\wedge} = \frac{\pi^{n(m-1)}}{\widetilde{\Gamma}_n(m)} \tag{4.139}$$

This is simply (4.134) divided by $(2\pi)^n$, which accounts for the arbitrary phases of the diagonal elements of $\mathbf{Q}$. An immediate corollary of this result is that whenever a function $h(\mathbf{Q})$ that **does not depend** on the particular phase of the columns of $\mathbf{Q}$ is integrated over $\widetilde{V}_{n,m}^{\subset}$, the following equality holds

$$\int_{\widetilde{V}_{n,m}^{\subset}}h(\mathbf{Q})(d\mathbf{Q})^{\wedge} = \frac{1}{(2\pi)^n}\int_{\widetilde{V}_{n,m}}h(\mathbf{Q})(d\mathbf{Q})^{\wedge} \tag{4.140}$$

The volumes for the unitary group and constrained unitary group are obtained by simply putting $n=m$ in the general expressions. Finally, it should be pointed out that the computed volumes are invariant under right and left transformations with unitary square matrices.

4.5.4 NONSINGULAR WISHART DISTRIBUTION

Returning to the problem at hand, and letting $\mathbf{A}=\mathbf{H}\mathbf{H}^H$, the challenge is to find the pdf of $\mathbf{A}$ from the pdf of channel matrix $\mathbf{H}$, rewritten here for convenience

$$f_{\mathbf{H}}(\mathbf{H}) = |\det(\pi\boldsymbol{\Sigma})|^{-n_T}\exp\left[-tr\left(\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{H}^H\right)\right] \tag{4.141}$$

Regarding the columns $\mathbf{h}_j$ of $\mathbf{H}$, as samples from a multivariate normal distribution, then we may write

$$\mathbf{A} = \mathbf{H}\mathbf{H}^H = \sum_{j=1}^{n_T}\mathbf{h}_j\mathbf{h}_j^H \tag{4.142}$$

which clearly shows that $\mathbf{A}$ is proportional to the *sample covariance matrix*

$$\mathbf{W} = \frac{1}{n_T - 1}\sum_{j=1}^{n_T}\mathbf{h}_j\mathbf{h}_j^H \tag{4.143}$$

The joint density of the elements of $\mathbf{W}$ in the real case was first computed by J. Wishart [42], and naturally $\mathbf{W}$ inherited his name: it is called the *Wishart matrix*. However, in order to derive the pdf of $\mathbf{A}$ we require the complex case, which is somewhat different. To that end it is useful to characterize the $n_R \times n_T$ channel matrix $\mathbf{H}$:

- since the rank of the $n_R \times n_R$ matrix $\mathbf{A}$ equals the rank of $\mathbf{H}$, and since $\mathbf{H}\mathbf{H}^H$ is positive semidefinite, matrix $\mathbf{A}$ is nonsingular (and thus positive definite) if and only if $n_T \geq n_R$;
- recalling that the columns of $\mathbf{H}$ are statistically independent, it follows that, in statistical terms, $\mathbf{H}$ contains n_R linearly independent columns; therefore, the rank of the matrix $\mathbf{H}^H$ is n_R, which means that it has a QR factorization

  $$\mathbf{H}^H = \mathbf{Q}\mathbf{R} \tag{4.144}$$

  where $\mathbf{Q}$ is an $n_T \times n_R$ unitary matrix and $\mathbf{R}$ is an $n_R \times n_R$ upper-triangular matrix with positive and real diagonal entries;

This transforms $\mathbf{A}$ into the product of lower-triangular and upper-triangular matrices: the Cholesky factorization

$$\mathbf{A} = \mathbf{R}^H\mathbf{Q}^H\mathbf{Q}\mathbf{R} = \mathbf{R}^H\mathbf{R} \tag{4.145}$$

and consequently the pdf of $\mathbf{A}$ can be expressed as

$$f_{\mathbf{A}}(\mathbf{A}) = f_{\mathbf{R}}(\mathbf{R})\frac{(d\mathbf{R})^{\wedge}}{(d\mathbf{A})^{\wedge}} = \left[\int_{\widetilde{V}_{n_R,n_T}} f_{\mathbf{QR}}(\mathbf{Q},\mathbf{R})(d\mathbf{Q})^{\wedge}\right]\frac{(d\mathbf{R})^{\wedge}}{(d\mathbf{A})^{\wedge}} \tag{4.146}$$

Moreover, using the Jacobian of the QR factorization in (4.105) it is possible to write

$$f_{\mathbf{QR}}(\mathbf{Q},\mathbf{R})(d\mathbf{Q})^{\wedge} = f_{\mathbf{H}}(\mathbf{R}^H\mathbf{Q}^H)\frac{(d\mathbf{H})^{\wedge}}{(d\mathbf{R})^{\wedge}} = f_{\mathbf{H}}(\mathbf{R}^H\mathbf{Q}^H)\prod_{k=1}^{n_R} r_{kk}^{(2n_T-2k+1)}(d\mathbf{Q})^{\wedge} \tag{4.147}$$

Hence, inserting (4.147) into (4.146) and expanding the pdf of $\mathbf{H}$, yields

$$f_{\mathbf{A}}(\mathbf{A}) = \left[|\det(\pi\boldsymbol{\Sigma})|^{-n_T}\exp\left[-tr\left(\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)\right]\prod_{k=1}^{n_R} r_{kk}^{(2n_T-2k+1)}\int_{\widetilde{V}_{n_R,n_T}}(d\mathbf{Q})^{\wedge}\right]\frac{(d\mathbf{R})^{\wedge}}{(d\mathbf{A})^{\wedge}} \tag{4.148}$$

Finally, inserting the volume of the Stiefel manifold (4.134) and the Jacobian of the Cholesky factorization (4.95) into (4.148), the pdf of matrix $\mathbf{A}$ is

$$\begin{aligned} f_{\mathbf{A}}(\mathbf{A}) &= \frac{\pi^{n_T n_R}}{|\det(\pi\boldsymbol{\Sigma})|^{n_T}\,\widetilde{\Gamma}_{n_R}(n_T)}\exp\left[-tr\left(\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)\right]\prod_{k=1}^{n_R} r_{kk}^{2(n_T-n_R)} \\ &= \frac{1}{|\det\boldsymbol{\Sigma}|^{n_T}\,\widetilde{\Gamma}_{n_R}(n_T)}\exp\left[-tr\left(\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)\right](\det\mathbf{A})^{n_T-n_R} \end{aligned} \tag{4.149}$$

This is indeed the pdf of the matrix $\mathbf{H}\mathbf{H}^H$ that was defined in the MIMO channel capacity expression (4.32), and it is known in the "realm" of multivariate statistics as the *complex Wishart distribution* $\mathbf{A} \sim \widetilde{\mathcal{W}}_{n_R}(n_T,\boldsymbol{\Sigma})$.

4.5.5 EIGENVALUE DISTRIBUTION

To take the expectation in the capacity expression (4.54), the distribution of the eigenvalues of $\mathbf{A}$ must be known. Even though the joint pdf of the eigenvalues of the Wishart matrix is already known (e.g. [43], and more recently [44]), its derivation has been done in a different context, disregarding the statistical peculiarities of the $\mathbf{h}_j$'s (e.g. the correlation between its elements), and is in fact somewhat intricate.

To round out the section, the author understood it was important to give a brief, yet lucid derivation. From (4.47) the eigenvalue decomposition is given by

$$\mathbf{A} = \mathbf{H}\mathbf{H}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \tag{4.150}$$

where $\mathbf{\Lambda} = diag(\lambda_1, \lambda_2, ..., \lambda_{n_R})$. To make the factorization unique, we fix $\mathbf{\Lambda}$ such that the eigenvalues are ordered from greatest to least $\lambda_1 > \lambda_2 > ... > \lambda_{n_R} > 0$, and write explicitly

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}^o\mathbf{Q}^H \tag{4.151}$$

We also impose no restriction to $\mathbf{Q}$ apart from a diagonal with real elements. The joint pdf of $\mathbf{\Lambda}^o$ and $\mathbf{Q}$ satisfies

$$f_{\mathbf{\Lambda Q}}(\mathbf{\Lambda}^o, \mathbf{Q})(d\mathbf{Q})^\wedge = f_{\mathbf{A}}(\mathbf{Q}\mathbf{\Lambda}^o\mathbf{Q}^H)\frac{(d\mathbf{A})^\wedge}{(d\mathbf{\Lambda}^o)^\wedge} \tag{4.152}$$

Expanding the Wishart pdf with (4.149) and inserting the Jacobian of the Eigenvalue transformation in (4.126), we get

$$f_{\mathbf{\Lambda Q}}(\mathbf{\Lambda}^o, \mathbf{Q})(d\mathbf{Q})^\wedge = \frac{\prod_{i<j}^{n_R}(\lambda_j - \lambda_i)^2 (\det\mathbf{\Lambda})^{n_T - n_R}}{|\det\mathbf{\Sigma}|^{n_T}\,\widetilde{\Gamma}_{n_R}(n_T)} \exp\left[-tr\left(\mathbf{\Sigma}^{-1}\mathbf{Q}\mathbf{\Lambda}^o\mathbf{Q}^H\right)\right](d\mathbf{Q})^\wedge \tag{4.153}$$

Once again, an integration is required to eliminate the $\mathbf{Q}$ dependence from (4.153), that is

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}^o) = \frac{\prod_{i=1}^{n_R}\lambda_i^{(n_T - n_R)}\prod_{i>j}^{n_R}(\lambda_j - \lambda_i)^2}{|\det\mathbf{\Sigma}|^{n_T}\,\widetilde{\Gamma}_{n_R}(n_T)} \int_{\widetilde{V}^{\subset}_{n_R,n_R}} \exp\left[-tr\left(\mathbf{\Sigma}^{-1}\mathbf{Q}\mathbf{\Lambda}^o\mathbf{Q}^H\right)\right](d\mathbf{Q})^\wedge \tag{4.154}$$

where, since $\mathbf{Q}$ is a unitary matrix with real diagonal elements, the integration is over the constrained Stiefel manifold $\widetilde{V}^{\subset}_{n_R,n_R}$ or, in light of (4.140), over the unitary group $\widetilde{U}(n_R)$.

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}^o) = \frac{\prod_{i=1}^{n_R}\lambda_i^{(n_T - n_R)}\prod_{i>j}^{n_R}(\lambda_j - \lambda_i)^2}{(2\pi)^{n_R}\,|\det\mathbf{\Sigma}|^{n_T}\,\widetilde{\Gamma}_{n_R}(n_T)} \int_{\widetilde{U}(n_R)} \exp\left[-tr\left(\mathbf{\Sigma}^{-1}\mathbf{Q}\mathbf{\Lambda}^o\mathbf{Q}^H\right)\right](d\mathbf{Q})^\wedge \tag{4.155}$$

Equation (4.155) gives the joint density function of the strictly positive and **ordered** eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_{n_R} > 0$ of $\mathbf{A} = \mathbf{H}\mathbf{H}^H$.

To find the unordered distribution of the eigenvalues one may notice that the ordered eigenvalues define a $(1/n_R!)$-th part of positive n_R-dimensional space. Thus, imposing no order to the eigenvalues is equivalent to extending the ordered $f_{\mathbf{\Lambda}}(\mathbf{\Lambda}^o)$ to the whole extent of positive n_R-dimensional space, which can be done by finding the pdfs of all permuta-

tions of the eigenvalues. Hence, we must have that

$$K\int_{\lambda_1,..,\lambda_{n_R}}\left[\sum_{p=(\lambda_1,..,\lambda_{n_R})} f_p(p)\right]d\lambda_1..d\lambda_{n_R} = 1 \tag{4.156}$$

where the $f_p(p)$'s are the non-overlapping pdfs of the ordered eigenvalues, and the summation is over all permutations of the eigenvalues. It then follows that $K = 1/n_R!$, and by consequence the pdf of the **unordered** eigenvalues is given by

$$\begin{aligned} f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) &= \frac{1}{n_R!}\sum_{p=(\lambda_1,..,\lambda_{n_R})} f_p(p) = \frac{1}{n_R!} f_{\mathbf{\Lambda}}(\mathbf{\Lambda}^o) \\ &= \frac{\prod_{i=1}^{n_R}\lambda_i^{(n_T-n_R)}\prod_{i<j}^{n_R}(\lambda_j-\lambda_i)^2}{n_R!(2\pi)^{n_R}|\det\mathbf{\Sigma}|^{n_T}\tilde{\Gamma}_{n_R}(n_T)}\int_{\tilde{U}(n_R)}\exp\left[-tr\left(\mathbf{\Sigma}^{-1}\mathbf{Q}\mathbf{\Lambda}^o\mathbf{Q}^H\right)\right](d\mathbf{Q})^\wedge \end{aligned} \tag{4.157}$$

which is simply an average of (4.155). This expression is not yet written in closed-form because the integral must still be evaluated. This integral depends on the covariance matrix $\mathbf{\Sigma} = E[\mathbf{h}_j\mathbf{h}_j^H]$ of the channel matrix columns or, in other words, it depends on the correlation between the fading processes at different receiver antennas.

4.5.6 INDEPENDENT FADING

It turns out that a special case of the matrix $\mathbf{\Sigma}$ makes the integration in (4.157) straightforward. When the fading processes at different receiver antennas are uncorrelated with one another (i.e. independent fading) and therefore $\mathbf{\Sigma} = E\left[|h_{ij}|^2\right]\mathbf{I}_{n_R} = \mathbf{I}_{n_R}$, substitution into the integral of (4.157) and using (4.134) leads to

$$\begin{aligned} \int_{\tilde{U}(n_R)}\exp[-tr(\mathbf{Q}\mathbf{\Lambda}^o\mathbf{Q}^H)](d\mathbf{Q})^\wedge &= \exp[-tr(\mathbf{\Lambda}^o)]\int_{\tilde{U}(n_R)}(d\mathbf{Q})^\wedge \\ &= \frac{2^{n_R}\pi^{n_Rn_R}}{\tilde{\Gamma}_{n_R}(n_R)}\exp\left(-\sum_{i=1}^{n_R}\lambda_i\right) \end{aligned} \tag{4.158}$$

which inserted into (4.157) leads to the joint pdf of the unordered eigenvalues in this particular fading scenario

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda})_{\mathbf{\Sigma}=\mathbf{I}} = \frac{\pi^{n_R(n_R-1)}}{n_R!}\frac{\prod_{i=1}^{n_R}\lambda_i^{n_T-n_R}\prod_{i<j}^{n_R}(\lambda_j-\lambda_i)^2}{\tilde{\Gamma}_{n_R}(n_T)\tilde{\Gamma}_{n_R}(n_R)}\exp\left(-\sum_{i=1}^{n_R}\lambda_i\right) \tag{4.159}$$

Recall that this distribution was obtained by always assuming that the number of receiver antennas was smaller than or equal to the number of transmitter antennas, i.e. $n_T \geq n_R$, otherwise the Wishart matrix $\mathbf{A} = \mathbf{H}\mathbf{H}^H$ would be singular and consequently would render the problem much difficult. However, in the case of independent fading at the receiver side this restriction can be relaxed. This is because the pdf of the channel matrix as given by (4.141) will now have the form

$$f_{\mathbf{H}}(\mathbf{H}) = \pi^{-n_Rn_T}\exp[-tr(\mathbf{H}\mathbf{H}^H)] = \pi^{-n_Rn_T}\exp[-tr(\mathbf{H}^H\mathbf{H})] \tag{4.160}$$

and as a result the Wishart matrix for $n_T < n_R$ may be written as $\mathbf{A} = \mathbf{H}^H\mathbf{H}$, which is clearly

nonsingular. To cover both cases one may define the variables

$$\begin{aligned} m &= \min(n_R, n_T) \\ n &= \max(n_R, n_T) \end{aligned} \tag{4.161}$$

and write the joint density of the unordered eigenvalues as

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda})_{\Sigma=\mathbf{I}} = \frac{\pi^{m(m-1)}}{m!} \frac{\prod_{i=1}^{m} \lambda_i^{n-m} \prod_{i<j}^{m} (\lambda_j - \lambda_i)^2}{\tilde{\Gamma}_m(n)\tilde{\Gamma}_m(m)} \exp\left(-\sum_{i=1}^{m} \lambda_i\right) \tag{4.162}$$

Now, noting that (4.162) is a symmetric function of the eigenvalues, the average MIMO channel capacity can be computed by means of (4.55), i.e.

$$\begin{aligned} C_{avg} &= \min(n_R, n_T) E\left[\log_2\left(1 + \frac{\gamma}{n_T}\lambda\right)\right] \\ &= m \int_0^{\infty} \log_2\left(1 + \frac{\gamma}{n_T}\lambda\right) f_{\lambda}(\lambda) d\lambda \end{aligned} \tag{4.163}$$

The only thing left to compute is the pdf of some eigenvalue, e.g. $\lambda = \lambda_1$, something that may be accomplished by integrating (4.162), for instance, with respect to $\lambda_2, ..., \lambda_m$. Thus

$$f_{\lambda}(\lambda)_{\Sigma=\mathbf{I}} = \frac{K_1}{m!} \int_{\lambda_2,...,\lambda_m} \prod_{i<j}^{m} (\lambda_j - \lambda_i)^2 \prod_{i=1}^{m} \lambda_i^{n-m} e^{-\lambda_i} \, d\lambda_2 ... d\lambda_m \tag{4.164}$$

A possible method to perform the integration in (4.164) is as follows. First, notice that the first factor can be put in determinant form using the following sequence of steps

$$\begin{aligned} \prod_{i<j}^{m} (\lambda_j - \lambda_i) &= [(\lambda_m - \lambda_{m-1})][(\lambda_m - \lambda_{m-2})(\lambda_{m-1} - \lambda_{m-2})]...[(\lambda_m - \lambda_1)...(\lambda_2 - \lambda_1)] \\ &= \begin{vmatrix} 1 & 1 \\ \lambda_{m-1} & \lambda_m \end{vmatrix} [(\lambda_m - \lambda_{m-2})(\lambda_{m-1} - \lambda_{m-2})]...[(\lambda_m - \lambda_1)...(\lambda_2 - \lambda_1)] \\ &= \begin{vmatrix} 1 & 1 & 1 \\ \lambda_{m-2} & \lambda_{m-1} & \lambda_m \\ \lambda_{m-2}^2 & \lambda_{m-1}^2 & \lambda_m^2 \end{vmatrix} ...[(\lambda_m - \lambda_1)...(\lambda_2 - \lambda_1)] = \begin{vmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_m \\ \vdots & & \vdots \\ \lambda_1^{m-1} & \cdots & \lambda_m^{m-1} \end{vmatrix} \end{aligned} \tag{4.165}$$

which equals the well known *Vandermonde determinant*. Then, carry out a sequence of row operations that may be useful to simplify the integration

$$\begin{aligned} |V| = \prod_{i<j}^{m} (\lambda_j - \lambda_i) &= \begin{vmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_m \\ \vdots & & \vdots \\ \lambda_1^{m-1} & \cdots & \lambda_m^{m-1} \end{vmatrix} = K_2 \begin{vmatrix} \varphi_1(\lambda_1) & \cdots & \varphi_1(\lambda_m) \\ \varphi_2(\lambda_1) & \cdots & \varphi_2(\lambda_m) \\ \vdots & & \vdots \\ \varphi_m(\lambda_1) & \cdots & \varphi_m(\lambda_m) \end{vmatrix} \\ &= K_2 \left[\sum_p \sigma(p) \prod_{i=1}^{m} \varphi_i(\lambda_{p_i}) \right] \end{aligned} \tag{4.166}$$

The eigenvalue density function then becomes

$$\begin{aligned} f_\lambda(\lambda)_{\Sigma=\mathbf{I}} &= \frac{K_1 K_2^2}{m!} \int_{\lambda_2,\ldots,\lambda_m} \left[\sum_{p,s} \sigma(p)\sigma(s) \prod_{i=1}^{m} \varphi_{p_i}(\lambda_i)\varphi_{s_i}(\lambda_i) \right] \prod_{i=1}^{m} \lambda_i^{n-m} e^{-\lambda_i} \, d\lambda_2 \ldots d\lambda_m \\ &= \frac{K_1 K_2^2}{m!} \left[\sum_{p,s} \sigma(p)\sigma(s)\varphi_{p_1}(\lambda_1)\varphi_{s_1}(\lambda_1)\lambda_1^{n-m}e^{-\lambda_1} \prod_{i=2}^{m} \left[\int_{\lambda_i} \varphi_{p_i}(\lambda_i)\varphi_{s_i}(\lambda_i)\lambda_i^{n-m}e^{-\lambda_i} \, d\lambda_i \right] \right] \end{aligned} \tag{4.167}$$

The integral inside (4.162) is recognized as the inner product of polynomial functions $\varphi_{p_i}(\lambda_i)$ with a weighting function of $\lambda_i^{n-m}e^{-\lambda_i}$. Thus, if the functions $\varphi_{p_i}(\lambda_i)$ are orthonormal, we get the simple result

$$\left\langle \varphi_{p_i}, \varphi_{s_i} \right\rangle = \int_{\lambda_i} \varphi_{p_i}(\lambda_i)\varphi_{s_i}(\lambda_i)\lambda_i^{n-m}e^{-\lambda_i} \, d\lambda_i = \delta_{p_i s_i} \tag{4.168}$$

where $\delta_{p_i s_i}$ is the Kronecker delta. The eigenvalue pdf now reduces to

$$f_\lambda(\lambda)_{\Sigma=\mathbf{I}} = \frac{K_1 K_2^2 (m-1)!}{m!} \sum_{k=1}^{m} \varphi_k^2(\lambda)\lambda^{n-m}e^{-\lambda} \tag{4.169}$$

where the constant may be found by interpreting (4.169) as a density function

$$\int_\lambda f_\lambda(\lambda)_{\Sigma=\mathbf{I}} d\lambda = \frac{K_1 K_2^2 (m-1)!}{m!} \sum_{k=1}^{m} \int_\lambda \varphi_k^2(\lambda)\lambda^{n-m}e^{-\lambda} d\lambda = K_1 K_2^2 = 1 \tag{4.170}$$

Therefore, the pdf of a single eigenvalue of the Wishart matrix in the case of independent fading is expressed as

$$f_\lambda(\lambda)_{\Sigma=\mathbf{I}} = \frac{1}{m} \sum_{k=1}^{m} \varphi_k^2(\lambda)\lambda^{n-m}e^{-\lambda} \tag{4.171}$$

for a set $\{\varphi_k(\lambda)\}$ of orthonormal polynomial functions. Recall from (4.162) that these functions were obtained by performing row operations on the Vandermonde determinant, which means that they are linear combinations of $1, \lambda, \lambda^2, \ldots, \lambda^{m-1}$. The latter are clearly independent functions because the Wronskian matrix is nonsingular

$$\det \begin{pmatrix} 1 & \lambda & \lambda^2 & \cdots & \lambda^{m-1} \\ 0 & 1 & 2\lambda & \cdots & (m-1)\lambda^{m-2} \\ 0 & 0 & 2 & \cdots & (m-1)(m-2)\lambda^{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (m-1)! \end{pmatrix} \neq 0 \tag{4.172}$$

and thus they may be orthonormalized as required. The orthonormalization can be accomplished using the well established Gram-Schmidt procedure. Writing

$$f_1(\lambda) = 1 \; ; \; f_2(\lambda) = \lambda \; ; \; \ldots \; ; \; f_m(\lambda) = \lambda^{m-1} \tag{4.173}$$

the orthonormalized functions are given by

$$\varphi_1(\lambda) = \frac{f_1(\lambda)}{\|f_1(\lambda)\|} \; ; \; \varphi_2(\lambda) = \frac{f_2(\lambda) - \langle f_2, \varphi_1 \rangle \varphi_1}{\|f_2(\lambda) - \langle f_2, \varphi_1 \rangle \varphi_1\|} \; ; \; \ldots \; ; \; \varphi_m(\lambda) = \frac{f_k(\lambda) - \sum_{i=1}^{m-1} \langle f_m, \varphi_i \rangle \varphi_i}{\left\| f_m(\lambda) - \sum_{i=1}^{m-1} \langle f_m, \varphi_i \rangle \varphi_i \right\|} \tag{4.174}$$

Using the notation $t(i) = (n-m+i)!$ and observing that $\langle \lambda^i, \lambda^j \rangle = t(i+j)$, these functions can be found recursively as follows:

- the first orthonormal function is computed as

$$\begin{aligned} & f_1(\lambda) = 1 \\ & \|f_1(\lambda)\| = \sqrt{\langle f_1, f_1 \rangle} = \left(\int_0^\infty f_1^2(\lambda) \lambda^{n-m} e^{-\lambda} d\lambda \right)^{1/2} = \sqrt{t(0)!} \\ & \varphi_1(\lambda) = \frac{f_1(\lambda)}{\|f_1(\lambda)\|} = \left(\frac{1}{t(0)!} \right)^{1/2} \end{aligned} \tag{4.175}$$

- the second orthonormal function is computed as

$$\begin{aligned} & f_2(\lambda) = \lambda \; ; \; \langle f_2, \varphi_1 \rangle = \int_0^\infty \frac{\lambda}{\sqrt{t(0)}} \lambda^{n-m} e^{-\lambda} d\lambda = \frac{t(1)}{\sqrt{t(0)}} \\ & \| f_2(\lambda) - \langle f_2, \varphi_1 \rangle \varphi_1 \| = \left[\int_0^\infty \left(\lambda - \frac{t(1)}{t(0)} \right)^2 \lambda^{n-m} e^{-\lambda} d\lambda \right]^{1/2} = \sqrt{t(1)} \\ & \varphi_2(\lambda) = \left(\frac{1}{t(1)!} \right)^{1/2} \left(\lambda - \frac{t(1)!}{t(0)!} \right) \end{aligned} \tag{4.176}$$

- the third orthonormal function is computed as

$$\begin{aligned} & f_3(\lambda) = \lambda^2 \; ; \; \langle f_3, \varphi_1 \rangle = \int_0^\infty \frac{\lambda^2}{\sqrt{t(0)}} \lambda^{n-m} e^{-\lambda} d\lambda = \frac{t(2)}{\sqrt{t(0)}} \\ & \langle f_3, \varphi_2 \rangle = \left(\frac{1}{t(1)} \right)^{1/2} \int_0^\infty \left(\lambda^3 - \frac{t(1)}{t(0)} \lambda^2 \right) \lambda^{n-m} e^{-\lambda} d\lambda = \frac{2t(2)}{\sqrt{t(1)}} \\ & \left\| f_3(\lambda) - \sum_{i=1}^{2} \langle f_3, \varphi_i \rangle \varphi_i \right\| = \left[\int_0^\infty \left(\lambda^2 - 2\frac{t(2)}{t(1)} \lambda + \frac{t(2)}{t(0)} \right)^2 \lambda^{n-m} e^{-\lambda} d\lambda \right]^{1/2} = \sqrt{2t(2)} \\ & \varphi_3(\lambda) = \left(\frac{2}{t(2)} \right)^{1/2} \left(\frac{\lambda^2}{2} - \frac{t(2)}{t(1)} \lambda + \frac{t(2)}{2t(0)} \right) \end{aligned} \tag{4.177}$$

- the same procedure leads to the fourth function

$$\varphi_4(\lambda) = \left(\frac{6}{t(3)} \right)^{1/2} \left(\frac{\lambda^3}{6} - \frac{t(3)}{2t(2)} \lambda^2 + \frac{t(3)}{2t(1)} \lambda - \frac{t(3)}{6t(0)} \right) \tag{4.178}$$

- and henceforth, by induction

$$\varphi_{k+1}(\lambda) = \left(\frac{k!}{t(k)} \right)^{1/2} \sum_{i=0}^{k} (-1)^{k-i} \frac{t(k)}{t(i)} \frac{\lambda^i}{i!(k-i)!} \qquad k = 0, \ldots, m-1 \tag{4.179}$$

which concludes the procedure. Having found the orthonormal polynomials of the pdf expression in (4.171), the next step is to insert the latter into (4.163), yielding

$$C_{avg} = \sum_{k=0}^{m-1} \int_0^\infty \varphi_{k+1}^2(\lambda) \lambda^{n-m} e^{-\lambda} \log_2 \left(1 + \frac{\gamma}{n_T} \lambda \right) d\lambda \tag{4.180}$$

Now, substituting (4.179) into (4.180) one obtains

$$C_{avg} = \sum_{k=0}^{m-1} k! t(k) \sum_{i,j=0}^{k} \frac{(-1)^{i+j}}{t(i)t(j)} \frac{1/[j!(k-j)!]}{i!(k-i)!} \int_0^\infty \lambda^{n-m+i+j} e^{-\lambda} \log_2 \left(1 + \frac{\gamma}{n_T} \lambda \right) d\lambda \tag{4.181}$$

To complete the derivation, the integral

$$\int_0^\infty \lambda^r e^{-\lambda} \log_2\left(1+\frac{\gamma}{n_T}\lambda\right)d\lambda = \log_2 e \int_0^\infty \lambda^r e^{-\lambda} \ln\left(1+\frac{\gamma}{n_T}\lambda\right)d\lambda \tag{4.182}$$

must be evaluated. Integrating by parts

$$\begin{aligned}&\int_0^\infty \lambda^r e^{-\lambda} \ln\left(1+\frac{\gamma}{n_T}\lambda\right)d\lambda = \\ &\left[\ln\left(1+\frac{\gamma}{n_T}\lambda\right)\int \lambda^r e^{-\lambda}d\lambda\right]_0^\infty - \int_0^\infty \frac{\gamma/n_T}{1+(\gamma/n_T)\lambda}\left[\int \lambda^r e^{-\lambda}d\lambda\right]d\lambda\end{aligned} \tag{4.183}$$

and using the equality in (3.45), leads to

$$\int_0^\infty \lambda^r e^{-\lambda} \ln\left(1+\frac{\gamma}{n_T}\lambda\right)d\lambda = \sum_{p=0}^{r}\frac{r!}{p!}\int_0^\infty \frac{\gamma/n_T}{1+(\gamma/n_T)\lambda}\lambda^p e^{-\lambda}d\lambda \tag{4.184}$$

Performing the change of variables $\lambda = n_T(\eta-1)/\gamma$

$$\int_0^\infty \lambda^r e^{-\lambda} \ln\left(1+\frac{\gamma}{n_T}\lambda\right)d\lambda = \sum_{p=0}^{r}\frac{r!}{p!}\frac{e^{n_T/\gamma}}{(\gamma/n_T)^p}\int_1^\infty \frac{1}{\eta}(\eta-1)^p e^{-\eta n_T/\gamma}d\eta \tag{4.185}$$

and using the binomial theorem

$$(\eta-1)^p = \sum_{l=0}^{p}\binom{p}{l}\eta^l(-1)^{p-l} \tag{4.186}$$

it follows that

$$\begin{aligned}\int_0^\infty \lambda^r e^{-\lambda} \ln\left(1+\frac{\gamma}{n_T}\lambda\right)d\lambda &= \sum_{p=0}^{r}\frac{r!}{p!}\frac{e^{n_T/\gamma}}{(\gamma/n_T)^p}\sum_{l=0}^{p}\binom{p}{l}(-1)^{p-l}\int_1^\infty \eta^{l-1}e^{-\eta n_T/\gamma}d\eta \\ &= \sum_{p=0}^{r}\frac{r!}{p!}e^{n_T/\gamma}\sum_{l=0}^{p}\binom{p}{l}\left(-\frac{n_T}{\gamma}\right)^{p-l}\Gamma(l, n_T/\gamma)\end{aligned} \tag{4.187}$$

where in the above expression

$$\Gamma(l, n_T/\gamma) = \int_{n_T/\gamma}^\infty \eta^{l-1}e^{-\eta}d\eta \tag{4.188}$$

is the incomplete gamma function. The final expression for the average capacity of the MIMO channel in independent Rayleigh fading conditions is thus given by

$$C_{avg} = \frac{e^{n_T/\gamma}}{\ln 2}\sum_{k=0}^{m-1}k!\,t(k)\sum_{i,j=0}^{k}(-1)^{i+j}\frac{t(i+j)}{t(i)t(j)}\frac{1/[\,j!(k-j)!\,]}{i!(k-i)!}\sum_{p=0}^{n-m+i+j}\frac{1}{p!}\sum_{l=0}^{p}\binom{p}{l}\left(-\frac{n_T}{\gamma}\right)^{p-l}\Gamma(l, n_T/\gamma) \tag{4.189}$$

where

$$\begin{aligned}\gamma &\rightarrow \text{average SNR at each receiver branch} \\ m &= \min(n_R, n_T) \\ n &= \max(n_R, n_T) \\ t(i) &= (n-m+i)!\end{aligned}$$

As a matter of curiosity, consider the case of a MISO system, i.e. let the transmitter have several antennas and the receiver have a single antenna. Then $m=1$, $n=n_T$ and the av-

erage channel capacity reduces to

$$C_{avg} = \frac{e^{n_T/\gamma}}{\ln 2} \sum_{p=0}^{n_T-1} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{n_T}{\gamma}\right)^{p-l} \Gamma(l, n_T/\gamma) \tag{4.190}$$

Equation (4.190) is exactly the capacity expression in (3.65) for the maximal-ratio combiner in a SIMO system, and here the effective diversity order Λ is replaced by n_T. This means, for instance, that in Rayleigh fading conditions a MISO system with n_T transmitter antennas has the same average information-theoretic capacity than the SIMO system with n_T receiver antennas and an MRC combiner at the output. This is clearly an extension of the result by Alamouti [54], which showed that in flat-fading channels and under a specific combining technique, a two-branch transmit diversity system is equivalent to a two-branch receive diversity MRC system, and hence provides the same diversity order. Moreover, in a SIMO system, i.e. $m = 1$ and $n = n_R$, the average channel capacity is

$$C_{avg} = \frac{e^{1/\gamma}}{\ln 2} \sum_{p=0}^{n_T-1} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{1}{\gamma}\right)^{p-l} \Gamma(l, 1/\gamma) \tag{4.191}$$

Still, since $\gamma = \Gamma_s / n_R$, where $n_R = \Lambda$ and Γ_s is the average SNR at the MRC output, (4.191) is also exactly equal to (3.65), showing that the maximal-ratio combiner is optimal from an information-theoretic perspective.

Now, returning to the general capacity expression in (4.189), an effective way to fully understand its behaviour is to produce several plots in diverse array dimensions at both the transmitter and receiver sides, each one as a function of the average SNR γ at each receiver branch. In this manner, the scaling of the capacity with the key variables at stake in a MIMO system will become more evident, and makes it possible to include relevant comparison curves to help in the analysis. More specifically, the plots in Figures 4.2 and 4.3 show how the capacity scales by fixing the number of receiver antennas and varying the number of transmitter antennas and the average SNR. They also incorporate the

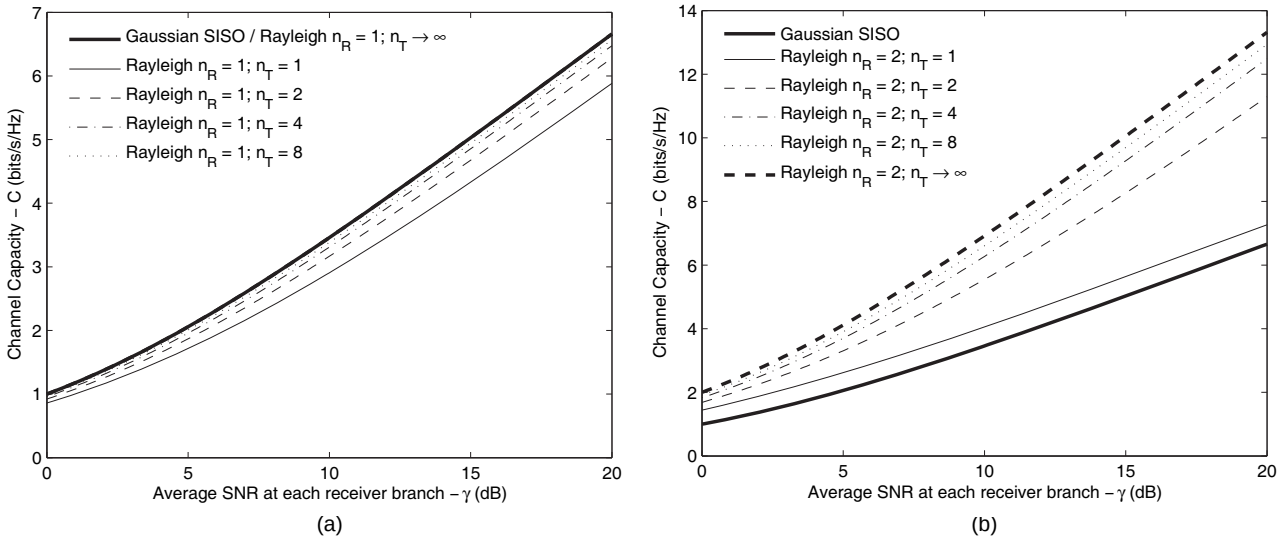

FIGURE 4.2 Average information-theoretic capacity of the MIMO channel as a function of the branch SNR at the receiver side, in the special cases of (a) one receiver branch, and (b) two receiver branches.

Gaussian SISO (Shannon) curve (Equation (4.36) with $|H|^2 = 1$), and the limiting capacity

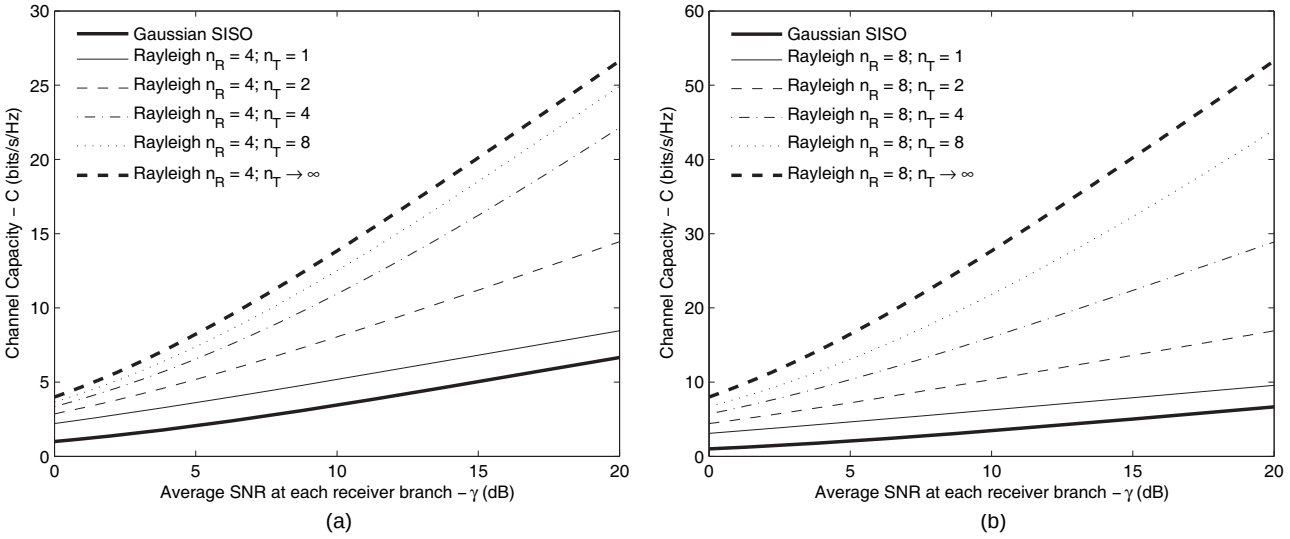

FIGURE 4.3 Average information-theoretic capacity of the MIMO channel as a function of the branch SNR at the receiver side, in the special cases of (a) four receiver branches, and (b) eight receiver branches.

for large n_T (Equation (4.40)). From these plots one may infer that:

- when the receiver side has a single antenna, the achievable capacity with a MISO system is always lower than the Gaussian SISO capacity, but is higher than the Rayleigh SISO when $n_T > 1$;
- in a Rayleigh fading environment, the average MISO capacity is equal to the capacity of the SIMO system with a maximal-ratio combiner, i.e. Figure 4.2(a) is equivalent to Figure 3.5;
- for a fixed n_R, the capacity is upper-bounded by the limit $C_{\lim}$ (4.40) as $n_T \to \infty$;
- the higher is the number of receiver antennas, the better the capacity scales with increasing n_T relative to $n_T = 1$; for example, at a branch SNR of 20 dB and with $n_R = 1$, the capacity can scale up to approximately 0.8 bits/cycle, for $n_R = 2$ it can scale up to 6 bits/cycle, for $n_R = 4$ it can scale up to 18 bits/cycle, and for $n_R = 8$ it can scale a maximum of 44 bits/cycle; however
- the higher is n_R, the higher must be n_T to get within some percentage of the limiting capacity; this fact is shown in Table 4.1, which gives the minimum n_T to reach 90% of the limiting capacity as $n_T \to \infty$;

TABLE 4.1 The number of transmitter antennas required to attain $C_{avg} \geq 0.9C_{\lim}$ (4.40).

SNR (dB) \ n_R	**1**	**2**	**4**	**8**
4	2	4	9	17
8	2	4	9	17
12	2	4	8	15
16	2	4	7	13
20	2	3	6	12

A different situation is depicted in the plots of Figures 4.4 and 4.5, which show how the capacity scales by fixing the number of transmitter antennas and varying the number of receiver antennas. From these cases one may conclude that:

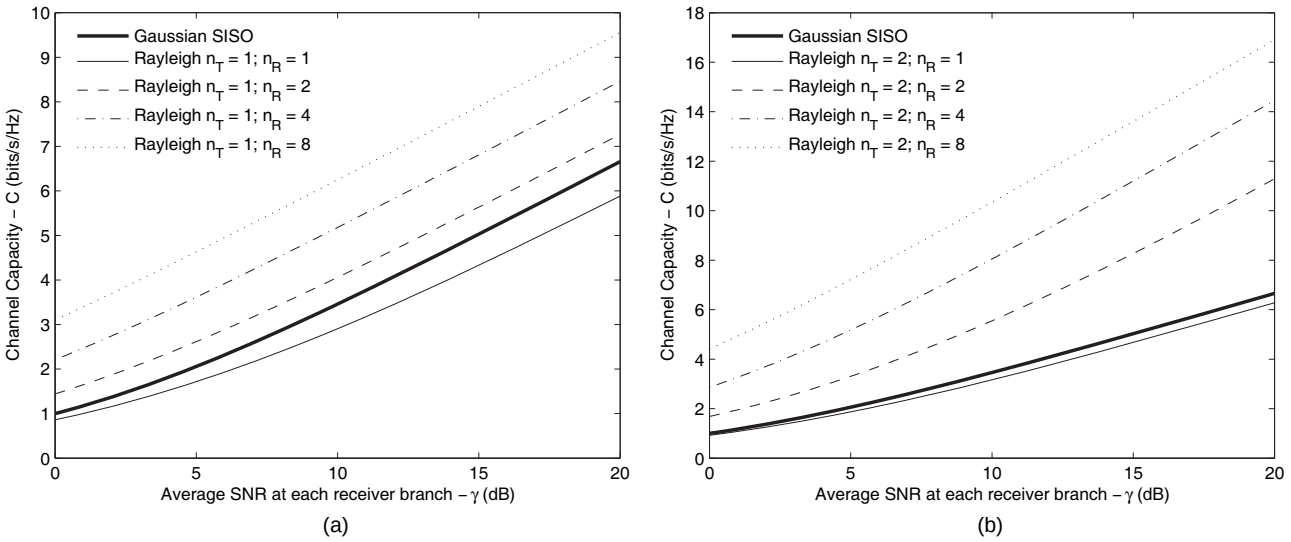

FIGURE 4.4 Average information-theoretic capacity of the MIMO channel as a function of the branch SNR at the receiver side, in the special cases of (a) one transmitter branch, and (b) two transmitter branches.

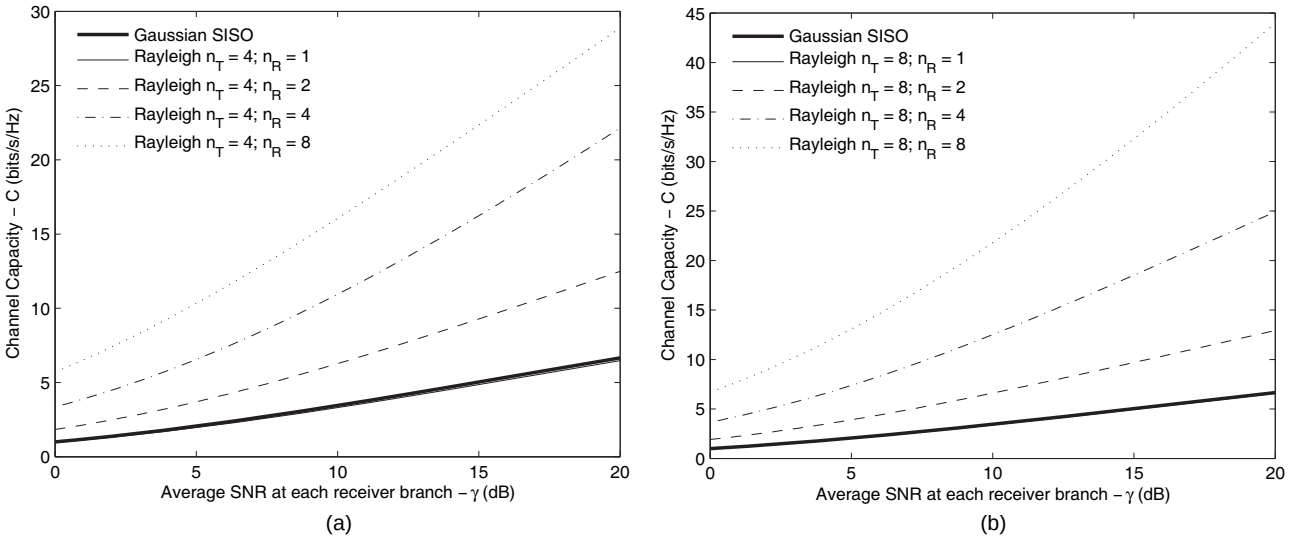

FIGURE 4.5 Average information-theoretic capacity of the MIMO channel as a function of the branch SNR at the receiver side, in the special cases of (a) four transmitter branches, and (b) eight transmitter branches.

- the capacity of the SIMO channel (Figure 4.4(a)) is higher than the Gaussian SISO capacity, provided that $n_R > 1$, contrary to the MISO channel (Figure 4.2(a)) where it is always lower; moreover
- it may seem that Figure 4.4(a) asserts that the maximal-ratio combining technique is deficient; in fact, the capacity curves of the SIMO system with MRC (Figure 3.5) are always lower that the curve of the Gaussian SISO capacity, whereas in Figure 4.4(a) they are higher when $n_R > 1$; this is explained by the different abscissas in the plots, Γ_s in Figure 3.5 and $\gamma = \Gamma_s / n_R$ in Figure 4.4(a); as explained earlier, the MRC is optimal;
- the higher is the number of receiver antennas, the better the capacity scales with increasing average SNR;
- at least in theory, the capacity is not upper-bounded with increasing n_R; among other things, this means that it is possible to attain the limiting capacity C_{lim} (4.40) as

$n_T \to \infty$, and hence that in certain situations it may be more convenient to increase n_R instead of n_T to achieve a greater capacity; as shown in Table 4.2, increasing the number of receiver antennas may convey theoretical capacities that cannot be obtained by increasing n_T solely;

TABLE 4.2 MIMO capacity comparison for a few particular array sizes, at $\gamma = 20$ dB.

Array size	C_{lim} (bits/s/Hz)	**Array size**	C_{avg} (bits/s/Hz)
$n_R = 1$; $n_T \to \infty$	6.7	$n_R = 2$; $n_T = 1$	7.3
$n_R = 2$; $n_T \to \infty$	13.3	$n_R = 4$; $n_T = 2$	14.5
$n_R = 4$; $n_T \to \infty$	26.6	$n_R = 7$; $n_T = 4$	27.8
$n_R = 8$; $n_T \to \infty$	53.3	$n_R = 13$; $n_T = 8$	54.3

Practical example. Consider the particular case of the downlink channel in the WCDMA-FDD air interface. The multiuser separation is achieved by spreading the user data with orthogonal channelization codes of length 4-512 chips, and a chip rate of $R_c = 3.84$ Mchips/s. Each chip is shaped by a *root-raised cosine* filter with roll-off factor $\alpha = 0.22$, and hence the total occupied bandwidth is $B = R_c(1+\alpha) \approx 5$ MHz. QPSK modulation is used for the downlink. With a single code, the maximum possible bit rate (including control symbols) is $2.R_c\,/\,4 = 1920$ kb/s, corresponding to a spectral efficiency per user of $\rho = 1.92\,/\,5 = 0.38$ bits/s/Hz. At the receiver, the multiuser signal is despread using the original codes, and the usual requirement for the *signal to noise-plus-interference* ratio after despreading is in the interval $0 \leq E_b\,/\,\mathcal{I}_0 \leq 5$ (dB), which is satisfactory for achieving a given error probability. The reason why the symbol rate may not be increased while maintaining the same bandwidth, is because the ratio $E_b\,/\,\mathcal{I}_0$ will also decrease if the transmitter power is not increased (e.g. duplicating the symbol rate reduces $E_b\,/\,\mathcal{I}_0$ by 3 dB), and this will have the effect of increasing the average error probability. Increasing the transmission power is not a option in a multiuser system because the interference power must be kept to a minimum. Even though it seems we have reached a deadlock, theoretical results show, however, that higher spectral efficiencies are certainly possible. For instance, at an average SNR of 5 dB, the capacities[1] that can be obtained with asymptotic MIMO arrangements ($n_R = n_T$) is summarized in Table 4.3. With a simple SISO system in a Rayleigh environment it is possible to obtain, on average, a spectral efficiency of 1.716 bits/s/Hz without receiver errors, which is more than the WCDMA spectral efficiency per user. Increasing the number of antennas dramatically increases the achievable ca-

1. **Computational note**. Numerical evaluation of the capacity expression in (4.189) may lead to incorrect results when the number of transmitter and/or receiver antennas is large. This happens because when a substantial number of additions and products is performed using finite-precision arithmetic, the final result may be inaccurate. Usually, software oriented to efficient numerical computation (such as Matlab), represents numbers using double-precision (64 bits) floating point binary values, meaning that large factorials and numerical integrals (for instance) will have limited precision. It is clearly the case of Equation (4.189). To achieve a good capacity precision, when the sum of n_R with n_T is larger than 18, Computer Algebra Software (CAS) should be used to perform the capacity calculation. For example, Mathematica or Maple are powerful CAS's that can be forced to use arbitrary precision arithmetic when performing calculations, producing results to the exact precision requested.

TABLE 4.3 Average capacities of asymptotic MIMO systems, for an average SNR at each receiver branch of 5 dB.

MIMO type	C_{avg} (bits/s/Hz)	C_{avg}/sub-channel	$C_{\lim}$ (bits/s/Hz)	$C_{\lim}$/sub-channel
Gaussian SISO	2.057	2.057	-	-
$n_R = 1$; $n_T = 1$	1.716	1.716	2.057	2.057
$n_R = 2$; $n_T = 2$	3.306	1.653	4.115	2.057
$n_R = 4$; $n_T = 4$	6.553	1.638	8.229	2.057
$n_R = 8$; $n_T = 8$	13.080	1.635	16.459	2.057
$n_R = 16$; $n_T = 16$	26.146	1.6341	32.918	2.057
$n_R = 32$; $n_T = 32$	52.286	1.6339	65.836	2.057

pacity. Notice that duplicating the number of transmitter and receiver antennas almost duplicates the average capacity C_{avg} and, as the number of antennas increases, the duplication becomes more accurate. This fact is a manifestation of the linearity of the capacity with increasing array sizes (i.e. when $\max(n_R, n_T)$ is large) and it is justified by (4.53). This is in fact what happens with $C_{\lim}$ (in the fourth column of the table), the limit of the capacity for very large n_T. Table 4.3 also shows the capacity per sub-channel (or eigenmode) of the MIMO system, which is essentially $C_{avg} \,/\, \min(n_R, n_T)$, revealing that it decreases monotonically and is always less than the Gaussian and Rayleigh SISO capacities. For further

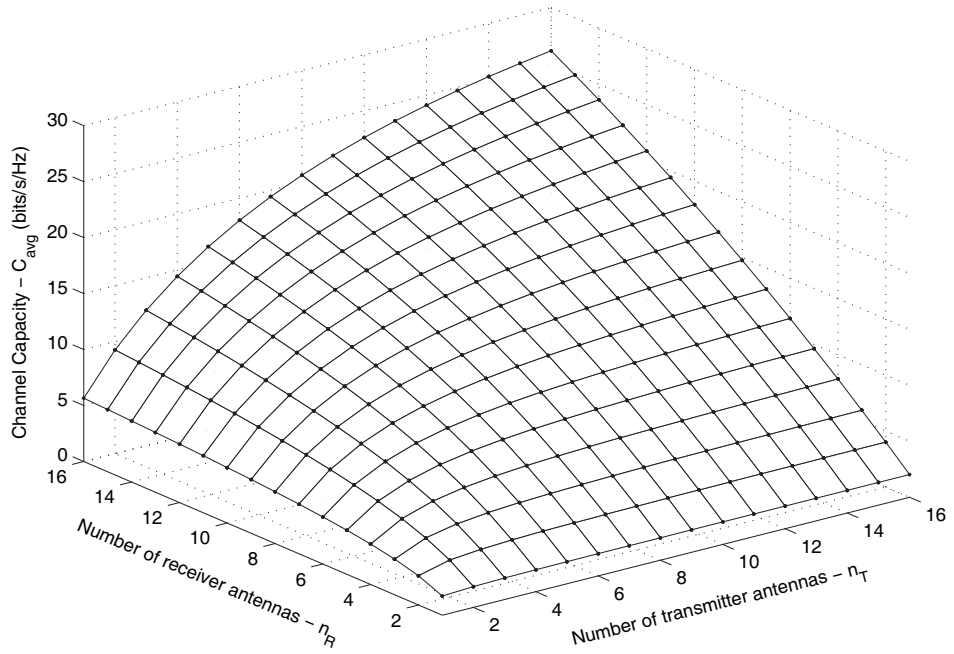

FIGURE 4.6 Average MIMO capacity as a function of the number of transmitter and receiver antennas, for an average SNR of 5 dB at each receiver branch.

reference, the achievable capacities as a function of the array sizes at the transmitter and receiver sides are depicted in Figures 4.6 and 4.7. The plots in Figure 4.7 are bidimensional replicas of Figure 4.6, and can be used to discover the array sizes that achieve a given information-theoretic capacity. All these plots assume independent fading, and thus repre-

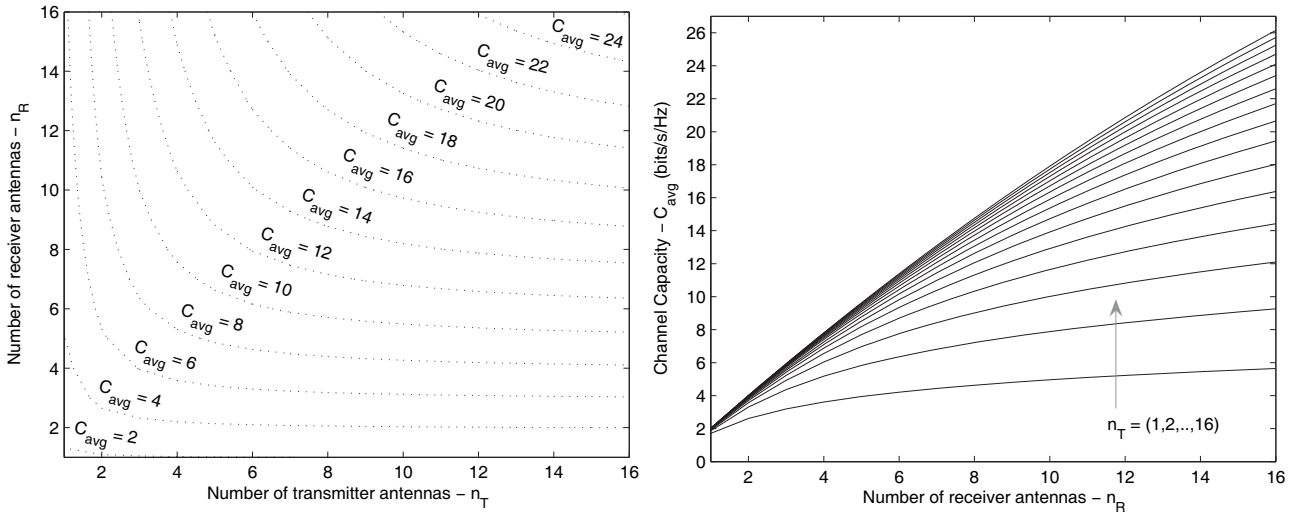

FIGURE 4.7 Capacity as a function of the number of transmitter and receiver antennas, for an average SNR at each receiver branch of 5 dB. The left figure retrieves some isolines from Figure 4.6.

sent a best-case scenario.

The bottom line of this discussion is that using appropriate signal processing techniques together with multi-antenna systems that circumvent, in the better possible way, the spatial correlation between the antennas (among other issues), it may be possible to equip multiuser wireless mobile networks with remarkable data transfer capabilities still unknown to the present time.

4.5.7 CORRELATED FADING

Estimation of the spectral efficiency limit (or capacity) of multiple-input multiple-output wireless channels under the most general conditions is a current issue within the research community. One of these conditions is the existence of signal correlation between the antennas of a moving receiver. The concept of MIMO channels has emerged from the principles of diversity and independence between the antennas of transmitters and receivers. However, practical evidence has shown that strict independence is difficult to attain, especially when either or both the transmitters and receivers are located in empty, isolated areas [55], [56]. It is expected that the antennas be spaced several wavelengths apart so that the correlation becomes negligible, but often these spacings are not feasible because of size restrictions in communication units. These restrictions will inevitably lead to an adjustment of the enormous capacities promised by [10], [11]. This adjustment must be theoretically quantified, and should be as generic as possible (i.e. irrespective of the correlation scenario).

The author has found several approaches to this problem in the literature, roughly divided into two groups: asymptotic (see, e.g., [57]) and exact (see, e.g., [58]). The second group shall be tackled. Some papers have appeared that successfully present some interesting results, but deal with the problem in a limited manner. For instance, [59] addresses the correlated case for large matrix sizes, solely. Another paper worth mentioning is [60]. It discusses power allocation when the correlation is transmitter-sided and partial channel state information (CSI) is known at the transmitter. One shall not dwell on this issue. Oth-

er approaches will be referred accordingly throughout the text.

Evaluating the MIMO capacity when the signals at the receiver antennas do not fade independently is a more complicated task. In this situation, the correlation matrix $\mathbf{\Sigma}$ is not the identity matrix as in the case of independent fading, and therefore the integral in the joint eigenvalue density is not readily obtained. The generic pdf is given by (4.157) as

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = \frac{\prod_{i=1}^{n_R} \lambda_i^{(n_T - n_R)} \prod_{i<j}^{n_R} (\lambda_j - \lambda_i)^2}{(2\pi)^{n_R}\, n_R! |\det \mathbf{\Sigma}|^{n_T}\, \tilde{\Gamma}_{n_R}(n_T)} \int_{\tilde{U}(n_R)} \exp\left[-tr\left(\mathbf{\Sigma}^{-1}\mathbf{Q}\mathbf{\Lambda}^{o}\mathbf{Q}^{H}\right)\right](d\mathbf{Q})^{\wedge} \tag{4.192}$$

The integral to be evaluated is the famous Itzykson-Zuber integral [61]. In what follows it will be shown that it is solvable by means of a powerful method that, at the same time, is beautiful and clear. It is based on the theory of group representations and character expansions.

Little was said so far about $\mathbf{\Sigma}$, except that it was given by

$$\mathbf{\Sigma} = E\left[\mathbf{h}_j \mathbf{h}_j^{H}\right] \tag{4.193}$$

where the $\mathbf{h}_j$'s are the columns of the channel matrix $\mathbf{H}$. From (4.193) one deduces the following about $\mathbf{\Sigma}$:

- it is self-adjoint or, in other words, it is an hermitian matrix;
- it has an eigenvalue decomposition $\mathbf{\Sigma} = \mathbf{E}\mathbf{\Upsilon}\mathbf{E}^{H}$, where $\mathbf{E}$ is an $n_R \times n_R$ unitary matrix and $\mathbf{\Upsilon} = diag(\upsilon_1,...,\upsilon_{n_R})$ with real eigenvalues;
- since it is a covariance matrix, one may expect that it can also be factorized as $\mathbf{\Sigma} = \mathbf{B}^{H}\mathbf{B}$, where $\mathbf{B}$ is a scaled matrix of multivariate samples (recall (4.143)); hence, all the eigenvalues of $\mathbf{\Sigma}$ are nonnegative because

$$\upsilon_i = \frac{\mathbf{x}_i^{H}\mathbf{B}^{H}\mathbf{B}\mathbf{x}_i}{\mathbf{x}_i^{H}\mathbf{x}_i} = \frac{\|\mathbf{B}\mathbf{x}_i\|^2}{\|\mathbf{x}_i\|^2} \geq 0 \tag{4.194}$$

- since it must be assumed from (4.192) that $\mathbf{\Sigma}$ is nonsingular, it must also follow that

$$\mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{x} = \mathbf{0} \tag{4.195}$$

  which implies that

$$\mathbf{A}\mathbf{x} = \upsilon\mathbf{x} \neq \mathbf{0} \quad \text{for } \mathbf{x} \neq \mathbf{0} \tag{4.196}$$

  and therefore, together with (4.194), one concludes that the eigenvalues of $\mathbf{\Sigma}$ must be real and positive;

Using the above conclusions, it is possible to rewrite the eigenvalue pdf in (4.192) in simplified form as

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon}) \int_{\tilde{U}(n_R)} \exp\left[-tr\left(\mathbf{\Upsilon}^{-1}\mathbf{E}^{H}\mathbf{Q}\mathbf{\Lambda}^{o}\mathbf{Q}^{H}\mathbf{E}\right)\right](d\mathbf{Q})^{\wedge} \tag{4.197}$$

where the introduced constant is given by

$$G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon}) = \frac{\prod_{i=1}^{n_R} \lambda_i^{n_T - n_R} \prod_{i<j}^{n_R} (\lambda_j - \lambda_i)^2}{(2\pi)^{n_R}\, n_R!\, \tilde{\Gamma}_{n_R}(n_T) \prod_{i=1}^{n_R} \upsilon_i^{n_T}} \tag{4.198}$$

Moreover, putting $\mathbf{\Lambda}^{o} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{H}$, where $\mathbf{P}$ is an orthogonal symmetric permutation matrix, writing $\mathbf{U} = \mathbf{E}^{H}\mathbf{Q}\mathbf{P}$, and noting that $(d\mathbf{U})^{\wedge} = |\det \mathbf{E}|^{2n_R}(\det \mathbf{P})^{2n_R}(d\mathbf{Q})^{\wedge} = (d\mathbf{Q})^{\wedge}$, leads to

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})\int_{\tilde{U}(n_R)} \exp\left[-tr\left(\mathbf{\Upsilon}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{H}\right)\right](d\mathbf{U})^{\wedge} \tag{4.199}$$

which shows that the eigenvalue pdf is a symmetric function of the eigenvalues of the covariance matrix $\mathbf{\Sigma}$ and the Wishart matrix $\mathbf{A}$.

To make the integration in (4.192) as easy as possible, one may substitute the integrand function by its series expansion, yielding

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})\sum_{k=0}^{\infty}\frac{1}{k!}\int_{\tilde{U}(n_R)}\left[tr\left(-\mathbf{\Upsilon}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{H}\right)\right]^{k}(d\mathbf{U})^{\wedge} \tag{4.200}$$

where the integrand is now a polynomial function in the elements of matrices $\mathbf{\Lambda}$ and $\mathbf{\Upsilon}$. Now, since the trace of a matrix equals the sum of its eigenvalues, we may notice that

$$\left[tr\left(-\mathbf{\Upsilon}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{H}\right)\right]^{k} = [tr(\mathbf{X})]^{k} = \left(x_1 + x_2 + ... + x_{n_R}\right)^{k} \tag{4.201}$$

where $x_1, x_2, ..., x_{n_R}$ are the eigenvalues of the matrix $\mathbf{X} = -\mathbf{\Upsilon}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{H}$. Expanding (4.201) results in a finite linear combination of monomials of degree k, i.e.

$$[tr(\mathbf{X})]^{k} = \sum d_{k_1...k_{n_R}} x_1^{k_1} x_2^{k_2} \cdots x_{n_R}^{k_{n_R}} \tag{4.202}$$

where the sum is over all n_R-tuples $(k_1, k_2, ..., k_{n_R})$ that sum to k. As $\mathbf{X}$ varies, the monomials in (4.202) generate the space V_k of all polynomials of degree k, which are homogeneous and symmetric in the eigenvalues of $\mathbf{X}$. Now, in the same way as any vector space can be decomposed into a set of irreducible invariant subspaces under a given linear operator, so does V_k has a decomposition into a set of irreducible subspaces V_κ, invariant under the unitary group. More formally, this means that

$$V_k = \bigoplus_{\kappa \vdash k} V_\kappa \tag{4.203}$$

where the symbol $\bigoplus_{\kappa \vdash k}$ denotes the *direct sum* of the subspaces over all distinct partitions $\kappa = (k_1, k_2, ..., k_{n_R})$, $0 \le k_1 \le k_2 \le ... \le k_{n_R}$, of the integer k. Only the distinct partitions of k need to be considered, given that the monomials in (4.202) can be grouped into equivalence classes in relation to the space they generate. Each subspace V_κ can be generated by a homogeneous symmetric polynomial $C_\kappa(\mathbf{X})$ (commonly referred in the literature as the *zonal polynomial*) in the eigenvalues of the matricial argument $\mathbf{X}$. Hence, (4.202) can be alternatively expressed as

$$[tr(\mathbf{X})]^{k} = \sum_{\kappa \vdash k} C_\kappa(\mathbf{X}) \tag{4.204}$$

The reason for expressing (4.202) as (4.204) is that a fundamental property of the zonal polynomials is given by the integral, proved in [43]:

$$\int_{\tilde{U}(n)} C_\kappa(\mathbf{A}\mathbf{U}\mathbf{B}\mathbf{U}^{H})(d\mathbf{U})^{\wedge} = \text{Volume}\left[\tilde{U}(n)\right]\frac{C_\kappa(\mathbf{A})C_\kappa(\mathbf{B})}{C_\kappa(\mathbf{I}_n)} \tag{4.205}$$

and thus the substitution of (4.204) into (4.200) gives

$$\begin{aligned} f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) &= \text{Volume}\big[\widetilde{U}(n)\big] G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon}) \sum_{k=0}^{\infty}\sum_{\kappa} \frac{C_\kappa\left(-\mathbf{\Upsilon}^{-1}\right) C_\kappa(\mathbf{\Lambda})}{C_\kappa\left(\mathbf{I}_{n_R}\right) k!} \\ &= \text{Volume}\big[\widetilde{U}(n)\big] G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})\, {}_0\tilde{F}_0(-\mathbf{\Upsilon}^{-1},\mathbf{\Lambda}) \end{aligned} \tag{4.206}$$

where ${}_0\tilde{F}_0(\mathbf{A},\mathbf{B})$ is the simplest case of the *complex hypergeometric function of two matrix arguments*, defined generically as

$$\begin{aligned} {}_p\tilde{F}_q(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{X},\mathbf{Y}) &= \sum_{k=0}^{\infty}\sum_{\kappa\vdash k} \frac{(a_1)_\kappa \cdots (a_p)_\kappa}{(b_1)_\kappa \cdots (b_q)_\kappa} \frac{C_\kappa(\mathbf{X}) C_\kappa(\mathbf{Y})}{C_\kappa(\mathbf{I}_n)\, k!} \\ (a)_\kappa &= \prod_{i=1}^{n} \Gamma(a-i+1+k_i)/\Gamma(a-i+1) \end{aligned} \tag{4.207}$$

Even though Equation (4.206) is an analytical expression for the joint pdf of the eigenvalues, it is expressed as a function of an infinite series which has poor convergence properties. In fact, the hypergeometric functions of matrix argument are hard to approximate, and require the evaluation of a very high number of its terms ([62], [63] and [64]. Discovering an alternative expression for ${}_0\tilde{F}_0(\mathbf{A},\mathbf{B})$ is therefore of practical importance. After some study, one realizes that the answer may be found in the context of the *group theory of representations and characters*. For a detailed discussion of these topics and much more besides, the author vividly recommends [65], [66], [67], and [68]. Some important results from this theory are now outlined:

1. $\text{GL}(\mathcal{V})$ denotes the group of invertible linear operators $\mathbf{T}$ (e.g. invertible matrices) that act on the vector space $\mathcal{V}$;

2. if $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for $\mathcal{V}$, $\mathbf{x} \in \mathcal{V}$, and $\mathbf{y} = \mathbf{T}(\mathbf{x})$, $\mathbf{T} \in \text{GL}(\mathcal{V})$, then we may also write $[\mathbf{y}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$, where $[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{y}]_{\mathcal{B}}$ denote the coordinate vectors with respect to $\mathcal{B}$, and $[\mathbf{T}]_{\mathcal{B}} = \left([\mathbf{T}(\mathbf{v}_1)]_{\mathcal{B}} \,|\, [\mathbf{T}(\mathbf{v}_2)]_{\mathcal{B}} \,|\cdots|\, [\mathbf{T}(\mathbf{v}_n)]_{\mathcal{B}}\right)$ is the coordinate matrix of $\mathbf{T}$ with respect to $\mathcal{B}$; furthermore we have the composition of linear operators

$$[\mathbf{T}_2 \circ \mathbf{T}_1]_{\mathcal{B}} = [\mathbf{T}_2]_{\mathcal{B}}[\mathbf{T}_1]_{\mathcal{B}} \tag{4.208}$$

3. a subspace $\mathcal{U}$ of $\mathcal{V}$ is an invariant subspace under $\mathbf{T}$ whenever $\mathbf{T}(\mathcal{U}) \subseteq \mathcal{U}$; in this case $\mathbf{T}$ can be considered as a linear operator on $\mathcal{U}$, and is denoted $\mathbf{T}_{/\mathcal{U}}$; the trivial subspaces of $\mathcal{V}$ are the subspace $\{0\}$ and $\mathcal{V}$ itself; a subspace $\mathcal{V}$ with no non-trivial invariant subspaces is called *irreducible*;

4. let $\mathcal{V}_i$ be an irreducible subspace of $\mathcal{V}$, and let $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_r$; if $\mathcal{B}_i$ is a basis for $\mathcal{V}_i$, then $\mathcal{B}' = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$ is a basis for $\mathcal{V}$; in this case, the coordinate matrix of a linear operator $\mathbf{T}$ with respect to $\mathcal{B}'$ can be expressed **equivalently** as a block diagonal matrix

$$[\mathbf{T}]_{\mathcal{B}'} = \mathbf{E}^{-1}[\mathbf{T}]_{\mathcal{B}}\mathbf{E} = \begin{pmatrix} \left[\mathbf{T}_{/\mathcal{V}_1}\right]_{\mathcal{B}_1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \left[\mathbf{T}_{/\mathcal{V}_r}\right]_{\mathcal{B}_r} \end{pmatrix} \tag{4.209}$$

where $\mathbf{E}$ is the change of basis matrix; in turn, this means that

$$[\mathbf{y}]_{\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}'}[\mathbf{x}]_{\mathcal{B}'} \rightarrow [\mathbf{y}]_{\mathcal{B}_i} = \left[\mathbf{T}_{/\mathcal{V}_i}\right]_{\mathcal{B}_i}[\mathbf{x}]_{\mathcal{B}_i} \tag{4.210}$$

i.e. the linear operator $\mathbf{T}$ can be decomposed into a set of irreducible linear operators $\mathbf{T}_{/\mathcal{V}_i}$;

5. the homomorphism $\rho : G \rightarrow \mathrm{GL}(\mathcal{V})$ is a linear action of a group G on the vector space $\mathcal{V}$; if $g \in G$, then $\rho(g) = \mathbf{T}_g$ is a linear operator on the vector space $\mathcal{V}$; ρ is called a *representation* of G;

6. from (**4.**) it follows that $\rho(g)$ can be decomposed into a set of linear operators $\rho(g)_{/\mathcal{V}_i}$; these linear operators are also representations because

$$\rho(gh)_{/\mathcal{V}_i} = [\rho(g) \circ \rho(h)]_{/\mathcal{V}_i} = \left[\rho(g)_{/\mathcal{V}_i}\right] \circ \left[\rho(h)_{/\mathcal{V}_i}\right] \tag{4.211}$$

and thus the homomorphisms $\rho_{/\mathcal{V}_i} : G \rightarrow \mathrm{GL}(\mathcal{V}_i) : g \rightarrow \rho(g)_{/\mathcal{V}_i}$ are the irreducible subrepresentations of ρ associated with the irreducible invariant subspaces of $\mathcal{V}$; in other words, there is an **equivalent** representation such that

$$[\rho(g)]_{\mathcal{B}'} = \mathbf{E}^{-1}[\rho(g)]_{\mathcal{B}}\mathbf{E} = \begin{pmatrix} \left[\rho_{/\mathcal{V}_1}(g)\right]_{\mathcal{B}_1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \left[\rho_{/\mathcal{V}_r}(g)\right]_{\mathcal{B}_r} \end{pmatrix} \tag{4.212}$$

7. now, for simplicity, write $\boldsymbol{\pi}_i(g) = \left[\rho_{/\mathcal{V}_i}(g)\right]_{\mathcal{B}_i}$ and call it the i-th irreducible matrix representation of the group element $g \in G$; then $\boldsymbol{\pi}_i(g)$ is equivalent to an irreducible unitary matrix representation, because writing

$$\boldsymbol{\pi}'_i(g) = \mathbf{E}^{-1}\boldsymbol{\pi}_i(g)\mathbf{E} \quad \rightarrow \quad \boldsymbol{\pi}'_i(g)\left[\boldsymbol{\pi}'_i(g)\right]^H = \mathbf{E}^{-1}\boldsymbol{\pi}_i(g)\mathbf{E}\mathbf{E}^H\left[\boldsymbol{\pi}_i(g)\right]^H\left[\mathbf{E}^{-1}\right]^H \tag{4.213}$$

and choosing $\mathbf{E} = \mathbf{U}\mathbf{D}^{1/2}$ where $\mathbf{U}$ is a unitary matrix and $\mathbf{D}$ is a diagonal matrix with real and positive entries, we have

$$\boldsymbol{\pi}'_i(g)\left[\boldsymbol{\pi}'_i(g)\right]^H = \mathbf{D}^{-1/2}\mathbf{U}^H\left(\boldsymbol{\pi}_i(g)\mathbf{U}\mathbf{D}\mathbf{U}^H\left[\boldsymbol{\pi}_i(g)\right]^H\right)\mathbf{U}\mathbf{D}^{-1/2} \tag{4.214}$$

- therefore, one simply has to define the eigenvalue decomposition

$$\mathbf{U}\mathbf{D}\mathbf{U}^H = \int_G \boldsymbol{\pi}_i(x)\left[\boldsymbol{\pi}_i(x)\right]^H dx \tag{4.215}$$

which substituted into (4.214) yields

$$\begin{aligned} \boldsymbol{\pi}'_i(g)\left[\boldsymbol{\pi}'_i(g)\right]^H &= \mathbf{D}^{-1/2}\mathbf{U}^H\left(\int_G \boldsymbol{\pi}_i(g)\boldsymbol{\pi}_i(x)\left[\boldsymbol{\pi}_i(g)\boldsymbol{\pi}_i(x)\right]^H dx\right)\mathbf{U}\mathbf{D}^{-1/2} \\ &= \mathbf{D}^{-1/2}\mathbf{U}^H\left(\int_G \boldsymbol{\pi}_i(gx)\left[\boldsymbol{\pi}_i(gx)\right]^H dx\right)\mathbf{U}\mathbf{D}^{-1/2} \\ &= \mathbf{D}^{-1/2}\mathbf{U}^H\left(\int_G \boldsymbol{\pi}_i(x)\left[\boldsymbol{\pi}_i(x)\right]^H dx\right)\mathbf{U}\mathbf{D}^{-1/2} = \mathbf{D}^{-1/2}\mathbf{D}\mathbf{D}^{-1/2} = \mathbf{I} \end{aligned} \tag{4.216}$$

proving that $\boldsymbol{\pi}'_i(g)$ is unitary; from this point on we shall assume that the irreducible matrix representations are all unitary;

8. let $\boldsymbol{\pi}_i(g)$ and $\boldsymbol{\pi}_j(g)$ be two inequivalent, unitary, and irreducible matrix representations

of a group G; also, let $U \subseteq G$ be a **compact topological group** (so that it is measurable), and $\mathbf{A}$ and $\mathbf{B}$ be generic matrices such that

$$\mathbf{B} = \int_U \boldsymbol{\pi}_i(u)\mathbf{A}\left[\boldsymbol{\pi}_j(u)\right]^H du \tag{4.217}$$

- multiply (4.217) on the left by $\boldsymbol{\pi}_i(x)$ and on the right by $[\boldsymbol{\pi}_i(x)]^H$, yielding

$$\begin{aligned}\boldsymbol{\pi}_i(x)\mathbf{B}\left[\boldsymbol{\pi}_j(x)\right]^H &= \boldsymbol{\pi}_i(x)\left(\int_U \boldsymbol{\pi}_i(u)\mathbf{A}\left[\boldsymbol{\pi}_j(u)\right]^H du\right)\left[\boldsymbol{\pi}_j(x)\right]^H \\ &= \int_U \boldsymbol{\pi}_i(xu)\mathbf{A}\left[\boldsymbol{\pi}_j(xu)\right]^H du = \int_U \boldsymbol{\pi}_i(u)\mathbf{A}\left[\boldsymbol{\pi}_j(u)\right]^H du = \mathbf{B}\end{aligned} \tag{4.218}$$

- thus one must have $\mathbf{B} = \lambda\delta_{ij}\mathbf{I}$, and choosing $\mathbf{A} = [\delta_{ln}]$

$$[\mathbf{B}]_{km} = \lambda\delta_{ij}\delta_{km} = \int_U [\boldsymbol{\pi}_i(u)]_{kl}\left[\boldsymbol{\pi}_j(u)\right]^*_{mn} du \tag{4.219}$$

- to find λ we put $i = j$ and $k = m$, and sum over k, which gives

$$\sum_{k=1}^{d_i}\int_U [\boldsymbol{\pi}_i(u)]_{kl}\left[\boldsymbol{\pi}_i(u)\right]^*_{kn} du = \lambda d_i \tag{4.220}$$

where d_i is the size of $\boldsymbol{\pi}_i(g)$, i.e. the dimension of the irreducible subspace $\mathcal{V}_i$; now use the simple identity

$$\int_U [\boldsymbol{\pi}_i(u)]^H \boldsymbol{\pi}_i(u)du = V_U\mathbf{I} \;\rightarrow\; \sum_{k=1}^{d_i}\int_U [\boldsymbol{\pi}_i(u)]_{kl}\left[\boldsymbol{\pi}_i(u)\right]^*_{kn} du = V_U\delta_{ln} \tag{4.221}$$

where V_U is the volume of U; comparison with (4.220) leads to $\lambda = V_U\delta_{ln}/d_i$, and final insertion into (4.219) reveals the so called **Schur orthogonality relations**

$$\left\langle [\boldsymbol{\pi}_i(u)]_{kl}, [\boldsymbol{\pi}_j(u)]_{mn}\right\rangle = \frac{1}{\mathrm{V}_U}\int_U [\boldsymbol{\pi}_i(u)]_{kl}\left[\boldsymbol{\pi}_j(u)\right]^*_{mn} du = \frac{1}{d_i}\delta_{ij}\delta_{km}\delta_{ln} \tag{4.222}$$

- these relations show that the entries $[\boldsymbol{\pi}_i(u)]_{kl}$ of irreducible representations of a compact topological group are not only independent functions in the space of all functions $f : U \rightarrow \mathbb{C}$, but are also orthogonal functions;

9. the *character* $\chi : G \rightarrow \mathbb{C}$ of a representation $\rho : G \rightarrow \mathrm{GL}(\mathcal{V})$ is defined as the trace of the coordinate matrix associated with $\rho(g)$, i.e.

$$\chi(g) = tr([\rho(g)]_{\mathcal{B}}) = tr(\boldsymbol{\pi}(g)) = \sum_{i=1}^{r} tr\left(\boldsymbol{\pi}_i(g)\right) = \sum_{i=1}^{r}\chi_{\mathcal{V}_i}(g) \tag{4.223}$$

where $g \in G$ and $\mathcal{B}$ is a basis for the vector space $\mathcal{V}$; $\chi_{\mathcal{V}_i}$ is the (irreducible) character associated with the irreducible representation $\rho_{\mathcal{V}_i}$; the character of the identity element gives the **dimension** of $\mathcal{V}_i$:

$$d_i = \chi_{\mathcal{V}_i}(1) = tr\left(\left[\rho_{\mathcal{V}_i}(1)\right]_{\mathcal{B}_i}\right) = tr\left(\mathbf{I}_{d_i}\right) \tag{4.224}$$

10. since the irreducible characters are given by

$$\chi_{\mathcal{V}_i}(g) = \sum_{k=1}^{d_i} [\boldsymbol{\pi}_i(g)]_{kk} \tag{4.225}$$

they are clearly orthonormal functions, because from (4.218) we have

$$\begin{aligned} \left\langle \chi_{\mathcal{V}_i}(g), \chi_{\mathcal{V}_j}(g) \right\rangle &= \frac{1}{\mathrm{V}_G} \int_G \chi_{\mathcal{V}_i}(g) \chi^*_{\mathcal{V}_j}(g) dg = \frac{1}{\mathrm{V}_G} \int_G \sum_{k=1}^{d_i} [\boldsymbol{\pi}_i(g)]_{kk} \sum_{p=1}^{d_j} \left[\boldsymbol{\pi}^*_j(g)\right]_{pp} dg \\ &= \sum_{k=1}^{d_i} \sum_{p=1}^{d_j} \frac{1}{\mathrm{V}_G} \int_G [\boldsymbol{\pi}_i(g)]_{kk} \left[\boldsymbol{\pi}^*_j(g)\right]_{pp} dg = \sum_{k=1}^{d_i} \sum_{p=1}^{d_j} \frac{1}{d_i} \delta_{ij} \delta_{kp} = \delta_{ij} \end{aligned} \tag{4.226}$$

11. another fundamental identity relating characters can be obtained as follows:

- let G be a group and $g_a, g_b \in G$; also, let $U \subseteq G$ be a compact topological group and let $u \in U$; now suppose we want to compute the integral

$$\int_U \chi_{\mathcal{V}_i}(g_a u g_b u^{-1}) du \tag{4.227}$$

- substitution of (4.225) yields

$$\begin{aligned} \int_U \chi_{\mathcal{V}_i}(g_a u g_b u^{-1}) du &= \sum_{k=1}^{d_i} \int_U [\boldsymbol{\pi}_i(g_a u g_b u^{-1})]_{kk} \, du \\ &= \sum_{k=1}^{d_i} \int_U [\boldsymbol{\pi}_i(g_a) \boldsymbol{\pi}_i(u) \boldsymbol{\pi}_i(g_b) \boldsymbol{\pi}_i^H(u)]_{kk} \, du \end{aligned} \tag{4.228}$$

- if the integrand is expanded as

$$[\boldsymbol{\pi}_i(g_a) \boldsymbol{\pi}_i(u) \boldsymbol{\pi}_i(g_b) \boldsymbol{\pi}_i^H(u)]_{kk} = \sum_{p=1}^{d_i} \sum_{q=1}^{d_i} \sum_{s=1}^{d_i} [\boldsymbol{\pi}_i(g_a)]_{kp} [\boldsymbol{\pi}_i(u)]_{pq} [\boldsymbol{\pi}_i(g_b)]_{qs} [\boldsymbol{\pi}_i^*(u)]_{ks} \tag{4.229}$$

then the integral can be rewritten as

$$\int_U \chi_{\mathcal{V}_i}(g_a u g_b u^{-1}) du = \sum_{k=1}^{d_i} \sum_{p=1}^{d_i} \sum_{q=1}^{d_i} \sum_{s=1}^{d_i} [\boldsymbol{\pi}_i(g_a)]_{kp} [\boldsymbol{\pi}_i(g_b)]_{qs} \int_U [\boldsymbol{\pi}_i(u)]_{pq} [\boldsymbol{\pi}_i^*(u)]_{ks} \, du \tag{4.230}$$

- now, the integration can be performed directly using the orthogonality relations in (4.218), yielding a nonzero result only for $p = k$ and $q = s$, i.e.

$$\begin{aligned} \int_U \chi_{\mathcal{V}_i}(g_a u g_b u^{-1}) du &= \mathrm{V}_U \frac{1}{d_i} \sum_{k=1}^{d_i} [\boldsymbol{\pi}_i(g_a)]_{kk} \sum_{q=1}^{d_i} [\boldsymbol{\pi}_i(g_b)]_{qq} \\ &= \mathrm{V}_U \frac{1}{d_i} tr(\boldsymbol{\pi}_i(g_a)) tr(\boldsymbol{\pi}_i(g_b)) = \mathrm{V}_U \frac{\chi_{\mathcal{V}_i}(g_a) \chi_{\mathcal{V}_i}(g_b)}{\chi_{\mathcal{V}_i}(1)} \end{aligned} \tag{4.231}$$

where V_U is the volume of the group U; this is a very important result that will be useful later;

12. a *class function* on G is a function $\varphi : G \to \mathbb{C}$ such that $\varphi(g) = \varphi(h)$ whenever $h = ugu^{-1}$, for $g, h, u \in G$; class functions are constant on the *conjugacy classes* $g^G = \{ugu^{-1} : g \in G\}$; clearly, the irreducible characters of a representation ρ are (linearly independent) class

functions since

$$\begin{aligned}\chi_{/\mathcal{V}_i}(h) &= tr([\rho_{/\mathcal{V}_i}(h)]_{\mathcal{B}_i}) = tr([\rho_{/\mathcal{V}_i}(u)\circ\rho_{/\mathcal{V}_i}(g)\circ\rho_{/\mathcal{V}_i}(u^{-1})]_{\mathcal{B}_i})\\ &= tr([\rho_{/\mathcal{V}_i}(u)]_{\mathcal{B}_i}[\rho_{/\mathcal{V}_i}(g)]_{\mathcal{B}_i}[\rho_{/\mathcal{V}_i}(u)]_{\mathcal{B}_i}^{-1})\\ &= tr([\rho_{/\mathcal{V}_i}(g)]_{\mathcal{B}_i}) = \chi_{/\mathcal{V}_i}(g)\end{aligned} \tag{4.232}$$

13. furthermore, it is not difficult to prove that the number of irreducible characters is equal to the number r of conjugacy classes, and hence they form an orthonormal basis for the space of all class functions on G, i.e.

$$\varphi(g) = \sum_{i=1}^{r} c_i \chi_{/\mathcal{V}_i}(g), \qquad c_i = \left\langle \varphi(g), \chi_{/\mathcal{V}_i}(g) \right\rangle \tag{4.233}$$

where putting $\varphi(g) = \delta(g,h)$, and using (4.226), gives the **completeness relation**

$$\sum_{i=1}^{r} \chi_{/\mathcal{V}_i}(g)\chi^{*}_{/\mathcal{V}_i}(h) = \mathrm{V}_G\,\delta(g,h) \tag{4.234}$$

Having described these fundamental aspects of the theory of representations and characters, it is now easy to find their application in the problem at hand. Recall that the main objective is to find a mathematically tractable expression for the function $f(\mathbf{X}) = \exp[tr(\mathbf{X})]$, $\mathbf{X} \in \mathrm{GL}(n_R, \mathbb{C})$. Letting $\mathcal{V}$ denote the function space generated by the function f, we may say that a transformation

$$\begin{aligned}\mathbf{X} &\rightarrow \mathbf{P}(\mathbf{X})\\ f(\mathbf{X}) &\rightarrow f\left(\mathbf{P}^{-1}(\mathbf{X})\right)\end{aligned} \tag{4.235}$$

where $\mathbf{P} \in \mathrm{GL}(n_R, \mathbb{C})$, acts on $\mathcal{V}$ by means of the function f. Hence it is reasonable to specify a mapping

$$\rho : \mathrm{GL}(n_R, \mathbb{C}) \rightarrow \mathrm{GL}(\mathcal{V}) : \mathbf{P} \mapsto \rho(\mathbf{P}) \tag{4.236}$$

which defines a representation of $\mathbf{P}$ in function space, but now by means of the linear operator $\rho(\mathbf{P})$ which will act on $\mathcal{V}$. As a result, we may:

- decompose $\mathcal{V}$ into a set of irreducible invariant subspaces $\mathcal{V}_\kappa$ under $\rho(\mathbf{P})$;
- for each κ, discover the irreducible representation $\rho_{/\mathcal{V}_\kappa}(\mathbf{P})$ associated with $\mathcal{V}_\kappa$;
- find a function basis $\mathcal{B}_\kappa = \{f_1, f_2, ..., f_{d_\kappa}\}$ for each $\mathcal{V}_\kappa$ such that $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots$, and consequently

$$\left[f\left(\mathbf{P}^{-1}(\mathbf{X})\right)\right]_{\mathcal{B}} = [\rho(\mathbf{P})(f(\mathbf{X}))]_{\mathcal{B}} = [\rho(\mathbf{P})]_{\mathcal{B}}[f(\mathbf{X})]_{\mathcal{B}} = \begin{pmatrix} \left[\rho_{/\mathcal{V}_1}(\mathbf{P})\right]_{\mathcal{B}_1} & \mathbf{0} & \cdots \\ \mathbf{0} & \left[\rho_{/\mathcal{V}_2}(\mathbf{P})\right]_{\mathcal{B}_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} [f(\mathbf{X})]_{\mathcal{B}_1} \\ [f(\mathbf{X})]_{\mathcal{B}_2} \\ \vdots \end{pmatrix} \tag{4.237}$$

- the matrix representation $\boldsymbol{\pi}_\kappa(\mathbf{P}) = \left[\rho_{/\mathcal{V}_\kappa}(\mathbf{P})\right]_{\mathcal{B}_\kappa}$ is called functional since its entries are coordinates with respect to $\mathcal{B}$ of functions in the elements of $\mathbf{P}$;

Considering, for instance, the case of a similarity transformation $\mathbf{P}(\mathbf{X}) = \mathbf{P}\mathbf{X}\mathbf{P}^{-1}$,

$\mathbf{P} \in \mathrm{GL}(n_R, \mathbb{C})$, it induces a transformation in function space given by

$$f\left(\mathbf{P}^{-1}(\mathbf{X})\right) = \exp\left[tr\left(\mathbf{P}^{-1}\mathbf{X}\mathbf{P}\right)\right] = \exp[tr(\mathbf{X})] = f(\mathbf{X}) \tag{4.238}$$

showing that f is a class function on $\mathrm{GL}(n_R, \mathbb{C})$, and hence from (4.233) it follows that

$$f(\mathbf{X}) = \exp[tr(\mathbf{X})] = \sum_{\kappa} c_\kappa \chi_{\mathcal{V}_\kappa}(\mathbf{X}) \tag{4.239}$$

The summation is, as before, over all irreducible representations indexed by κ. The functions $\chi_{\mathcal{V}_\kappa}(\mathbf{X}) = tr([\rho_{\mathcal{V}_\kappa}(\mathbf{X})]_{\mathcal{B}_\kappa})$ are the characters of the irreducible representations $\rho_{\mathcal{V}_\kappa}(\mathbf{X})$ of the linear group $\mathrm{GL}(n_R, \mathbb{C})$ of invertible $n_R \times n_R$ matrices $\mathbf{X}$ of complex entries. These can be found from [69] to be given by the Schur polynomials

$$\chi_{\mathcal{V}_\kappa}(\mathbf{X}) = s_\kappa(\mathbf{x}) = \frac{\Delta_\kappa(\mathbf{x})}{\Delta(\mathbf{x})} = \frac{\det\left[x_i^{k_j + n_R - j}\right]}{\det\left[x_i^{n_R - j}\right]} \tag{4.240}$$

where $\mathbf{x} = (x_1, x_2, ..., x_{n_R})$ is the vector of eigenvalues of the matrix $\mathbf{X}$, $\Delta(\mathbf{x}) = \det\left[x_i^{n_R - j}\right]$ is the *Vandermonde determinant*[1], and $\Delta_\kappa(\mathbf{x}) = \det\left[x_i^{k_j + n_R - j}\right]$ is the *generalized Vandermonde determinant*. In fact, it is well known (see e.g. [68], [70]) that the Schur polynomials constitute a basis for the vector space of homogeneous symmetric polynomials, which is exactly what is needed for each term in the power series expansion

$$\exp[tr(\mathbf{X})] = \sum_{k=0}^{\infty} \frac{1}{k!} [tr(\mathbf{X})]^k \tag{4.241}$$

At this point, we need to deduce the expansion of $f(\mathbf{X}) = \exp[tr(\mathbf{X})]$ into characters as given by (4.239), this way proving the decomposition of the function space $\mathcal{V}$ into irreducible invariants. First, since the trace of the matrix $\mathbf{X}$ equals the sum of its eigenvalues $(x_1, x_2, ..., x_{n_R})$, the function can be rewritten as

$$f(\mathbf{X}) = \exp[tr(\mathbf{X})] = e^{x_1} e^{x_2} \cdots e^{x_{n_R}} \tag{4.242}$$

Multiplying on both sides by the Vandermonde determinant in the eigenvalues x_i, yields

$$\det\left[x_i^{n_R - j}\right] \exp[tr(\mathbf{X})] = \det\left[x_i^{n_R - j}\right] e^{x_1} e^{x_2} \cdots e^{x_{n_R}} = \det\left[x_i^{n_R - j} e^{x_i}\right] \tag{4.243}$$

Moreover, expanding the exponential inside the determinant into series of powers gives

$$\det\left[x_i^{n_R - j}\right] \exp[tr(\mathbf{X})] = \det\left[\sum_{k=0}^{\infty} \frac{1}{k!} x_i^{k + n_R - j}\right] \tag{4.244}$$

To simplify (4.239) the determinant must be expanded using the standard sum over all permutations p of $(1, 2, ..., n_R)$

$$\det[a_{ij}] = \sum_p \epsilon(p) \prod_{i=1}^{n_R} a_{ip_i} = \sum_p \epsilon(p) \prod_{j=1}^{n_R} a_{p_j j} \tag{4.245}$$

---

1. The notation $\det[x_{ij}]$ refers to the determinant of the square matrix $\mathbf{X}$ whose entry in the i-th row and j-th column is x_{ij}.

where for briefness of notation $\epsilon(p) = \text{sign}(p)$. This yields

$$\begin{aligned}\det\left[x_i^{n_R-j}\right]\exp[tr(\mathbf{X})] &= \sum_p \epsilon(p)\prod_{j=1}^{n_R}\sum_{k_j=0}^{\infty}\frac{1}{k_j!}x_{p_j}^{k_j+n_R-j} \\ &= \sum_p \epsilon(p)\sum_{k_1=0}^{\infty}\cdots\sum_{k_{n_R}=0}^{\infty}\prod_{j=1}^{n_R}\frac{1}{k_j!}x_{p_j}^{k_j+n_R-j}\end{aligned} \tag{4.246}$$

Interchanging the order of summation in (4.246) we obtain

$$\begin{aligned}\det\left[x_i^{n_R-j}\right]\exp[tr(\mathbf{X})] &= \sum_{k_1=0}^{\infty}\cdots\sum_{k_{n_R}=0}^{\infty}\left[\prod_{j=1}^{n_R}\frac{1}{k_j!}\right]\left[\sum_p \epsilon(p)\prod_{j=1}^{n_R}x_{p_j}^{k_j+n_R-j}\right] \\ &= \sum_{k_1=0}^{\infty}\cdots\sum_{k_{n_R}=0}^{\infty}\left[\prod_{j=1}^{n_R}\frac{1}{k_j!}\right]\det\left[x_i^{k_j+n_R-j}\right]\end{aligned} \tag{4.247}$$

Now it is helpful to make the temporary change of variables

$$t_j = k_j + n_R - j \tag{4.248}$$

which produces

$$\det\left[x_i^{n_R-j}\right]\exp[tr(\mathbf{X})] = \sum_{t_1=n_R-1}^{\infty}\sum_{t_2=n_R-2}^{\infty}\cdots\sum_{t_{n_R}=0}^{\infty}\left[\prod_{j=1}^{n_R}\frac{1}{(t_j-n_R+j)!}\right]\det\left[x_i^{t_j}\right] \tag{4.249}$$

We notice from (4.246) that whenever two columns of the matrix $\left[x_i^{t_j}\right]$ are equal, the determinant $\det\left[x_i^{t_j}\right]$ will be zero, and hence the outer sum may be restricted to distinct t_j's. Moreover, instead of considering all sequences of distinct t_j's, one may consider solely the ordered sequences, and then sum over all permutations of each sequence, i.e.

$$\det\left[x_i^{n_R-j}\right]\exp[tr(\mathbf{X})] = \sum_{t_1>t_2>\cdots>t_{n_R}\geq 0}\sum_p \epsilon(p)\left[\prod_{j=1}^{n_R}\frac{1}{(t_{p_j}-n_R+j)!}\right]\det\left[x_i^{t_j}\right] \tag{4.250}$$

The factor $\epsilon(p)$ must be introduced because the sign of the determinant $\det\left[x_i^{t_j}\right]$ depends upon the order of the t_j's. The inner sum is now recognized as a determinant, and thus we write (4.250) as

$$\det\left[x_i^{n_R-j}\right]\exp[tr(\mathbf{X})] = \sum_{t_1>t_2>\cdots>t_{n_R}\geq 0}\det\left[\frac{1}{(t_j-n_R+i)!}\right]\det\left[x_i^{t_j}\right] \tag{4.251}$$

Finally, the change of variables in (4.250) can be reverted, which clearly leads to weakly decreasing sequences $k_1 \geq k_2 \geq \cdots \geq k_{n_R} \geq 0$ indexing the sum, i.e.

$$\det\left[x_i^{n_R-j}\right]\exp[tr(\mathbf{X})] = \sum_{k_1\geq k_2\geq\cdots\geq k_{n_R}\geq 0}\det\left[\frac{1}{(k_j+i-j)!}\right]\det\left[x_i^{k_j+n_R-j}\right] \tag{4.252}$$

Division by the Vandermonde determinant $\det\left[x_i^{n_R-j}\right]$ and substitution of (4.240) gives the final result for the character expansion

$$\begin{aligned}\exp[tr(\mathbf{X})] &= \sum_{k_1\geq k_2\geq\cdots\geq k_{nR}\geq 0}\det\left[\frac{1}{(k_j+i-j)!}\right]\frac{\det\left[x_i^{k_j+n_R-j}\right]}{\det\left[x_i^{n_R-j}\right]} \\ &= \sum_\kappa \det\left[\frac{1}{(k_j+i-j)!}\right]\chi_{\mathcal{V}_\kappa}(\mathbf{X})\end{aligned} \tag{4.253}$$

where the summation is over all weakly decreasing partitions $\kappa = (k_1, k_2, ..., k_{n_R})$, $k_1 \geq k_2 \geq \cdots \geq k_{n_R} \geq 0$, into no more than n_R parts. We now reach the conclusion that the irreducible representations in the space $\mathcal{V}$ generated by the function $f(\mathbf{X}) = \exp[tr(\mathbf{X})]$ are labelled by the partitions κ of all integers into no more than n_R parts. Naturally, the coefficients in the expansion are given by

$$c_\kappa = \det\left[\frac{1}{(k_j + i - j)!}\right] = \det\left[\frac{\prod_{p=i+1}^{n_R}(k_j + p - j)}{(k_j + n_R - j)!}\right] = \left[\prod_{p=1}^{n_R}\frac{1}{(k_p + n_R - p)!}\right]\det\left[(k_j + n_R - j)^{n_R - i}\right] \quad (4.254)$$

Using the character expansion in (4.253) it is now easy to compute the integral within the eigenvalue pdf of (4.199). Write

$$\begin{aligned} f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) &= G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})\int_{\widetilde{U}(n_R)}\exp\left[-tr\left(\mathbf{\Upsilon}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H\right)\right](d\mathbf{U})^\wedge \\ &= G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})\sum_\kappa c_\kappa\int_{\widetilde{U}(n_R)}\chi_{\mathcal{V}_\kappa}(-\mathbf{\Upsilon}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H)(d\mathbf{U})^\wedge \end{aligned} \quad (4.255)$$

and the integration of the exponential reduces to the integration of the irreducible character. Since the integration is over the unitary group, and the latter is a compact topological group, the general result of (4.231) applies directly, i.e.

$$\int_{\widetilde{U}(n_R)}\chi_{\mathcal{V}_\kappa}(-\mathbf{\Upsilon}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H)(d\mathbf{U})^\wedge = \text{Volume}\left[\widetilde{U}(n_R)\right]\frac{\chi_{\mathcal{V}_\kappa}(-\mathbf{\Upsilon}^{-1})\chi_{\mathcal{V}_\kappa}(\mathbf{\Lambda})}{\chi_{\mathcal{V}_\kappa}(\mathbf{I}_{n_R})} \quad (4.256)$$

and since the volume of the unitary group is already known from (4.134), we have

$$\begin{aligned} f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) &= \frac{2^{n_R}\pi^{n_R^2}}{\widetilde{\Gamma}_{n_R}(n_R)}G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})\sum_\kappa c_\kappa\frac{\chi_{\mathcal{V}_\kappa}(-\mathbf{\Upsilon}^{-1})\chi_{\mathcal{V}_\kappa}(\mathbf{\Lambda})}{\chi_{\mathcal{V}_\kappa}(\mathbf{I}_{n_R})} \\ &= \frac{2^{n_R}\pi^{n_R^2}}{\widetilde{\Gamma}_{n_R}(n_R)}\frac{G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]\det\left[\lambda_i^{n_R-j}\right]}\sum_\kappa c_\kappa\frac{\det\left[(-\upsilon_i^{-1})^{k_j+n_R-j}\right]\det\left[\lambda_i^{k_j+n_R-j}\right]}{\chi_{\mathcal{V}_\kappa}(\mathbf{I}_{n_R})} \end{aligned} \quad (4.257)$$

where the υ_i's are the eigenvalues of $\mathbf{\Upsilon}$ (and $\mathbf{\Sigma}$) and the λ_i's are the eigenvalues of $\mathbf{\Lambda}$. Notice that character of the identity matrix could not be obtained directly because of its indeterminate form

$$\chi_{\mathcal{V}_\kappa}(\mathbf{I}_{n_R}) = \lim_{\mathbf{X}\to\mathbf{I}_{n_R}}\chi_{\mathcal{V}_\kappa}(\mathbf{X}) = \lim_{x_i\to 1}\frac{\det\left[x_i^{k_j+n_R-j}\right]}{\det\left[x_i^{n_R-j}\right]} = \frac{0}{0} \quad (4.258)$$

It is however possible to compute the limit with the help of L'Hôpital's rule

$$\lim_{\mathbf{X}\to\mathbf{I}_{n_R}}\chi_{\mathcal{V}_\kappa}(\mathbf{X}) = \lim_{x_i\to 1}\frac{\frac{\partial^{r_1}}{\partial x_1}\cdots\frac{\partial^{r_{n_R}}}{\partial x_{n_R}}\left(\det\left[x_i^{k_j+n_R-j}\right]\right)}{\frac{\partial^{r_1}}{\partial x_1}\cdots\frac{\partial^{r_{n_R}}}{\partial x_{n_R}}\left(\det\left[x_i^{n_R-j}\right]\right)} \quad (4.259)$$

and the rule of determinant differentiation

$$\frac{\partial\det\mathbf{A}}{\partial x} = \det\mathbf{D}_1 + \det\mathbf{D}_2 + \cdots + \det\mathbf{D}_n \quad (4.260)$$

where $\mathbf{D}_i$ is identical to $\mathbf{A}$ except that the entries in the i-th row are replaced by derivatives with respect to x. Since the Vandermonde matrix has a different variable in each row, differentiating the Vandermonde determinant is equivalent to differentiating the respective row of the Vandermonde matrix, i.e.

$$\frac{\partial^{r_k}}{\partial x_k}\Big(\det\big[\,x_i^{t_j}\,\big]\Big) = \det\left[\prod_{p=0}^{r_k-1}(t_j-p)^{\delta_{ik}}\,x_i^{t_j-r_k\delta_{ik}}\right] \tag{4.261}$$

Now, since the x_i's are all unity in the limit of (4.259), the only possibility for having a non-zero determinant in both numerator and denominator is to differentiate a different number of times with respect to each x_i, say, differentiate $n_R - i$ times with respect to x_i, i.e. $r_k = n_R - k$. Thus

$$\frac{\partial^{n_R-1}}{\partial x_1}\cdots\frac{\partial^{0}}{\partial x_{n_R}}\Big(\det\big[\,x_i^{t_j}\,\big]\Big) = \det\left[\prod_{p=0}^{n_R-i-1}(t_j-p)\,x_i^{t_j-(n_R-i)}\right] \tag{4.262}$$

We are now in conditions to compute the two limits in (4.259). We have

$$\begin{aligned}\lim_{x_i\to 1}\frac{\partial^{n_R-1}}{\partial x_1}\cdots\frac{\partial^{0}}{\partial x_{n_R}}\Big(\det\big[\,x_i^{n_R-j}\,\big]\Big) &= \det\left[\prod_{p=0}^{n_R-i-1}(n_R-j-p)\right]\\ &= \det\big[(n_R-j)^{n_R-i}\big] = \prod_{p=1}^{n_R}(n_R-p)!\end{aligned} \tag{4.263}$$

and

$$\begin{aligned}\lim_{x_i\to 1}\frac{\partial^{n_R-1}}{\partial x_1}\cdots\frac{\partial^{0}}{\partial x_{n_R}}\Big(\det\big[\,x_i^{k_j+n_R-j}\,\big]\Big) &= \det\left[\prod_{p=0}^{n_R-i-1}(k_j+n_R-j-p)\right]\\ &= \det\big[(k_j+n_R-j)^{n_R-i}\big]\end{aligned} \tag{4.264}$$

Substitution of (4.263) and (4.264) into (4.259) gives the character of the identity matrix and dimension of $\mathcal{V}_\kappa$ as

$$d_\kappa = \chi_{\mathcal{V}_\kappa}(\mathbf{I}_{n_R}) = \lim_{\mathbf{X}\to\mathbf{I}_{n_R}}\chi_{\mathcal{V}_\kappa}(\mathbf{X}) = \frac{\det\big[(k_j+n_R-j)^{n_R-i}\big]}{\det\big[(n_R-j)^{n_R-i}\big]} \tag{4.265}$$

Comparing this expression with the coefficients expression in (4.254), the latter can be rewritten as follows

$$\begin{aligned}c_\kappa &= \left[\prod_{p=1}^{n_R}\frac{1}{(k_p+n_R-p)!}\right]\det\big[(n_R-j)^{n_R-i}\big]\,\chi_{\mathcal{V}_\kappa}(\mathbf{I}_{n_R})\\ &= \left[\prod_{p=1}^{n_R}\frac{1}{(k_p+n_R-p)!}\right]\left[\prod_{p=1}^{n_R-1}p!\right]\chi_{\mathcal{V}_\kappa}(\mathbf{I}_{n_R})\end{aligned} \tag{4.266}$$

and therefore, replacing into (4.257) yields

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = \frac{\dfrac{2^{n_R}\pi^{n_R^2}}{\widetilde{\Gamma}_{n_R}(n_R)}G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})}{\det\big[(-\upsilon_i^{-1})^{n_R-j}\big]\det\big[\lambda_i^{n_R-j}\big]}\left[\prod_{p=1}^{n_R-1}p!\right]\sum_{\kappa}\left[\prod_{p=1}^{n_R}\frac{1}{(k_p+n_R-p)!}\right]\det\big[(-\upsilon_i^{-1})^{k_j+n_R-j}\big]\det\big[\lambda_i^{k_j+n_R-j}\big] \tag{4.267}$$

Recall that the inner summation is given by

$$\sum_{\kappa} \sim \sum_{k_1 \geq k_2 \geq \cdots \geq k_{n_R} \geq 0} \tag{4.268}$$

and that by making the change of variables of (4.248) one obtains a strictly ordered indexation of the sum, or more formally

$$\begin{aligned} k_j + n_R - j \quad &\rightarrow \quad t_j \\ \sum_{k_1 \geq k_2 \geq \cdots \geq k_{n_R} \geq 0} \quad &\mapsto \quad \sum_{t_1 > t_2 > \cdots > t_{n_R} \geq 0} \end{aligned} \tag{4.269}$$

As such, the pdf expression in (4.247) can be written more simply as

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = \frac{\dfrac{2^{n_R}\pi^{n_R^2}}{\widetilde{\Gamma}_{n_R}(n_R)} G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]\det\left[\lambda_i^{n_R-j}\right]}\left[\prod_{p=1}^{n_R-1} p!\right]\sum_{t_1>t_2>\cdots>t_{n_R}\geq 0}\left[\prod_{p=1}^{n_R}\frac{1}{t_p!}\right]\det\left[(-\upsilon_i^{-1})^{t_j}\right]\det\left[\lambda_i^{t_j}\right] \tag{4.270}$$

In order to further simplify the sum we need to expand the determinants inside the summation using (4.245). Expanding the one on the left yields the sum

$$\sum_{t_1>t_2>\cdots>t_{n_R}\geq 0}\left[\sum_q \epsilon(q)\prod_{i=1}^{n_R}(-\upsilon_i^{-1})^{t_{q_i}}\right]\left[\prod_{p=1}^{n_R}\frac{1}{t_p!}\right]\det\left[\lambda_i^{t_j}\right] \tag{4.271}$$

Here, we see that for each sequence $t_1 > t_2 > \cdots > t_{n_R} \geq 0$, the summand is a complete set of permutations, and hence it may be included in the outer sum by disregarding the order of the t_i's. Since for each permutation of the t_i's, the sign of $\det\left[\lambda_i^{t_j}\right]$ changes, it obviously cancels the sign of the inner permutation $\epsilon(q)$, and thus the sum simplifies to

$$\sum_{t_1=0}^{\infty}\cdots\sum_{t_{n_R}=0}^{\infty}\left[\prod_{p=1}^{n_R}\frac{1}{t_p!}(-\upsilon_p^{-1})^{t_p}\right]\det\left[\lambda_i^{t_j}\right] \tag{4.272}$$

The final step involves expanding the remaining determinant and rearranging terms, which gives

$$\begin{aligned} &\sum_{t_1=0}^{\infty}\cdots\sum_{t_{n_R}=0}^{\infty}\left[\prod_{i=1}^{n_R}\frac{1}{t_i!}(-\upsilon_i^{-1})^{t_i}\right]\left[\sum_q \epsilon(q)\prod_{i=1}^{n_R}\lambda_{q_i}^{t_i}\right] = \\ &= \sum_q \epsilon(q)\prod_{i=1}^{n_R}\left[\sum_{t_i=0}^{\infty}\frac{1}{t_i!}\left(-\lambda_{q_i}/\upsilon_i\right)^{t_i}\right] = \sum_q \epsilon(q)\prod_{i=1}^{n_R} e^{-\lambda_{q_i}/\upsilon_i} = \det\left[e^{-\lambda_i/\upsilon_j}\right] \end{aligned} \tag{4.273}$$

Surprisingly, one realizes that the expression for the joint pdf of the eigenvalues of the Wishart matrix is as simple as

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = \frac{2^{n_R}\pi^{n_R^2}}{\widetilde{\Gamma}_{n_R}(n_R)} G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})\left[\prod_{p=1}^{n_R-1} p!\right]\frac{\det\left[e^{-\lambda_i/\upsilon_j}\right]}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]\det\left[\lambda_i^{n_R-j}\right]} \tag{4.274}$$

Notice that this expression only depends upon three matrix determinants, which are easily computed in contrast with the infinite series of (4.206). The interrogation that now emerges is that of if the equation is sufficiently amenable to mathematical manipulation, since it must be integrated with respect to the eigenvalues in $\mathbf{\Lambda}$. Expanding $G_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Upsilon})$ in (4.198)

and observing that

$$\det\left[\lambda_i^{n_R-j}\right] = \prod_{i<j}^{n_R}(\lambda_i - \lambda_j) \tag{4.275}$$

the joint density can be reexpressed as

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = \frac{1}{n_R!}H_{n_R,n_T}(\mathbf{\Upsilon})\det\left[e^{-\lambda_i/\upsilon_j}\right]\det\left[\lambda_i^{n_R-j}\right]\prod_{i=1}^{n_R}\lambda_i^{n_T-n_R} \tag{4.276}$$

where

$$H_{n_R,n_T}(\mathbf{\Upsilon}) = \frac{\pi^{n_R(n_R-1)}}{\widetilde{\Gamma}_{n_R}(n_R)\widetilde{\Gamma}_{n_R}(n_T)}\left[\prod_{p=1}^{n_R-1}p!\right]\frac{1}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]\prod_{i=1}^{n_R}\upsilon_i^{n_T}} \tag{4.277}$$

is a function that only depends on the MIMO channel dimensions and on the eigenvalues υ_i of the covariance matrix $\mathbf{\Sigma}$. It is clear from (4.276) that, in the correlated MIMO scenario, the joint pdf of the eigenvalues is not a symmetric function of the eigenvalues. Hence, one requires (4.54) to compute the average capacity, i.e.

$$C_{avg} = \sum_{k=1}^{n_R}E\left[\log_2\left(1+\frac{\gamma}{n_T}\lambda_k\right)\right] = \sum_{k=1}^{n_R}\int_{\mathbf{\Lambda}}\log_2\left(1+\frac{\gamma}{n_T}\lambda_k\right)f_{\mathbf{\Lambda}}(\mathbf{\Lambda})(d\mathbf{\Lambda})^{\wedge} \tag{4.278}$$

Expanding the pdf yields

$$C_{avg} = \frac{1}{n_R!}H_{n_R,n_T}(\mathbf{\Upsilon})\sum_{k=1}^{n_R}\int_{\lambda_1,\ldots,\lambda_{n_R}}\det\left[\lambda_i^{n_R-j}\right]\det\left[\exp(-\lambda_i/\upsilon_j)\right]\log_2\left(1+\frac{\gamma}{n_T}\lambda_k\right)\prod_{i=1}^{n_R}\lambda_i^{n_T-n_R}d\lambda_1\ldots d\lambda_{n_R} \tag{4.279}$$

Expanding the determinants yet again leads to

$$\begin{aligned} C_{avg} &= \frac{1}{n_R!}H_{n_R,n_T}(\mathbf{\Upsilon})\sum_{k=1}^{n_R}\int_{\lambda_1,\ldots,\lambda_{n_R}}\left[\sum_p\epsilon(p)\sum_q\epsilon(q)\prod_{i=1}^{n_R}\lambda_i^{n_R-p_i}e^{-\lambda_i/\upsilon_{q_i}}\right]\log_2\left(1+\frac{\gamma}{n_T}\lambda_k\right)\prod_{i=1}^{n_R}\lambda_i^{n_T-n_R}d\lambda_1\ldots d\lambda_{n_R} \\ &= \frac{1}{n_R!}H_{n_R,n_T}(\mathbf{\Upsilon})\sum_{k=1}^{n_R}\sum_p\epsilon(p)\sum_q\epsilon(q)\prod_{i=1}^{n_R}\int_0^{\infty}\lambda^{n_T-p_i}e^{-\lambda/\upsilon_{q_i}}\left[\log_2\left(1+\frac{\gamma}{n_T}\lambda\right)\right]^{\delta_{kq_i}}d\lambda \end{aligned} \tag{4.280}$$

where δ_{kq_i} is the kronecker delta. Remembering Equation (4.245), one acknowledges that the inner sum is in fact an expanded determinant where q_i may index the matrix rows, and accordingly, a more compact version of Equation (4.280) is

$$C_{avg} = \frac{1}{n_R!}H_{n_R,n_T}(\mathbf{\Upsilon})\sum_{k=1}^{n_R}\sum_p\epsilon(p)\det\left[\int_0^{\infty}\lambda^{n_T-p_i}e^{-\lambda/\upsilon_i}\left[\log_2\left(1+\frac{\gamma}{n_T}\lambda\right)\right]^{\delta_{ki}}d\lambda\right] \tag{4.281}$$

Moreover, each permutation p induces a column interchange in the matrix of the inner determinant, which brings in a factor of +1 or -1 and cancels $\epsilon(p)$. Therefore, since there are $n_R!$ permutations, we may write the average capacity as

$$C_{avg} = H_{n_R,n_T}(\mathbf{\Upsilon})\sum_{k=1}^{n_R}\det\left[\int_0^{\infty}\lambda^{n_T-j}e^{-\lambda/\upsilon_i}\left[\log_2\left(1+\frac{\gamma}{n_T}\lambda\right)\right]^{\delta_{ki}}d\lambda\right] \tag{4.282}$$

For efficient evaluation of this expression, a closed form solution of the integral is clearly

required. First we make the change of variable $\lambda \rightarrow \upsilon_i \lambda$ and arrive at

$$C_{avg} = H_{n_R,n_T}(\boldsymbol{\Upsilon}) \sum_{k=1}^{n_R} \det\left[\upsilon_i^{n_T-j+1} \int_0^\infty \lambda^{n_T-j} e^{-\lambda} \left[\log_2\left(1 + \frac{\upsilon_i \gamma}{n_T}\lambda\right)\right]^{\delta_{ki}} d\lambda \right] \tag{4.283}$$

Now, one knows from (4.187) that

$$\int_0^\infty \lambda^r e^{-\lambda} \log_2\left(1 + \frac{\gamma}{n_T}\lambda\right) d\lambda = r! \frac{e^{n_T/\gamma}}{\ln 2} \sum_{p=0}^{r} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{n_T}{\gamma}\right)^{p-l} \Gamma(l, n_T/\gamma) \tag{4.284}$$

and for the case $\delta_{ki} = 0$, the basic identity

$$\int_0^\infty \lambda^r e^{-\lambda} d\lambda = \Gamma(r+1) = r! \tag{4.285}$$

may be used. Hence, we write with no further delay that the average capacity of the MIMO channel in a Rayleigh fading environment with branch correlation solely at the receiver is given by

$$C_{avg} = H_{n_R,n_T}(\boldsymbol{\Upsilon}) \sum_{k=1}^{n_R} \det\left[\frac{(n_T-j)!}{\upsilon_i^{-(n_T-j+1)}} \left[\frac{e^{n_T/(\upsilon_i\gamma)}}{\ln 2} \sum_{p=0}^{n_T-j} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{n_T}{\upsilon_i\gamma}\right)^{p-l} \Gamma\left[l, \frac{n_T}{\upsilon_i\gamma}\right] \right]^{\delta_{ki}} \right] \tag{4.286}$$

where $\Gamma(a,b)$ is the incomplete gamma function defined in (4.188), $H_{n_R,n_T}(\boldsymbol{\Upsilon})$ is given in (4.277), and the remaining variables are

$$\begin{aligned}
\gamma &\rightarrow \text{average SNR at each receiver branch} \\
n_T &\rightarrow n\text{umber of transmitter antennas} \\
n_R &\rightarrow n\text{umber of receiver antennas} \\
\upsilon_i &\rightarrow \text{eigenvalues of the } n_R \times n_R \text{ covariance matrix } \boldsymbol{\Sigma} \\
\delta_{ki} &\rightarrow \text{kronecker delta}
\end{aligned}$$

- **Two-sided Independent Fading scenario**

The average capacity under independent fading, i.e. when $\boldsymbol{\Sigma} = \mathbf{I}_{n_R}$, has been computed in [71] in terms of the integral of Laguerre polynomials, and a closed form expression has been derived in (4.189). With the help of the mathematical concepts introduced so far it is possible to devise an alternative closed-form expression for the capacity in this particular scenario. Notice that this result cannot be obtained directly from (4.286). Instead, one has to replace $\chi_{\mathcal{V}_\kappa}(-\boldsymbol{\Sigma}^{-1})$ by $\chi_{\mathcal{V}_\kappa}(-\mathbf{I}_{n_R})$ in (4.257) and then proceed by the same lines as (4.258)-(4.276). We find that

$$\chi_{\mathcal{V}_\kappa}(-\mathbf{I}_{n_R}) = \frac{\det\left[(k_j + n_R - j)^{n_R-i}(-1)^{k_j+i-j}\right]}{\det\left[(n_R-j)^{n_R-i}(-1)^{i-j}\right]} \tag{4.287}$$

and after some manipulations that

$$f_{\boldsymbol{\Lambda}}^{\boldsymbol{\Sigma}=\mathbf{I}}(\boldsymbol{\Lambda}) = \frac{1}{n_R!} H_{n_R,n_T}(\mathbf{1}) \det\left[(-1)^{n_R-j} \sum_{t=0}^{\infty} \frac{(-1)^t t^{n_R-j}}{t!} \lambda_i^t \right] \det\left[\lambda_i^{n_T-j}\right] \tag{4.288}$$

where

$$H_{n_R,n_T}(\mathbf{1}) = \frac{\pi^{n_R(n_R-1)}}{\tilde{\Gamma}_{n_R}(n_R)\tilde{\Gamma}_{n_R}(n_T)}\left[\prod_{p=1}^{n_R-1} p!\right]\frac{1}{\det\left[(n_R-j)^{n_R-i}(-1)^{i-j}\right]} \tag{4.289}$$

The eigenvalue pdf can be simplified by noting the equality

$$\sum_{k=0}^{\infty}\frac{(-1)^k k^n}{k!}x^k = e^{-x}\sum_{k=1}^{n}(-1)^k a_{k,n}x^k \tag{4.290}$$

where

$$a_{k,n} = \sum_{0<r_1\le\cdots\le r_{n-k}}^{k} r_1 r_2 \ldots r_{n-k} \tag{4.291}$$

which leads to

$$f_{\mathbf{\Lambda}}^{\Sigma=\mathbf{I}}(\mathbf{\Lambda}) = \frac{1}{n_R!}H_{n_R,n_T}(\mathbf{1})\det\left[e^{-\lambda_i}\sum_{t=1}^{n_R-j}(-1)^{t+n_R-j}a_{t,n_R-j}\lambda_i^t\right]\det\left[\lambda_i^{n_T-j}\right] \tag{4.292}$$

Now, completing the same steps as (4.278)-(4.286), the average MIMO capacity under independent fading is found to be

$$\begin{aligned} C_{avg}^{\Sigma=\mathbf{I}} = {} & H_{n_R,n_T}(\mathbf{1})\sum_{k=1}^{n_R}\det\left[\sum_{t=1}^{n_R-i}(-1)^{t+n_R-i}a_{t,n_R-i}(t+n_T-j)!\cdot\right. \\ & \left.\cdot\left[\frac{e^{n_T/\gamma}}{\ln 2}\sum_{p=0}^{t+n_T-j}\frac{1}{p!}\sum_{l=0}^{p}\binom{p}{l}\left(-\frac{n_T}{\gamma}\right)^{p-l}\Gamma\left[l,\frac{n_T}{\gamma}\right]\right]^{\delta_{ki}}\right] \end{aligned} \tag{4.293}$$

This equation is only valid for $n_T \ge n_R$. The coefficients $a_{k,n}$ can be computed using a simple algorithm presented in Appendix A.

4.5.8 SPATIAL CORRELATION MODELLING

Before materializing the capacity expression into a readable plot, it is necessary to adopt a model for the covariance matrix $\mathbf{\Sigma}$. Several correlation models have been proposed in the literature. One of them is the exponential correlation model, which was introduced in [74] and has been extensively used by researchers in the performance analysis of space diversity techniques. In this section we shall follow an alternative approach.

Recall that if $\mathbf{h}_j$ is a column from the MIMO channel response matrix, then $\mathbf{\Sigma}$ is given by

$$\mathbf{\Sigma} = E\left[\mathbf{h}_j\mathbf{h}_j^H\right] = \begin{pmatrix} E\left[h_{1j}h_{1j}^*\right] & \cdots & E\left[h_{1j}h_{n_Rj}^*\right] \\ \vdots & \ddots & \vdots \\ E\left[h_{n_Rj}h_{1j}^*\right] & \cdots & E\left[h_{n_Rj}h_{n_Rj}^*\right] \end{pmatrix} \tag{4.294}$$

which means that we only need to determine the general entry $E\left[h_{ij}^*h_{rj}\right]$. To this end, Equation (4.67) can be used directly, i.e.

$$E\left[h_{ij}(t)h_{rj}^*(t)\right] = E\left[\cos\left(\beta d_{ir}\cos(\phi_{jk}+\theta_{ir})\right)\right] - jE\left[\sin\left(\beta d_{ir}\cos(\phi_{jk}+\theta_{ir})\right)\right] \tag{4.295}$$

where β is the phase constant, ϕ_{jk} is the angle between the direction of motion and the k-th incoming wave, θ_{ir} is the angle between the direction of motion and the line joining the i-th and r-th receiver antennas, and d_{ir} is the distance between the antennas. To complete the expectation process, one may reasonably assume a rich scattering environment where the multipath waves arrive from every direction around the receiver. This is a good approximation for urban environments cluttered with objects of different dimensions, such as buildings, cars, and trees, which scatter the waves in all directions (isotropic scattering). Moreover, there is no line-of-sight in these scenarios, justifying the Rayleigh model. The angle spread is exactly 2π radians and the density function of the ϕ_{jk}'s is uniform from 0 to 2π, i.e.

$$p(\phi_{ij}) = \frac{1}{2\pi} \qquad 0 \leq \phi_{ij} \leq 2\pi \tag{4.296}$$

When the scattering is non-isotropic (e.g. because of the "street-canyon" effect), the width of the angle-of-arrival (AOA) is not 2π and (4.296) may become a rough approximation. In [75], a generalization of (4.296) has been proposed, and is based on the *von Mises* density function

$$p(\phi_{jk}) = \frac{1}{2\pi I_0(\kappa)} \mathrm{e}^{\kappa \cos(\phi_{jk} - \phi_c)} \qquad -\pi \leq \phi_{jk} \leq \pi \tag{4.297}$$

where ϕ_c is the mean AOA, $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind, and $\kappa \in [0, \infty[$ controls the width of the AOA.

We shall consider solely the case $\kappa = 0$, which corresponds to the isotropic scattering model (ISM). It follows that the in-phase correlation is

$$\begin{aligned} E\big[\cos\big(\beta d_{ir} \cos(\phi_{jk} + \theta_{ir})\big)\big] &= \frac{1}{2\pi} \int_0^{2\pi} \cos\big(\beta d_{ir} \cos(\phi_{jk} + \theta_{ir})\big) d\phi_{jk} \\ &= \frac{1}{2\pi} \int_{\theta_{ir}}^{2\pi + \theta_{ir}} \cos\big(\beta d_{ir} \cos(\phi_{jk})\big) d\phi_{jk} = J_0(\beta d_{ir}) \end{aligned} \tag{4.298}$$

and that the quadrature correlation is

$$E\big[\sin\big(\beta d_{ir} \cos(\phi_{jk} + \theta_{ir})\big)\big] = \frac{1}{2\pi} \int_0^{2\pi} \sin\big(\beta d_{ir} \cos(\phi_{jk} + \theta_{ir})\big) d\phi_{jk} = 0 \tag{4.299}$$

and thus the entries of the covariance matrix are given by

$$[\mathbf{\Sigma}]_{ir} = E\big[h_{ij}^*(t) h_{rj}(t)\big] = J_0(\beta d_{ir}) \tag{4.300}$$

where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind. Although Equation (4.296) also depends on θ_{ir}, the angle between the direction of motion and the lines joining each pair of antennas, taking the expectation in (4.295) makes it clear that the correlation between the channel responses at two different antennas is independent of θ_{ir}. This is a distinctive feature of the uniform distribution of wave arrival.

From the plot of (4.300) in Figure 4.8, it is seen that it is possible to achieve low correlation by placing the antennas near the zeros of the Bessel function. The zeros are approximately equispaced by 0.5λ, and the first zero occurs at 0.38λ, which means that for 1-D or 2-D arrays with more than 2 or 3 receiver antennas, respectively, it is not possible to attain zero correlation between all the antennas. One should, however, expect an im-

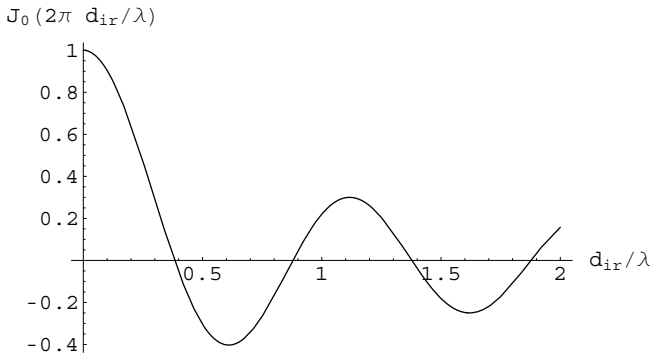

FIGURE 4.8 Receiver spatial correlation for the isotropic scattering model.

provement in capacity by arranging the antennas so that the average correlation between them is minimized. In practice, and depending on the particular mobile scenario, the spatial correlation function between any pair of antennas will diverge from the ISM. Nevertheless, if the receiver is able to accurately track the channel matrix, it may also be able to approximate (4.300) by time-averaging the product $h_{ij}(t)h_{rj}^*(t)$ during some multiple of the channel coherence time. Now, and for instance, if the receiver is equipped with an array with many closely-spaced antennas, it may select (for signal detection) the ones that achieve the lowest average correlation be-tween all pairs of antennas, and thus, in principle, improve the capacity.

4.5.9 NUMERICAL RESULTS AND DISCUSSION

Due to receiver size constraints, the antenna arrays of typical mobile units must be restricted in their end-to-end dimensions. Therefore, it seems reasonable to confine the analysis to receiver arrays of fixed maximum length, say $d_{\max} = 1.2\lambda$, assume a minimum separation between antennas of $d_{\min} = 0.15\lambda$, and a maximum of eight antennas. In the case of linear equispaced arrays, there are two options to increase the number of antennas from one to eight: 1) comply with the maximum size, and evenly distribute the antennas within; and 2) comply with the minimum separation, thereby reducing the maximum length. In what follows, we will use the first option as a case study, from which more general conclusions may be drawn. The entries of the covariance matrix are thus given by

$$[\mathbf{\Sigma}]_{ir} = J_0(\beta |i - r| d_{\max} / n_R) \tag{4.301}$$

with which $\mathbf{\Sigma}$ can be built and its eigenvalues be found. The capacity in the special cases of $n_R = 2$ and $n_R = 4$ is plotted in Figure 4.9, where the curves for the Gaussian and Rayleigh SISO channels are also included for comparison. It is clear that the capacity in the correlated case is always lower than in the independent case, and this reduction becomes more significant when the number of receiver antennas increases. For instance, at $\gamma = 10$ dB, and for $n_R = 2$ and $n_T = 8$, the capacity decreases by approximately 0.20 bits/s/Hz relative to the independent case, while for $n_R = 4$ the loss is more than 1 bits/s/Hz. It is noticeable, however, that the utmost increase in capacity is obtained by increasing n_R, and increasing n_T only provides a slight capacity gain. It is also seen that the higher is n_R, the better the capacity scales with increasing SNR, which means that the MIMO system takes superior advantage of receive diversity. Moreover, and despite the presence of cor-

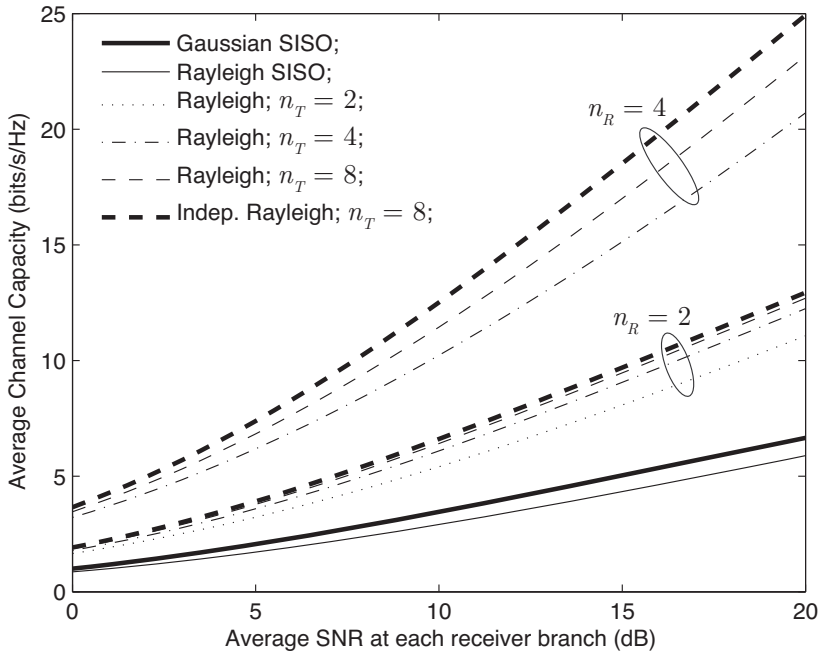

FIGURE 4.9 Average capacity of the MIMO channel as a function of the branch SNR at the receiver side, in the special cases of two and four receiver antennas.

relation, it is evident that the use of multiple antenna receivers/transmitters offers remarkably improved capacities as compared to SISO configurations. Even with a scenario as simple as $\gamma = 10$ dB, $n_R = 2$ and $n_T = 2$, and a Rayleigh fading environment, we are granted a capacity gain of 2.5 bits/s/Hz which almost doubles the capacity of the Rayleigh SISO channel, and for $n_R = 4$ we watch an increase of 251%. Nonetheless, recall that the length of the receiver array is being restricted to $d_{\max} = 1.2\lambda$, and that larger capacities are certainly possible by increasing this length.

Figure 4.10 illustrates the case $n_R = 8$ and in addition incorporates the limiting situation of $n_T \to \infty$, as given by (4.40). One concludes that it is not possible to achieve the capacity under independent fading even with an infinite number of transmitter antennas. In fact, doubling n_T offers a negligible increase in capacity. Even so, we find that at $\gamma = 10$ dB and $n_R = n_T = 8$, the capacity in the correlated case is 521% superior to the Rayleigh SISO capacity. These remarkable capacities are obtained without ever increasing the length of the array, but by fitting all the antennas within $d_{\max} = 1.2\lambda$, which is a very important consideration if the receiver unit is to be portable and, hence, small-sized.

A question that now arises is: What is the best fit of the antennas within $d_{\max}$ so that the capacity is maximized? So far we have been using equispaced linear arrays, and have also been assuming that the receiver has the same number of detectors as the number of antennas. As previously mentioned, if the receiver has several equispaced antennas, it may select the ones that achieve the highest capacity. For example, imagine a receiver with a linear array of eight equispaced antennas, but that only has four signal detectors available. Then it must choose the four antennas that will maximize capacity. Maximizing (4.286) directly over $\mathbf{\Sigma}$ does not seem an easy problem. We know, however, that (4.35) is maximized when $\mathbf{\Sigma} = \mathbf{I}_{n_R}$ (see [76]), which is consistent with the fact that the independent fading scenarios yield the highest capacities. Intuitively, one may now expect to achieve higher ca-

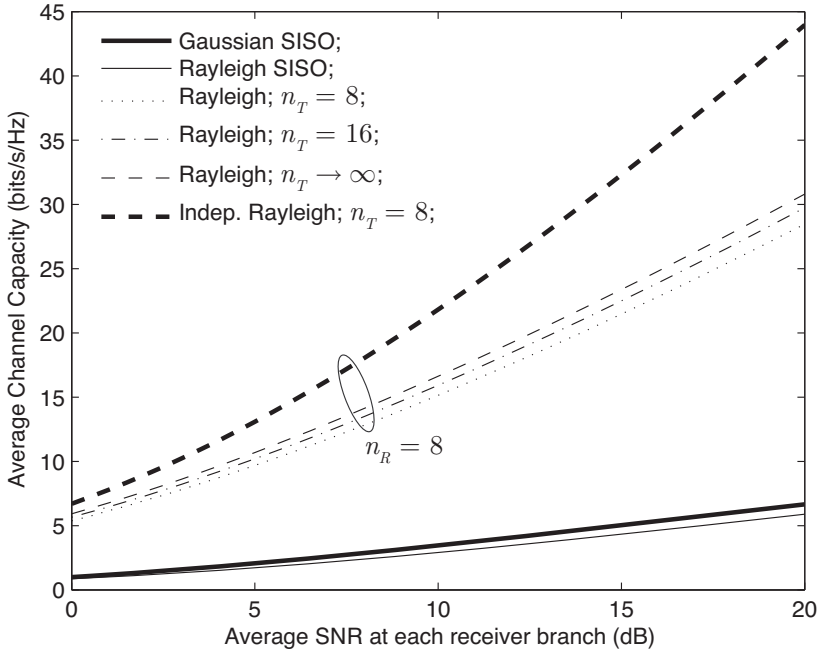

FIGURE 4.10 Average capacity of the MIMO channel as a function of the branch SNR at the receiver side, in the special case of eight receiver antennas.

pacities by reducing the average correlation between all pairs of receiver antennas or, similarly, by minimizing the sum of the absolute values of the entries of $\mathbf{\Sigma}$, i.e.

$$\max_{\mathbf{\Sigma}} C_{avg} \leftrightarrow \min_{\mathbf{\Sigma}} \sum_{i,r} |[\mathbf{\Sigma}]_{ir}| \tag{4.302}$$

Notice that (4.302) is equivalent to minimizing the Frobenius norm of $\mathbf{\Sigma}$. Since the receiver has to choose four of eight available antennas, it must find all 4×4 principal submatrices of $\mathbf{\Sigma}$, that is, all submatrices that lie on the same set of rows and columns, and then apply (4.302) to each of these. It then chooses the principal submatrix with smaller Frobenius norm, and finally selects the antennas corresponding to the original rows (and columns) of $\mathbf{\Sigma}$. A simple algorithm for finding the antenna positions within $d_{\max}$ that maximize capacity is given in Appendix B. For the example in consideration, the antenna positions are the subset {1, 3, 6, 8}. When the receiver has 16 equispaced antennas and 4 detectors, the positions are {1, 6, 11, 16} and, in the case of 32 equispaced antennas, {1, 11, 22, 32}. Figure 4.11 shows the capacity improvement in the ISM model by following this procedure. Since an increase in the number of receiver antennas is accompanied by an enhanced spatial resolution, we perceive a substantial increase in available capacity. By employing 32 receiver antennas and effectively using four of them, it is almost possible to attain the capacity under independent fading. This a very interesting result, yet it is not generalizable to situations where the receiver has a greater number of detectors. For instance, if the receiver has eight detectors, then the optimum antenna positions are {1, 2, 6, 7, 11, 12, 13, 16} and {1, 2, 11, 12, 21, 22, 31, 32} for the cases of 16 and 32 receiver antennas, respectively. From the plot in Figure 4.12 it is seen that the capacity improvement is subtle, and approximating the capacity under independent fading is not feasible. This happens because with eight detectors, the spatial resolution is already high, and any antennas added to the configuration will be severely correlated with the ones already

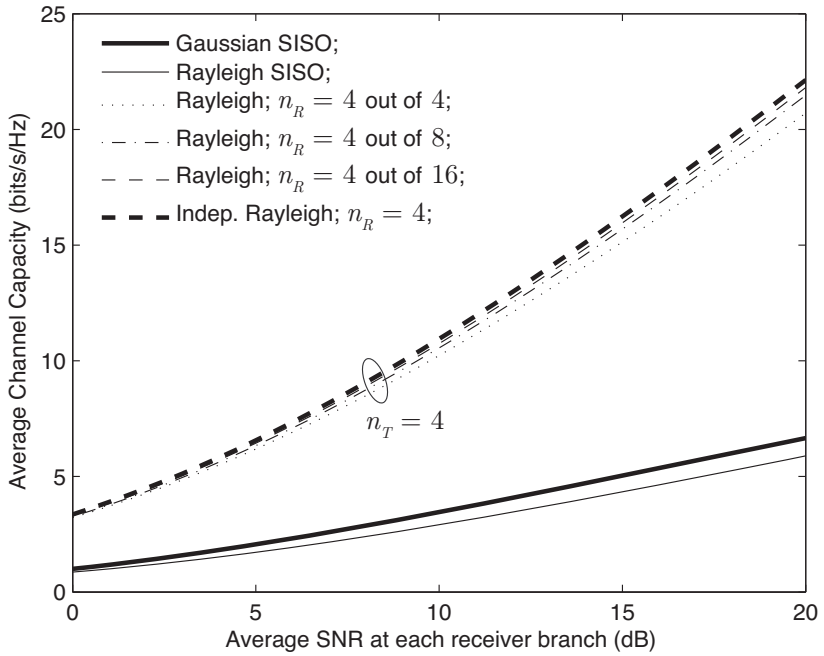

FIGURE 4.11 Optimizing the capacity of the MIMO channel by selecting four receiver antennas from an equispaced linear array with eight and sixteen antennas.

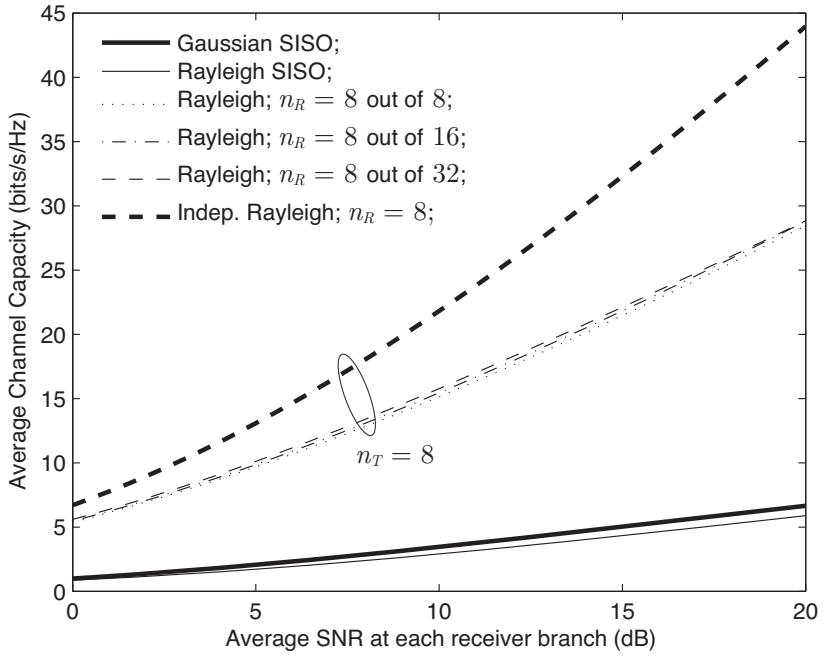

FIGURE 4.12 Optimizing the capacity of the MIMO channel by selecting eight receiver antennas from an equispaced linear array with sixteen and thirty two antennas.

present. Nevertheless, this exposition shows that equispaced linear arrays do not offer the maximum available capacity, and that by using some simple signal processing and calculation techniques, any MIMO system can optimize the data rates delivered to users, and the spectral efficiency of the whole communication system.

4.5.10 CONCLUSION

Section 4.5 has presented a clear and complete derivation of the closed-form expression for the average capacity of MIMO channels with $n_T \geq n_R$, in the realistic scenario of signal correlation between different receiver antennas. It did so by introducing some simple concepts from the theory of linear transformations and the theory of group representations, making the exposition virtually self-contained. The main result was extended to independent fading scenarios, yielding a new closed-form expression. Additionally, plots of the capacity for the isotropic scattering model in distinct MIMO arrangements were obtained, showing a substantial capacity reduction in the presence of correlation. Finally, it was proved that equispaced linear arrays do not attain the maximum capacity, and that by equipping the receiver with more antennas than signal detectors it becomes possible to realize capacity gains.

4.6 *Capacity when the Wishart Matrix is Singular*

In Section 4.5 an exact, closed-form, analytic solution was found for the particular case of no more receiver antennas than transmitter antennas ($n_T \geq n_R$). Although important and useful, that solution did not contemplate the case $n_T < n_R$, rendering it somewhat non-universal. Correlation is being assumed solely at one side of the communication link, and consequently one must assume that the antennas at the other side are sufficiently spaced so as to provide independent transmission/reception diversity and avoid mutual coupling effects. This is often a reasonable assumption if the radio links consist of a base station (BS) at one end and a mobile unit (MU) at the other end, provided that the antennas at the base station have separations of several wavelengths. In such a scenario, the total number of antennas at the base station must be limited, possibly being less than the number of antennas at the mobile unit. Recall that, even under correlation conditions and fixed array lengths, the utmost increase in capacity is obtained by increasing the number of receiver antennas.

In this section we fill the aforementioned gap by following an approach identical to the previous, and try to extend it to transmitter-sided correlation and two-sided independent fading. It is advisable for the reader to be acquainted with Section 4.5, since some of the results and techniques that will be used here (without proof) rely on material that has been properly introduced there. Namely:

1. if $\mathbf{Q}$ is a complex generic $m \times n$ $(m \geq n)$ random matrix with statistically independent entries, it clearly defines a $2mn$-dimensional Euclidean space; if, in addition, matrix $\mathbf{Q}$ satisfies the condition $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$, it is implicit that a set of n^2 independent equations must be satisfied, which restricts the original space to a topological space of dimension $2mn - n^2$; this topological space, denoted by $\widetilde{V}_{n,m}$, is closed and bounded, hence compact, and is called the *complex Stiefel manifold*; if $m = n$, then the Stiefel manifold is called the unitary group, and is denoted by $\widetilde{V}_{n,n} \sim \widetilde{U}(n)$;
2. since $\widetilde{V}_{n,m}$ is compact, it is measurable and has a finite volume; this volume is given by

$$\text{Volume}\left[\widetilde{V}_{n,m}\right] = \int_{\widetilde{V}_{n,m}} (d\mathbf{Q})^{\wedge} = \frac{2^n \pi^{mn}}{\tilde{\Gamma}_n(m)} \tag{4.303}$$

where $\tilde{\Gamma}_n(m)$ is the complex multivariate gamma function defined in (4.133);

3. adding the extra restriction to $\mathbf{Q}$ of real diagonal elements (i.e. fixing the phase of the columns so that the diagonal is real), one obtains the restricted Stiefel manifold $\widetilde{V}_{n,m}^{\subset}$, which has the volume

$$\text{Volume}\left[\widetilde{V}_{n,m}^{\subset}\right] = \int_{\widetilde{V}_{n,m}^{\subset}} (d\mathbf{Q})^{\wedge} = \frac{\pi^{n(m-1)}}{\tilde{\Gamma}_n(m)} \tag{4.304}$$

4. the following equality was also obtained:

$$\int_{\widetilde{V}_{n,m}^{\subset}} h(\mathbf{Q})(d\mathbf{Q})^{\wedge} = \frac{1}{(2\pi)^n} \int_{\widetilde{V}_{n,m}} h(\mathbf{Q})(d\mathbf{Q})^{\wedge} \tag{4.305}$$

which is valid provided that the function $h(\mathbf{Q})$ does not depend on the particular phase of the columns of $\mathbf{Q}$;

5. in addition, the following integral over the unitary group has been evaluated

$$\begin{aligned} \int_{\widetilde{U}(n)} e^{tr(\mathbf{AUBU}^H)} (d\mathbf{U})^{\wedge} = \\ = \text{Volume}\left[\widetilde{U}(n)\right] \frac{\det\left[(i-1)!e^{a_i b_i}\right]}{\det\left[a_i^{n-j}\right]\det\left[b_i^{n-j}\right]}, \end{aligned} \tag{4.306}$$

where $\mathbf{A}$ and $\mathbf{B}$ are generic, full-rank, $n \times n$ complex matrices, and $\{a_i\}$ and $\{b_i\}$ are their respective eigenvalues.

4.6.1 Statistical Characterization

We are solely interested in studying frequency-flat multipath wireless channels, that is, channels that have a small delay spread as compared to the period of the transmitted symbols. If the channel is frequency-selective, it is assumed that it can be converted to frequency-flat. Also, for typical data rates, it is a good approximation to assume that the channel is slowly-fading, i.e. virtually time-invariant during a symbol period. Under these conditions, it is possible to model each MIMO channel transmission with the equation

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} \tag{4.307}$$

where $\mathbf{x}$ is the $n_T \times 1$ input vector, $\mathbf{y}$ is the $n_R \times 1$ output vector, $\mathbf{n}$ is the $n_R \times 1$ AWGN noise vector and $\mathbf{H}$ is the $n_R \times n_T$ random channel matrix. The average (ergodic) capacity of the MIMO channel was determined to be given by

$$C_{avg} = E_{\mathbf{H}}\left[\log_2 \det\left(\mathbf{I}_{n_R} + \frac{\gamma}{n_T}\mathbf{HH}^H\right)\right] \tag{4.308}$$

where γ is the average SNR at each receiver branch, n_T and n_R are the number of transmitter and receiver antennas, respectively, and $E_{\mathbf{H}}[\cdot]$ denotes the expectation over $\mathbf{H}$. We call matrix $\mathcal{A} = \mathbf{HH}^H$ the Wishart matrix and let $f_{\mathbf{H}}(\mathbf{H})$ denote the joint density function

of $\mathbf{H}$'s entries. Using the eigenvalue decomposition $\boldsymbol{\mathcal{A}} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H$, where $\mathbf{Q}$ is a unitary matrix, we may also write (4.308) as

$$C_{avg} = \sum_{k=1}^{m} \int_{\boldsymbol{\Lambda}} \log_2\left(1 + \frac{\gamma}{n_T}\lambda_k\right) f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})(d\boldsymbol{\Lambda})^{\wedge} \tag{4.309}$$

where m is the number of non-zero eigenvalues of $\boldsymbol{\mathcal{A}}$, λ_k is some eigenvalue, $\boldsymbol{\Lambda} = diag(\lambda_1,..,\lambda_m)$, $f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})$ is the joint pdf of the eigenvalues, which must be derived from $f_{\mathbf{H}}(\mathbf{H})$, and $(d\boldsymbol{\Lambda})^{\wedge}$ is the volume element in the eigenvalue coordinates. Assuming a Rayleigh fading environment, and denoting by $\boldsymbol{\Sigma}_{\mathrm{T}}$ and $\boldsymbol{\Sigma}_{\mathrm{R}}$ the transmit and receive correlation matrices, respectively, it was found that if correlation exists only at the receiver side, then

$$f_{\mathbf{H}}^{\boldsymbol{\Sigma}_{\mathrm{T}}=\mathbf{I}_{n_T}}(\mathbf{H}) = [\det(\pi\boldsymbol{\Sigma}_{\mathrm{R}})]^{-n_T} \exp\left[-tr\left(\boldsymbol{\Sigma}_{\mathrm{R}}^{-1}\mathbf{H}\mathbf{H}^H\right)\right] \tag{4.310}$$

If correlation exists only at the transmitter side, then it is easy to verify that

$$f_{\mathbf{H}}^{\boldsymbol{\Sigma}_{\mathrm{R}}=\mathbf{I}_{n_R}}(\mathbf{H}) = [\det(\pi\boldsymbol{\Sigma}_{\mathrm{T}})]^{-n_R} \exp\left[-tr\left(\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{H}^H\mathbf{H}\right)\right] \tag{4.311}$$

It was previously said that the capacity for the case $n_T \geq n_R$ was computed in Section 4.5. The problem with the case $n_T < n_R$ is that the Wishart matrix $\boldsymbol{\mathcal{A}} = \mathbf{H}\mathbf{H}^H$ becomes rank-deficient (i.e. not full-rank, hence singular), which leads to serious complications in the derivation of the eigenvalue pdf $f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})$. In the following subsections we show that it is possible to overcome these difficulties.

4.6.2 VOLUME ELEMENTS OF TRANSFORMATIONS

Before proceeding with the derivation of the density functions $f_{\boldsymbol{\mathcal{A}}}(\boldsymbol{\mathcal{A}})$ and $f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})$, one must first obtain the relations between the volume elements associated with the following transformations:

1. **The Cholesky factorization** $\mathbf{S} = \mathbf{R}^H\mathbf{R}$ of the hermitian, positive definite, $n \times n$ matrix $\mathbf{S}$, where $\mathbf{R}$ is an upper-triangular $n \times n$ matrix with positive and real diagonal entries. From (4.95) we have

$$(d\mathbf{S})^{\wedge} = 2^n \prod_{k=1}^{n} r_{kk}^{2n-2k+1} (d\mathbf{R})^{\wedge} \tag{4.312}$$

2. **The QR factorization** $\mathbf{S} = \mathbf{Q}\mathbf{R} = \mathbf{Q}\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\end{pmatrix}$ of the complex $m \times n$ $(n > m)$ matrix $\mathbf{S}$ of rank m, where $\mathbf{Q}$ is a unitary $m \times m$ matrix, $\mathbf{R}_1$ is an upper-triangular $m \times m$ matrix, and $\mathbf{R}_2$ is an $m \times (n-m)$ matrix. Each matrix has the following number of functionally independent entries: matrix $\mathbf{S}$ - $2mn$; matrix $\mathbf{Q}$ - m^2 (due to $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$); matrix $\mathbf{R}_1$ - $m^2 + m$; and matrix $\mathbf{R}_2$ - $2mn - 2m^2$. To make the factorization a one-to-one mapping we force the diagonal elements of $\mathbf{R}_1$ to be real, so that the right side of the factorization has a total of $2mn$ functionally independent entries. Using (4.86), we write

$$\begin{aligned}(d\mathbf{S})^{\wedge} &= |\det\mathbf{Q}^H|^{2n}(d\mathbf{S})^{\wedge} = (\mathbf{Q}^H d\mathbf{S})^{\wedge} \\ &= \left(\mathbf{Q}^H d\mathbf{Q}\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\end{pmatrix} + \begin{pmatrix}d\mathbf{R}_1 & d\mathbf{R}_2\end{pmatrix}\right)^{\wedge}.\end{aligned} \tag{4.313}$$

Also, we have

$$\left[\mathbf{Q}^H d\mathbf{Q}\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\end{pmatrix}\right]_{ij} = \begin{cases}\sum_{k=1}^{j}[\,\mathbf{q}_i^H d\mathbf{q}_k\,]r_{kj} & j \le m\\ \sum_{k=1}^{m}[\,\mathbf{q}_i^H d\mathbf{q}_k\,]r_{kj} & j > m\end{cases} \tag{4.314}$$

Since $\mathbf{Q}^H d\mathbf{Q}$ is skew-hermitian, and noting that products of repeated differentials are zero, i.e. $(\,\mathbf{q}_i^H d\mathbf{q}_k\,)^\wedge \wedge (\,\mathbf{q}_i^H d\mathbf{q}_k\,)^\wedge = (\,\mathbf{q}_i^H d\mathbf{q}_k\,)^\wedge \wedge (\,\mathbf{q}_k^H d\mathbf{q}_i\,)^\wedge = 0$, the entries in (4.314) for $j > m$ do not contribute to the exterior product, and only the last term of the diagonal and sub-diagonal entries of $\mathbf{Q}^H d\mathbf{Q}\mathbf{R}_1$ is not redundant. It follows that

$$\left(\mathbf{Q}^H d\mathbf{Q}\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\end{pmatrix}\right)^\wedge = \prod_{k=1}^{m} r_{kk}^{2m-2k+1} \bigwedge_{i\ge j}^{m}(\,\mathbf{q}_i^H d\mathbf{q}_j\,)^\wedge \tag{4.315}$$

and because

$$\bigwedge_{i\ge j}^{m}(\,\mathbf{q}_i^H d\mathbf{q}_j\,)^\wedge = (\,\mathbf{Q}^H d\mathbf{Q}\,)^\wedge = (\,d\mathbf{Q}\,)^\wedge \tag{4.316}$$

we hence have

$$(\,d\mathbf{S}\,)^\wedge = \prod_{k=1}^{m} r_{kk}^{2m-2k+1}\,(\,d\mathbf{Q}\,)^\wedge \wedge (\,d\mathbf{R}\,)^\wedge \tag{4.317}$$

3. The Congruence transformation $\mathbf{S} = \mathbf{B}^H\mathbf{A}\mathbf{B}$ between two hermitian $n \times n$ matrices $\mathbf{A}$ and $\mathbf{S}$ of rank $q < n$, where $\mathbf{B}$ is a fixed nonsingular $n \times n$ matrix. It is assumed that there is an $n \times q$ matrix $\mathbf{X}$ such that $\mathbf{A}$ factorizes as $\mathbf{A} = \mathbf{X}^H\mathbf{X}$. The Congruence transformation can be written in block form as

$$\begin{pmatrix}\mathbf{S}_{11} & \mathbf{S}_{12}\\ \mathbf{S}_{12}^H & \mathbf{S}_{22}\end{pmatrix} = \begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{pmatrix}^H \begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{12}^H & \mathbf{A}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{pmatrix} \tag{4.318}$$

where $\mathbf{S}_{11}$ and $\mathbf{A}_{11}$ are hermitian, positive definite, $q \times q$ matrices of rank q. Matrix $\mathbf{S}$ contains $q(q+1) + 2q(n-q)$ functionally independent elements, which correspond, for instance, to the elements of $\mathbf{S}_{12}$ plus the diagonal and above diagonal elements of $\mathbf{S}_{11}$. The same applies to matrix $\mathbf{A}$. The volume elements of $\mathbf{S}$ and $\mathbf{A}$ are thus given by

$$\begin{aligned}(\,d\mathbf{S}\,)^\wedge &= (\,d\mathbf{S}_{11}\,)^\wedge \wedge (\,d\mathbf{S}_{12}\,)^\wedge\\ (\,d\mathbf{A}\,)^\wedge &= (\,d\mathbf{A}_{11}\,)^\wedge \wedge (\,d\mathbf{A}_{12}\,)^\wedge\end{aligned} \tag{4.319}$$

We need to find

$$\frac{(\,d\mathbf{S}\,)^\wedge}{(\,d\mathbf{A}\,)^\wedge} = \frac{(\,d\mathbf{S}_{11}\,)^\wedge \wedge (\,d\mathbf{S}_{12}\,)^\wedge}{(\,d\mathbf{A}_{11}\,)^\wedge \wedge (\,d\mathbf{A}_{12}\,)^\wedge} = \frac{(\,d\mathbf{S}_{11}\,)^\wedge}{(\,d\mathbf{A}_{11}\,)^\wedge}\frac{(\,d\mathbf{S}_{12} \mid \mathbf{S}_{11}\,)^\wedge}{(\,d\mathbf{A}_{12}\,)^\wedge} \tag{4.320}$$

where $(d\mathbf{S}_{12} \mid \mathbf{S}_{11})^{\wedge}$ denotes the differential of matrix $\mathbf{S}_{12}$ given $\mathbf{S}_{11}$. From (4.318) it is found that

$$\begin{aligned}\mathbf{S}_{11} &= \mathbf{B}_{11}^H\mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{B}_{21}^H\mathbf{A}_{12}^H\mathbf{B}_{11} + \mathbf{B}_{11}^H\mathbf{A}_{12}\mathbf{B}_{21} + \mathbf{B}_{21}^H\mathbf{A}_{22}\mathbf{B}_{21}\\ \mathbf{S}_{12} &= \mathbf{B}_{11}^H\mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{B}_{21}^H\mathbf{A}_{12}^H\mathbf{B}_{12} + \mathbf{B}_{11}^H\mathbf{A}_{12}\mathbf{B}_{22} + \mathbf{B}_{21}^H\mathbf{A}_{22}\mathbf{B}_{22}\\ &= \mathbf{B}_{11}^H\mathbf{A}_{12}\left(\mathbf{B}_{22} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12}\right) + \mathbf{S}_{11}\mathbf{B}_{11}^{-1}\mathbf{B}_{12} -\\ &\qquad -\mathbf{B}_{21}^H\mathbf{A}_{22}\mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12} + \mathbf{B}_{21}^H\mathbf{A}_{22}\mathbf{B}_{22}\end{aligned} \tag{4.321}$$

Differentiating these equations, and using (4.86) and (4.115), yields

$$(d\mathbf{S}_{11})^{\wedge} = |\det\mathbf{B}_{11}|^{2q}\,(d\mathbf{A}_{11})^{\wedge} \tag{4.322}$$

and

$$(d\mathbf{S}_{12} \mid \mathbf{S}_{11})^{\wedge} = |\det\mathbf{B}_{11}|^{2(n-q)-2q}\,|\det\mathbf{B}|^{2q}\,(d\mathbf{A}_{11})^{\wedge} \tag{4.323}$$

where the following property of the determinant has been used:

$$\det\mathbf{B} = \det\begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{pmatrix} = \det\mathbf{B}_{11}\,\det\left(\mathbf{B}_{22} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12}\right) \tag{4.324}$$

Consequently, we find

$$(d\mathbf{S})^{\wedge} = |\det\mathbf{B}_{11}|^{2(n-q)}\,|\det\mathbf{B}|^{2q}\,(d\mathbf{A})^{\wedge} \tag{4.325}$$

4. The Eigenvalue decomposition $\mathbf{S} = \mathbf{Q}_1\mathbf{\Lambda}\mathbf{Q}_1^H$ of the hermitian $n \times n$ matrix $\mathbf{S}$ of rank $q < n$, where $\mathbf{\Lambda}$ is a $q \times q$ diagonal matrix with the non-zero eigenvalues of $\mathbf{S}$, and $\mathbf{Q}_1$ is an $n \times q$ semi-unitary matrix (such that $\mathbf{Q}_1^H\mathbf{Q}_1 = \mathbf{I}$). Alternatively, the decomposition can be rewritten as

$$\mathbf{S} = \overline{\mathbf{Q}}\begin{pmatrix}\mathbf{\Lambda} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\overline{\mathbf{Q}}^H \tag{4.326}$$

where $\overline{\mathbf{Q}} = \begin{pmatrix}\mathbf{Q}_1 & \mathbf{Q}_1^{\perp}\end{pmatrix}$ and $\mathbf{Q}_1^{\perp}$ is the orthogonal complement of $\mathbf{Q}_1$. Defining $\mathbf{Q}_1^H = \begin{pmatrix}\mathbf{Q}_{11}^H & \mathbf{Q}_{12}^H\end{pmatrix}$, where $\mathbf{Q}_{11}$ is $q \times q$, by (4.325) we know that

$$\begin{aligned}\left(\overline{\mathbf{Q}}^H\mathbf{dS}\overline{\mathbf{Q}}\right)^{\wedge} &= |\det\mathbf{Q}_{11}|^{2(n-q)}\,(\mathbf{dS})^{\wedge}\\ &= \left(\begin{pmatrix}d\mathbf{\Lambda}\\ \mathbf{0}\end{pmatrix} + \overline{\mathbf{Q}}^H d\overline{\mathbf{Q}}\begin{pmatrix}\mathbf{\Lambda}\\ \mathbf{0}\end{pmatrix} + \left(\overline{\mathbf{Q}}^H d\overline{\mathbf{Q}}\begin{pmatrix}\mathbf{\Lambda}\\ \mathbf{0}\end{pmatrix}\right)^H\right)^{\wedge}.\end{aligned} \tag{4.327}$$

It is easily checked that, for $i \le n$ and $j \le q$, we have

$$\left[\overline{\mathbf{Q}}^H d\overline{\mathbf{Q}}\begin{pmatrix}\mathbf{\Lambda}\\ \mathbf{0}\end{pmatrix} + \left(\overline{\mathbf{Q}}^H d\overline{\mathbf{Q}}\begin{pmatrix}\mathbf{\Lambda}\\ \mathbf{0}\end{pmatrix}\right)^H\right]_{ij} = \overline{\mathbf{q}}_i^H d\overline{\mathbf{q}}_j(\lambda_j - \lambda_i) \tag{4.328}$$

Moreover, since $\overline{\mathbf{Q}}^H d\overline{\mathbf{Q}}$ is skew hermitian, it follows that $\overline{\mathbf{q}}_i^H d\overline{\mathbf{q}}_j = -\overline{\mathbf{q}}_j^H d\overline{\mathbf{q}}_i$, and thus only the diagonal and subdiagonal elements given by (4.328) are relevant when taking exterior

products (i.e. $i \geq j$). This leads to

$$\begin{aligned}\left(\overline{\mathbf{Q}}^H \mathbf{dS}\overline{\mathbf{Q}}\right)^{\wedge} = \overbrace{\prod_{j=1}^{q}\lambda_j^{2(n-q)} \bigwedge_{j=1}^{q}\bigwedge_{i=q+1}^{n}\left(\mathbf{q}_i^H d\mathbf{q}_j\right)^{\wedge}}^{j<q\leq i\leq n} \wedge \\ \wedge \overbrace{\prod_{i>j}^{q}(\lambda_j-\lambda_i)^2 \bigwedge_{i>j}^{q}\left(\mathbf{q}_i^H d\mathbf{q}_j\right)^{\wedge}}^{j<i\leq q} \wedge \overbrace{\bigwedge_{i=1}^{q} d\lambda_i \wedge \left(\mathbf{q}_i^H d\mathbf{q}_i\right)^{\wedge}}^{i=j},\end{aligned} \tag{4.329}$$

which, after plugging exterior product factors together, and using the identities

$$\begin{aligned}(d\mathbf{Q}_1)^{\wedge} = \left(\overline{\mathbf{Q}}^H d\mathbf{Q}_1\right)^{\wedge} = \bigwedge_{j=1}^{q}\bigwedge_{i=j}^{n}\left(\mathbf{q}_i^H d\mathbf{q}_j\right)^{\wedge} \\ \det\left[\lambda_i^{q-j}\right] = \prod_{i<j}^{q}(\lambda_i-\lambda_j),\end{aligned} \tag{4.330}$$

yields

$$\left(\overline{\mathbf{Q}}^H \mathbf{dS}\overline{\mathbf{Q}}\right)^{\wedge} = (\det\boldsymbol{\Lambda})^{2(n-q)}\left(\det\left[\lambda_i^{q-j}\right]\right)^2 (d\mathbf{Q}_1)^{\wedge}\wedge(d\boldsymbol{\Lambda})^{\wedge} \tag{4.331}$$

Substituting this result back into (4.327), one reaches the final relation between the volume elements

$$(\mathbf{dS})^{\wedge} = |\det\mathbf{Q}_{11}|^{2(q-n)}\frac{\left(\det\left[\lambda_i^{q-j}\right]\right)^2}{(\det\boldsymbol{\Lambda})^{2(q-n)}}(d\mathbf{Q}_1)^{\wedge}\wedge(d\boldsymbol{\Lambda})^{\wedge} \tag{4.332}$$

4.6.3 The Wishart and Eigenvalue Densities

We now determine the pdf of the Wishart matrix $\boldsymbol{\mathcal{A}} = \mathbf{H}\mathbf{H}^H$. Since we are considering the case $n_T < n_R$ we note that, in statistical terms, matrix $\mathbf{H}$ contains n_T linearly independent columns. Consequently, we can use the factorization $\mathbf{H}^H = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q}$ is a unitary $n_T \times n_T$ matrix and $\mathbf{R} = \begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\end{pmatrix}$ is an $n_T \times n_R$ matrix with $\mathbf{R}_1$ $n_T \times n_T$ upper-triangular. Expressing the joint pdf of $\mathbf{Q}$ and $\mathbf{R}$ as

$$f_{\mathbf{QR}}(\mathbf{Q},\mathbf{R})(d\mathbf{Q})^{\wedge} = f_{\mathbf{H}}^{\boldsymbol{\Sigma}_T=\mathbf{I}_{n_T}}(\mathbf{R}^H\mathbf{Q}^H)\frac{(d\mathbf{H})^{\wedge}}{(d\mathbf{R})^{\wedge}} \tag{4.333}$$

and inserting (4.317) gives

$$f_{\mathbf{QR}}(\mathbf{Q},\mathbf{R})(d\mathbf{Q})^{\wedge} = f_{\mathbf{H}}^{\boldsymbol{\Sigma}_T=\mathbf{I}_{n_T}}(\mathbf{R}^H\mathbf{Q}^H)\prod_{k=1}^{n_T} r_{kk}^{2n_T-2k+1}(d\mathbf{Q})^{\wedge} \tag{4.334}$$

Now, since $\boldsymbol{\mathcal{A}}$ is given by

$$\boldsymbol{\mathcal{A}} = \mathbf{R}^H\mathbf{R} = \begin{pmatrix}\mathbf{R}_1^H\mathbf{R}_1 & \mathbf{R}_1^H\mathbf{R}_2 \\ \mathbf{R}_2^H\mathbf{R}_1 & \mathbf{R}_2^H\mathbf{R}_2\end{pmatrix} = \begin{pmatrix}\boldsymbol{\mathcal{A}}_{11} & \boldsymbol{\mathcal{A}}_{12} \\ \boldsymbol{\mathcal{A}}_{12}^H & \boldsymbol{\mathcal{A}}_{22}\end{pmatrix} \tag{4.335}$$

where $\boldsymbol{\mathcal{A}}_{11}$ is $n_T \times n_T$, we notice the dependence $\boldsymbol{\mathcal{A}}_{22} = \boldsymbol{\mathcal{A}}_{12}^H\boldsymbol{\mathcal{A}}_{11}^{-1}\boldsymbol{\mathcal{A}}_{12}$, which means that

$$f_{\boldsymbol{\mathcal{A}}}(\boldsymbol{\mathcal{A}})(d\boldsymbol{\mathcal{A}})^{\wedge} = f_{\boldsymbol{\mathcal{A}}_{11}\boldsymbol{\mathcal{A}}_{12}}(\boldsymbol{\mathcal{A}}_{11},\boldsymbol{\mathcal{A}}_{12})(d\boldsymbol{\mathcal{A}}_{11})^{\wedge}\wedge(d\boldsymbol{\mathcal{A}}_{12})^{\wedge} \tag{4.336}$$

Hence, we can also write

$$\begin{aligned} f_{\mathcal{A}}(\mathcal{A}) &= \frac{(d\mathbf{R}_1)^{\wedge} \wedge (d\mathbf{R}_2)^{\wedge}}{(d\mathcal{A}_{11})^{\wedge} \wedge (d\mathcal{A}_{12})^{\wedge}} f_{\mathbf{R}_1\mathbf{R}_2}(\mathbf{R}_1, \mathbf{R}_2) \\ &= \frac{(d\mathbf{R}_1)^{\wedge}}{(d\mathcal{A}_{11})^{\wedge}} \frac{(d\mathbf{R}_2 \mid \mathbf{R}_1)^{\wedge}}{(d\mathcal{A}_{12})^{\wedge}} \int_{\tilde{V}_{n_T,n_T}} f_{\mathbf{QR}}(\mathbf{Q}, \mathbf{R})(d\mathbf{Q})^{\wedge}. \end{aligned} \tag{4.337}$$

From result (4.312) for the Cholesky factorization

$$\frac{(d\mathbf{R}_1)^{\wedge}}{(d\mathcal{A}_{11})^{\wedge}} = \left[2^{n_T} \prod_{k=1}^{n_T} r_{kk}^{2n_T - 2k + 1} \right]^{-1} \tag{4.338}$$

and from (4.86)

$$\frac{(d\mathbf{R}_2 \mid \mathbf{R}_1)^{\wedge}}{(d\mathcal{A}_{12})^{\wedge}} = |\det \mathbf{R}_1|^{2(n_T - n_R)} = (\det \mathcal{A}_{11})^{n_T - n_R} \tag{4.339}$$

Expanding (4.310) in (4.334), inserting the result into (4.337), and then using (4.338) and (4.339), we find that the pdf of the complex, singular Wishart matrix is given by

$$f_{\mathcal{A}}(\mathcal{A}) = \frac{\pi^{n_T(n_T - n_R)}}{(\det \mathbf{\Sigma}_{\mathrm{R}})^{n_T} \tilde{\Gamma}_{n_T}(n_T)} e^{-tr(\mathbf{\Sigma}_{\mathrm{R}}^{-1}\mathcal{A})} (\det \mathcal{A}_{11})^{n_T - n_R} \tag{4.340}$$

It is now possible to derive the joint pdf $f_{\mathbf{\Lambda}}(\mathbf{\Lambda})$ of the non-zero eigenvalues of $\mathcal{A}$. Since $rank(\mathcal{A}) = n_T$, it has n_T non-zero eigenvalues. Thus, we define the Eigenvalue decomposition $\mathcal{A} = \mathbf{Q}_1 \mathbf{\Lambda}^{\mathrm{o}} \mathbf{Q}_1^H$, where $\mathbf{\Lambda}^{\mathrm{o}}$ is the $n_T \times n_T$ diagonal matrix of the ordered eigenvalues of $\mathbf{\Lambda}$, and $\mathbf{Q}_1$ is a semi-unitary $n_R \times n_T$ $(n_T < n_R)$ matrix with real diagonal. Furthermore, we write the joint pdf of $\mathbf{\Lambda}^{\mathrm{o}}$ and $\mathbf{Q}_1$ as

$$f_{\mathbf{\Lambda Q}_1}(\mathbf{\Lambda}^{\mathrm{o}}, \mathbf{Q}_1)(d\mathbf{Q}_1)^{\wedge} = f_{\mathcal{A}}(\mathbf{Q}_1 \mathbf{\Lambda}^{\mathrm{o}} \mathbf{Q}_1^H) \frac{(d\mathcal{A})^{\wedge}}{(d\mathbf{\Lambda}^{\mathrm{o}})^{\wedge}} \tag{4.341}$$

Substitution of (4.332) and (4.340) yields

$$\begin{aligned} f_{\mathbf{\Lambda Q}_1}(\mathbf{\Lambda}^{\mathrm{o}}, \mathbf{Q}_1)(d\mathbf{Q}_1)^{\wedge} &= \frac{\pi^{n_T(n_T - n_R)}}{(\det \mathbf{\Sigma}_{\mathrm{R}})^{n_T} \tilde{\Gamma}_{n_T}(n_T)} \cdot \\ &\cdot e^{-tr\left(\mathbf{\Sigma}_{\mathrm{R}}^{-1} \overline{\mathbf{Q}} \begin{pmatrix} \mathbf{\Lambda}^{\mathrm{o}} & 0 \\ 0 & 0 \end{pmatrix} \overline{\mathbf{Q}}^H\right)} \prod_{j=1}^{n_T} \lambda_j^{n_R - n_T} \left(\det\left[\lambda_i^{n_T - j}\right]\right)^2 (d\mathbf{Q}_1)^{\wedge}, \end{aligned} \tag{4.342}$$

where $\overline{\mathbf{Q}} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_1^{\perp} \end{pmatrix}$ and $\mathbf{Q}_1^{\perp}$ is the orthogonal complement of $\mathbf{Q}_1$. Equation (4.342) can be integrated over the restricted Stiefel manifold $\tilde{V}_{n_T,n_R}^{\subset}$ to give

$$\begin{aligned} f_{\mathbf{\Lambda}}(\mathbf{\Lambda}^{\mathrm{o}}) &= G'_{n_R,n_T}(\mathbf{\Lambda}^{\mathrm{o}}, \mathbf{\Sigma}_{\mathrm{R}}) \cdot \\ &\cdot \int_{\tilde{V}_{n_T,n_R}^{\subset}} e^{-tr\left\{\mathbf{\Sigma}_{\mathrm{R}}^{-1} \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_1^{\perp} \end{pmatrix} \begin{pmatrix} \mathbf{\Lambda}^{\mathrm{o}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_1^{\perp} \end{pmatrix}^H\right\}} (d\mathbf{Q}_1)^{\wedge}, \end{aligned} \tag{4.343}$$

where

$$G'_{m,n}(\mathbf{\Lambda}, \mathbf{\Sigma}) = \frac{\pi^{n(n-m)} \left(\det\left[\lambda_i^{n-j}\right]\right)^2 (\det \mathbf{\Lambda})^{m-n}}{(\det \mathbf{\Sigma})^n \tilde{\Gamma}_n(n)} \tag{4.344}$$

Averaging over all permutations, applying (4.305), and making the transformation $\mathbf{U}_1 = \mathbf{Q}_1\mathbf{P}$, where $\mathbf{P}$ is $n_T \times n_T$ orthogonal and $\mathbf{\Lambda}^{\mathrm{o}} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^H$, we are led to

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = \frac{G'_{n_R,n_T}(\mathbf{\Lambda},\mathbf{\Sigma}_{\mathrm{R}})}{n_T\,!(2\pi)^{n_T}} \cdot \cdot \int_{\widetilde{V}_{n_T,n_R}} e^{-tr\left\{\mathbf{\Sigma}_{\mathrm{R}}^{-1}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}\begin{pmatrix}\mathbf{\Lambda} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}^H\right\}} (d\mathbf{U}_1)^{\wedge} \tag{4.345}$$

At this point, we require a closed-form expression for the integral in (4.345). This problem can be solved by extending (4.306) to the case of a singular $\mathbf{B}$ matrix, and noting that in (4.345) we are in fact integrating over a confined domain, more specifically the quotient space $\widetilde{U}(n_R)/\widetilde{U}(n_R - n_T)$. Ignoring the singularity of matrix $\mathbf{D} = diag(\mathbf{\Lambda},\mathbf{0})$, we could thus write

$$\int_{\widetilde{V}_{n_T,n_R}} e^{-tr\left\{\mathbf{\Sigma}_{\mathrm{R}}^{-1}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}\mathbf{D}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}^H\right\}} (d\mathbf{U}_1)^{\wedge} = = \mathrm{Volume}\left[\widetilde{V}_{n_T,n_R}\right]\frac{\det\left[(i-1)!e^{-d_i/\upsilon_j}\right]}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]\det\left[d_i^{n_R-j}\right]} \tag{4.346}$$

where $\{\upsilon_i\}$ and $\{d_i\}$ are the eigenvalues of $\mathbf{\Sigma}_{\mathrm{R}}$ and $\mathbf{D}$, respectively. Since $n_R - n_T$ eigenvalues of $\mathbf{D}$ are zero, say $d_{n_T+1},..,d_{n_R}$, (4.346) is the indetermination $0/0$. The indetermination can be dealt with using the fact that the integrand function is uniformly convergent as the eigenvalues of $\mathbf{D}$ approach zero, which allows us to write

$$\int_{\widetilde{V}_{n_T,n_R}} e^{-tr\left\{\mathbf{\Sigma}_{\mathrm{R}}^{-1}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}\mathbf{D}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}^H\right\}} (d\mathbf{U}_1)^{\wedge} = = \frac{\mathrm{Volume}\left[\widetilde{V}_{n_T,n_R}\right]}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]}\lim_{d_{n_T+1},..,d_{n_R}\to 0}\frac{\det\left[(i-1)!e^{-d_i/\upsilon_j}\right]}{\det\left[d_i^{n_R-j}\right]}. \tag{4.347}$$

It is now possible to use L'Hôpital's rule by differentiating both numerator and denominator inside the limit, $n_R - i$ times with respect to d_i, for $i > n_T$. We get

$$\int_{\widetilde{V}_{n_T,n_R}} e^{-tr\left\{\mathbf{\Sigma}_{\mathrm{R}}^{-1}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}\begin{pmatrix}\mathbf{\Lambda} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_1^{\perp}\end{pmatrix}^H\right\}} (d\mathbf{U}_1)^{\wedge} = = \frac{\mathrm{Volume}\left[\widetilde{V}_{n_T,n_R}\right]}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]}\left[\prod_{p=n_R-n_T}^{n_R-1} p!\right]\frac{\det\left[\left(-\upsilon_j\right)^{(i-n_R)\sigma_{in_T}} e^{-\lambda_i/\upsilon_j}\right]}{(\det\mathbf{\Lambda})^{n_R-n_T}\det\left[\lambda_i^{n_T-j}\right]}, \tag{4.348}$$

where σ_{ab} has value 1 if $a > b$, and 0 otherwise. If we substitute (4.348) into (4.345) and rearrange the result, we finally find that the joint eigenvalue density is

$$f_{\mathbf{\Lambda}}(\mathbf{\Lambda}) = \frac{H'_{n_R,n_T}(\boldsymbol{\upsilon}_{\mathrm{R}})}{\left(\det\left[\lambda_i^{n_T-j}\right]\right)^{-1}}\det\left[\left(-\upsilon_j\right)^{(i-n_R)\sigma_{in_T}} e^{-\lambda_i/\upsilon_j}\right] \tag{4.349}$$

where $\boldsymbol{\upsilon}_{\mathrm{R}}$ is the vector of eigenvalues of $\mathbf{\Sigma}_{\mathrm{R}}$,

$$H'_{m,n}(\boldsymbol{\upsilon}) = \frac{\pi^{n(n-m)}}{n!(2\pi)^n\widetilde{\Gamma}_n(n)} \cdot \cdot \frac{\mathrm{Volume}\left[\widetilde{V}_{n,m}\right]}{\det\left[(-1)^{m-j}\upsilon_i^{n-m+j}\right]}\left[\prod_{p=m-n}^{m-1} p!\right], \tag{4.350}$$

and $\text{Volume}\left[\widetilde{V}_{n,m}\right]$ is obtained from (4.303).

4.6.4 CAPACITY UNDER RECEIVER-SIDED CORRELATED FADING

Having found the joint pdf of the eigenvalues of the complex singular Wishart matrix, we now derive the average capacity of the MIMO wireless channel for the scenario $\mathbf{\Sigma}_{\text{T}} = \mathbf{I}_{n_T}$ and $n_T < n_R$. To that end we evaluate (4.309) with (4.349), yielding

$$\begin{aligned} C_{avg} &= H'_{n_R,n_T}(\boldsymbol{\upsilon}_{\text{R}}) \sum_{k=1}^{n_T} \int_{\lambda_1} \cdots \int_{\lambda_{n_T}} \log_2\left(1+\frac{\gamma}{n_T}\lambda_k\right) \cdot \\ &\cdot \det\left[\lambda_i^{n_T-j}\right]_{n_T\times n_T} \det\left[\left(-\upsilon_j\right)^{(i-n_R)\sigma_{in_T}} e^{-\lambda_i/\upsilon_j}\right] d\lambda_{n_T}...d\lambda_1. \end{aligned} \tag{4.351}$$

Expanding the determinants, the multiple integral becomes

$$\begin{aligned} &\int_{\lambda_1} \cdots \int_{\lambda_{n_T}} \log_2\left(1+\frac{\gamma}{n_T}\lambda_k\right) \sum_{q}^{n_R} \epsilon(q) \sum_{p}^{n_T} \epsilon(p) \cdot \\ &\cdot \left[\prod_{l=n_T+1}^{n_R} \left(-\upsilon_{q_l}^{-1}\right)^{n_R-l}\right] \prod_{i=1}^{n_T} \lambda_i^{n_T-p_i} e^{-\lambda_i/\upsilon_{q_i}} d\lambda_{n_T}...d\lambda_1. \end{aligned} \tag{4.352}$$

In (4.352), we switch integration and summation operations, and incorporate the logarithm within the product, so that (4.351) becomes

$$\begin{aligned} C_{avg} &= H'_{n_R,n_T}(\boldsymbol{\upsilon}_{\text{R}}) \sum_{k=1}^{n_T} \sum_{q}^{n_R} \epsilon(q) \left[\prod_{l=n_T+1}^{n_R} \left(-\upsilon_{q_l}^{-1}\right)^{n_R-l}\right] \cdot \\ &\cdot \det\left[\int_0^{\infty} \lambda^{n_T-j} e^{-\lambda/\upsilon_{q_i}} \left[\log_2\left(1+\frac{\gamma}{n_T}\lambda\right)\right]^{\delta_{kj}} d\lambda\right]_{n_T\times n_T} \end{aligned} \tag{4.353}$$

where δ_{ab} is the Kronecker function. Keep in mind that the summation is over all n_R-tuples $q = (q_1,..,q_{n_R})$ (bounded by $0 < q_i \le n_R$), or equivalently, over all permutations of the natural sequence $(1,..,n_R)$.

Following the same steps as in (4.182)-(4.189) (essentially, integrating by parts and then using the binomial theorem) the integral in (4.353) can be expressed in closed-form. The average capacity formula becomes

$$\begin{aligned} C_{avg}^{\mathbf{\Sigma}_{\text{R}}/n_T<n_R} &= H'_{n_R,n_T}(\boldsymbol{\upsilon}_{\text{R}}) \cdot \sum_{k=1}^{n_T} \sum_{q}^{n_R} \epsilon(q) \left[\prod_{l=n_T+1}^{n_R} \left(-\upsilon_{q_l}^{-1}\right)^{n_R-l}\right] \det\left[\frac{(n_T-j)!}{\upsilon_{q_i}^{-(n_T-j+1)}} \cdot \right. \\ &\left. \cdot \left[\frac{e^{n_T/(\upsilon_{q_i}\gamma)}}{\ln 2} \sum_{p=0}^{n_T-j} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{n_T}{\upsilon_{q_i}\gamma}\right)^{p-l} \Gamma\left[l, \frac{n_T}{\upsilon_{q_i}\gamma}\right]\right]^{\delta_{kj}}\right]_{n_T\times n_T} \end{aligned} \tag{4.354}$$

where $\Gamma(l,1/a)$ is the incomplete gamma function. This completes the derivation for $n_T < n_R$. The expression for the case $n_T \ge n_R$ is given in (4.286).

Evidence has shown that these capacity expressions are prone to numerical overflow/underflow when the eigenvalues of the covariance matrix are very small and/or n_T is large. This leads to a large $x_i = n_T/(\upsilon_i\gamma)$, and the consequence is that e^{x_i} overflows and $\Gamma[l,x_i]$

underflows. For $l > 0$ we deal with this problem by noting that

$$\Gamma[l,x] = e^{-x}(l-1)!\sum_{k=0}^{l-1}\frac{x^k}{k!} \qquad l > 0 \text{ (integer)} \tag{4.355}$$

and hence we compute $e^{x_i}\Gamma[l,x_i]$. For $l = 0$ and say, $x_i > 10^3$, it may be suitable to use the asymptotic expansion

$$\Gamma[0,x] \sim e^{-x}\sum_{k=0}^{N}(-1)^k\frac{k!}{x^{k+1}} \qquad (x \to \infty) \tag{4.356}$$

where N should be chosen sufficiently large (e.g. $N > 10$ is usually adequate).

- **Two-sided Independent Fading scenario**

Equation (4.293) provides a closed-form expression for the average capacity of a MIMO channel in a two-sided independent fading scenario when $n_T \geq n_R$. Now, we try to find a similar expression for $n_T < n_R$. In (4.348), we draw attention to the limit

$$\lim_{\upsilon_1,..,\upsilon_{n_R}\to 1}\frac{\det\left[(-\upsilon_i)^{(j-n_R)\sigma_{jn_T}}\,e^{-\lambda_j/\upsilon_i}\right]}{\det\left[(-\upsilon_i^{-1})^{n_R-j}\right]} = \frac{0}{0} \tag{4.357}$$

Making the change of variables $x_i = -\upsilon_i^{-1}$, and using the identity

$$f(x) = x^a e^{bx} \;\rightarrow\; f^{(n)}(x) = e^{bx}\sum_{k=0}^{n}(a)_k\binom{n}{k}b^{n-k}x^{a-k} \tag{4.358}$$

where $(a)_k$ is the *falling factorial*

$$(a)_k = \prod_{p=0}^{k-1}(a-p) \tag{4.359}$$

we apply L'Hôpital's rule and differentiate the numerator and denominator of (4.357) $n_R - i$ times with respect to x_i, thus getting

$$\lim_{x_1,..,x_{n_R}\to -1}\frac{\det\left[x_i^{(n_R-j)\sigma_{jn_T}}e^{x_i\lambda_j}\right]}{\det\left[x_i^{n_R-j}\right]} = \lim_{x_1,..,x_{n_R}\to -1}\frac{\det\left[e^{\lambda_j x_i}\sum_{s=0}^{n_R-i}\binom{n_R-i}{s}\frac{((n_R-j)\sigma_{jn_T})_s}{\lambda_j^{s+i-n_R}x_i^{s-(n_R-j)\sigma_{jn_T}}}\right]}{\det\left[(n_R-j)_{n_R-i}x_i^{i-j}\right]} \tag{4.360}$$

Setting $\{x_i\} = -1$ and using the identity

$$\det\left[(n_R-j)_{n_R-i}(-1)^{i-j}\right] = \det\left[(n_R-j)^{n_R-i}\right] \tag{4.361}$$

(4.360) is equivalent to

$$\frac{\det\left[e^{-\lambda_i}\sum_{s=0}^{n_R-j}\binom{n_R-j}{k}\frac{((n_R-i)\sigma_{in_T})_s}{\lambda_i^{s+j-n_R}(-1)^{s-(n_R-j)\sigma_{jn_T}}}\right]}{\det\left[(n_R-j)^{n_R-i}\right]} \tag{4.362}$$

Replacing (4.357) by (4.362) in (4.348) we find that the joint pdf of the eigenvalues is given by

$$f_{\mathbf{\Lambda}}^{\mathbf{\Sigma}=\mathbf{I}}(\mathbf{\Lambda}) = H'_{n_R,n_T} \det\left[\lambda_i^{n_T-j}\right] \det\left[e^{-\lambda_i} \sum_{s=0}^{n_R-j} \binom{n_R-j}{k} \frac{((n_R-i)\sigma_{in_T})_s}{\lambda_i^{s+j-n_R}(-1)^{s-(n_R-j)\sigma_{jn_T}}}\right] \tag{4.363}$$

where

$$H'_{m,n} = \frac{\pi^{n(n-m)}}{n!(2\pi)^n \tilde{\Gamma}_n(n)} \frac{\text{Volume}\left[\tilde{V}_{n,m}\right]}{\det\left[(m-j)^{m-i}\right]} \left[\prod_{p=m-n}^{m-1} p!\right] \tag{4.364}$$

By the same reasoning as in (4.351)-(4.354), the capacity under independent fading for $n_T < n_R$ is found to be

$$\begin{aligned} C_{avg}^{\mathbf{\Sigma}=\mathbf{I}/\,n_T<n_R} = H'_{n_R,n_T} \sum_{k=1}^{n_T} \sum_{q} \epsilon(q) \left[\prod_{l=n_T+1}^{n_R} (-1)^{q_l-l}(n_R-l)_{n_R-q_l}\right] \det[(n_R+n_T-j-q_i)!\cdot \\ \cdot \left[\frac{e^{n_T/\gamma}}{\ln 2} \sum_{p=0}^{n_R+n_T-j-q_i} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{n_T}{\gamma}\right)^{p-l} \Gamma\left[l, \frac{n_T}{\gamma}\right]\right]^{\delta_{kj}}\Bigg]_{n_T \times n_T} \end{aligned} \tag{4.365}$$

Evaluating (4.293) and (4.365) for several parameter choices and comparing the capacity values with the results from [71], an exact match was observed, which clearly validates the procedure employed to derive the capacity formulas.

A final remark is required. Despite the fact that the configurations $n_T \geq n_R$ and $n_T < n_R$ have been treated independently in this thesis, the author has proved (see [72]) that it is indeed possible to perform a unified derivation that contemplates both situations. The universal complex Wishart density function obtained therein provides the necessary momentum for the derivation of a set of all-embracing analytic formulas for the average capacity of the MIMO wireless channel under receiver-sided correlated fading.

4.6.5 CAPACITY UNDER TRANSMITTER-SIDED CORRELATED FADING

When correlation exists only between signals departing from different transmitter antennas (i.e. $\mathbf{\Sigma}_{\text{R}} = \mathbf{I}_{n_R}$), which may happen if the transmitter antennas are close to one another, then the capacity derivation must follow from the pdf in (4.311). The situation is a little bit different because of the dependence on the quadratic form $\mathbf{H}^H\mathbf{H}$, which is different from (4.308). However, using the properties of block determinants, we may write

$$\det\left[\begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{A}^H & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_m + \mathbf{A}\mathbf{A}^H & \mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ -\mathbf{A}^H & \mathbf{I}_n \end{pmatrix}\right] = \det(\mathbf{I}_m + \mathbf{A}\mathbf{A}^H) = \det(\mathbf{I}_n + \mathbf{A}^H\mathbf{A}) \tag{4.366}$$

where $\mathbf{A}$ is $m \times n$. Hence, assuming that no channel state information is available at the transmitter, (4.308) is equivalent to

$$C_{avg} = E_{\mathbf{H}}\left[\log_2 \det\left(\mathbf{I}_{n_T} + \frac{\gamma}{n_T}\mathbf{H}^H\mathbf{H}\right)\right] \tag{4.367}$$

where we have to note that, for simplicity, the optimization with respect to the input covariance matrix is being avoided. For $n_T \leq n_R$, the quadratic form $\mathbf{H}^H\mathbf{H}$ is full-rank, and therefore if we write $\mathbf{G} = \mathbf{H}^H$, then $\mathbf{G}$ has n_T linearly independent columns and the average

capacity is evaluated by eigendecomposing $\mathbf{G}\mathbf{G}^H$. As a result, we simply have to exchange $\mathbf{\Sigma}_{\mathrm{R}}$ by $\mathbf{\Sigma}_{\mathrm{T}}$, n_R by n_T, and finally make $\gamma \rightarrow (n_R / n_T)\gamma$ in (4.286) to obtain the desired capacity expression

$$C_{avg}^{\mathbf{\Sigma}_{\mathrm{T}}/n_T \leq n_R} = H_{n_T,n_R}(\boldsymbol{\upsilon}_{\mathrm{T}}) \sum_{k=1}^{n_T} \det\left[\frac{(n_R - j)!}{\upsilon_i^{-(n_R-j+1)}} \cdot \right.$$
$$\left. \cdot \left[\frac{e^{n_T/(\upsilon_i\gamma)}}{\ln 2} \sum_{p=0}^{n_R-j} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{n_T}{\upsilon_i\gamma}\right)^{p-l} \Gamma\left[l, \frac{n_T}{\upsilon_i\gamma}\right] \right]^{\delta_{ki}} \right], \tag{4.368}$$

where $\boldsymbol{\upsilon}_{\mathrm{T}}$ is the vector of eigenvalues of $\mathbf{\Sigma}_{\mathrm{T}}$. Finally, when $n_T > n_R$ and there is no fading correlation at the receiver, matrix $\mathbf{G}\mathbf{G}^H$ is rank-deficient. This was exactly the condition aimed to be solved by this section, and for which the solution was found to be (4.354). Therefore, we again exchange $\mathbf{\Sigma}_{\mathrm{R}}$ by $\mathbf{\Sigma}_{\mathrm{T}}$, n_R by n_T, and finally make $\gamma \rightarrow (n_R / n_T)\gamma$ in (4.354) to obtain

$$C_{avg}^{\mathbf{\Sigma}_{\mathrm{T}}/n_T > n_R} = H'_{n_T,n_R}(\boldsymbol{\upsilon}_{\mathrm{T}}) \sum_{k=1}^{n_R} \sum_{q}^{n_T} \epsilon(q) \left[\prod_{l=n_R+1}^{n_T} \left(-\upsilon_{q_l}^{-1}\right)^{n_T-l} \right] \det\left[\frac{(n_R - j)!}{\upsilon_{q_i}^{-(n_R-j+1)}} \cdot \right.$$
$$\left. \cdot \left[\frac{e^{n_T/(\upsilon_{q_i}\gamma)}}{\ln 2} \sum_{p=0}^{n_R-j} \frac{1}{p!} \sum_{l=0}^{p} \binom{p}{l} \left(-\frac{n_T}{\upsilon_{q_i}\gamma}\right)^{p-l} \Gamma\left[l, \frac{n_T}{\upsilon_{q_i}\gamma}\right] \right]^{\delta_{kj}} \right]_{n_R \times n_R} \tag{4.369}$$

Equations (4.368) and (4.369) are the consequence of capacity duality in single-sided correlated MIMO systems, that is, the fact that the capacity in one direction for $\gamma = \gamma_0 N$ equals the capacity in the other direction for $\gamma = \gamma_0 M$, where N and M are the respective numbers of transmitting antennas.

4.6.6 Numerical Results and Discussion

We shall consider several distinct situations in the subsequent capacity analysis. Nevertheless, at the correlated side of the communication link, the analysis will be always confined to linear antenna arrays of fixed maximum length $d_{\max} = 1.2\lambda$, and a maximum of eight antennas. Moreover, the antennas will be considered evenly distributed within $d_{\max}$. When isotropic scattering is assumed, the entries of the covariance matrix are given by

$$\left[\mathbf{\Sigma}_{\mathrm{R/T}}\right]_{ir} = J_0(\beta(i - r)\, d_{\max} / n_{R/T}) \tag{4.370}$$

where $J_0(\cdot)$ the zeroth-order Bessel function of the first kind, $\beta = 2\pi / \lambda$ is the phase constant, and $n_{R/T}$ is the number of antennas at the correlated side of the link.

First, we consider the downlink capacity of a typical cellular mobile communication system, and assume that correlation is present only at the mobile unit. The scattering is isotropic around the MU, mutual coupling effects are ignored, and the received SNR is fixed at $\gamma = 5$ dB. Figure 4.13 shows a 3-dimensional capacity plot as a function of the number of transmitter and receiver antennas, which includes the scenario of two-sided independent fading for comparison. The capacity values were obtained using expressions (4.286), (4.354), (4.293) and (4.365). We conclude that the capacity under receiver-sided correlated fading always lags behind the capacity under two-sided independent fading, and this condition becomes more pronounced as the number of antennas increases. Given that

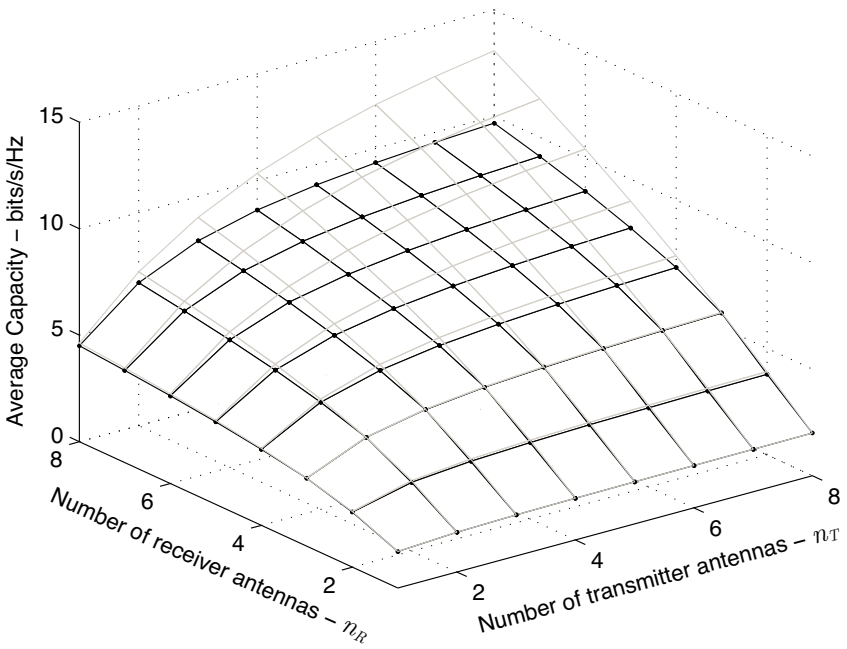

FIGURE 4.13 Average MIMO channel capacity as a function of n_R and n_T, for $\gamma = 5$ dB. The black line is for receiver-sided correlated fading with isotropic scattering, and the gray line is for two-sided independent fading.

the capacity is limited as $n_T \to \infty$ (4.35), and due to the imposed $d_{\max}$ at the receiver, adding more antennas will eventually become negligible, since the capacity increase will tend to stagnate. A different plot is shown in Figure 4.14 which confirms this behaviour. We

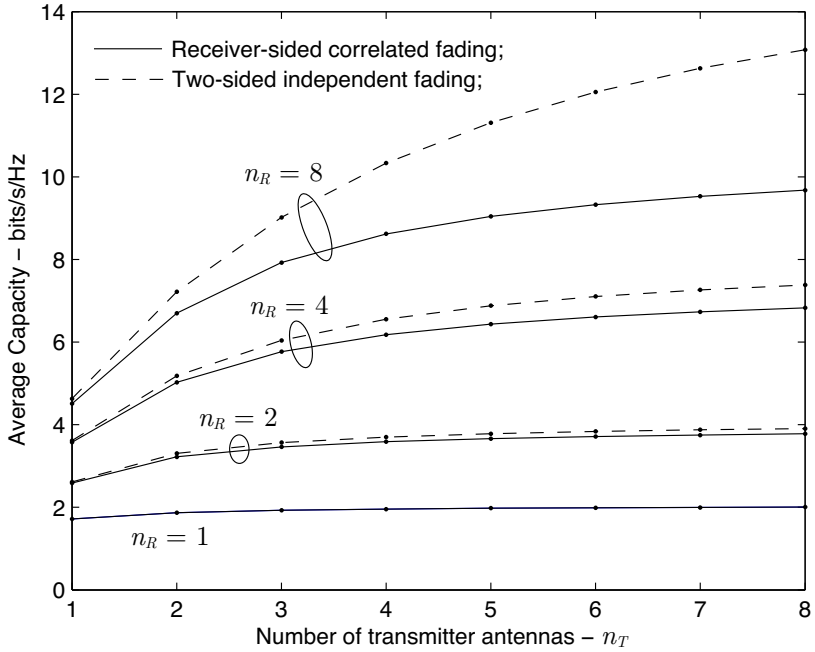

FIGURE 4.14 Average MIMO channel capacity as a function of n_T, for several values of n_R, and $\gamma = 5$ dB. For receiver-sided correlated fading, isotropic scattering is assumed.

observe that in the correlated scenarios the most significant capacity increase occurs while $n_T < n_R$, and that substantial stagnation occurs otherwise. Moreover, we observe that to attain an equivalent capacity increase as by switching from $n_R = 2$ to $n_R = 4$, we need to

switch from $n_R = 4$ to $n_R = 8$. This means that under receiver-sided correlated fading with fixed $d_{\max}$, adding more receiver antennas has a diminished impact on the capacity as compared to two-sided independent fading. In fact, the higher is n_R, the lesser is the impact of increasing the former, and one should expect this impact to eventually become irrelevant.

To analyse the effects of transmitter-sided correlation, we now consider the uplink of the cellular communication system, and plot the capacity as a function of the number of mobile unit antennas n_{MU}, and the number of base station antennas n_{BS}. This is shown in Figure 4.15, where the downlink capacity is also included for comparison. As expected,

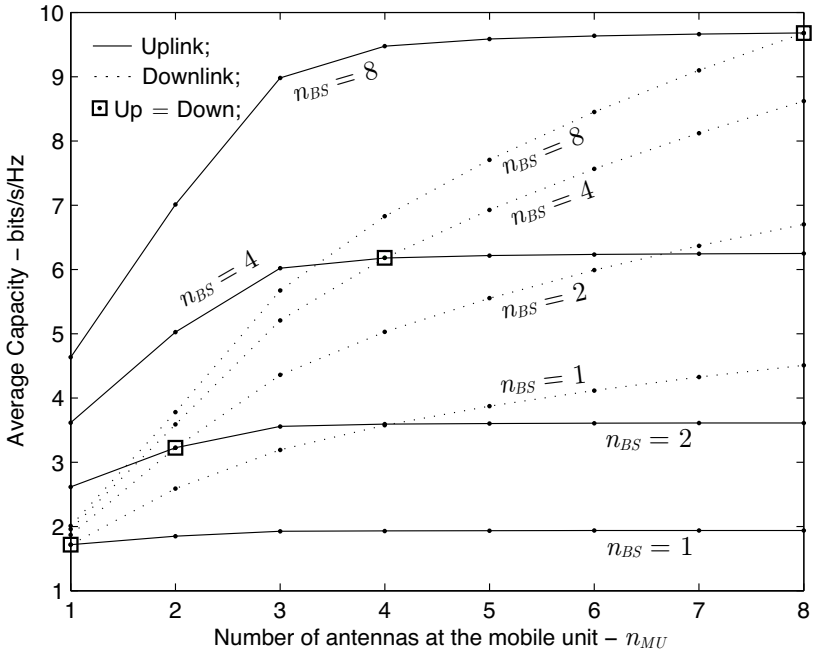

FIGURE 4.15 Uplink and downlink average MIMO channel capacities as a function of n_{MU}, for several values of n_{BS}, and $\gamma = 5$ dB. Correlated fading and isotropic scattering are assumed for the mobile unit only.

we notice that the uplink capacity is higher than the downlink capacity if $n_{BS} > n_{MU}$, and the opposite if $n_{BS} < n_{MU}$, while an exact match occurs if $n_{BS} = n_{MU}$. Naturally, this only happens because γ is assumed equal on both sides. We also note a peculiar behaviour of the uplink capacity, as it is virtually non-increasing for $n_{MU} > 3$. This fact makes it more difficult to increase the uplink capacity by increasing n_{MU} than to increase the downlink capacity by increasing n_{BS}. If one wishes to obtain uplink capacity enlargements, either one improves γ, or one has to increase n_{BS}. Moreover, since the downlink capacity is most often the more demanding, one may usually require that $n_{MU} \geq n_{BS}$, where n_{BS} approximately fixes the uplink capacity, and n_{MU} varies so as to achieve the desired downlink capacity.

To conclude, we study the implications of receiver-sided correlation with non-isotropic scattering on the average capacity of the MIMO channel. Non-isotropy is worth considering, since it may frequently apply to the propagation conditions of some sub-urban/urban environments. To that end, we use the correlation model proposed in [73], for

which

$$[\mathbf{\Sigma}_{\mathrm{R}}]_{ir} = I_0\left(\sqrt{(\kappa^2 - \beta^2 d_{ir}^2 + j2\beta\kappa d_{ir}\cos\phi_c}\right)/I_0(\kappa) \tag{4.371}$$

where $\phi_c \in [-\pi, \pi)$ is the mean angle-of-arrival (AOA) relative to the direction of motion, $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind, $\kappa \in [0, \infty[$ controls the width of the AOA, and for the array under consideration we have $d_{ir} = (i-r)\,d_{\max}/n_R$. Equation (4.365) is easily obtained from (4.295) and (4.297), by setting $\theta_{ir} = \pi$ (i.e. the mobile unit moves in parallel with the axis of the linear array). Average capacity curves in the configuration $n_T = n_R$ are shown in Figure 4.16. We observe that non-isotropic scat-

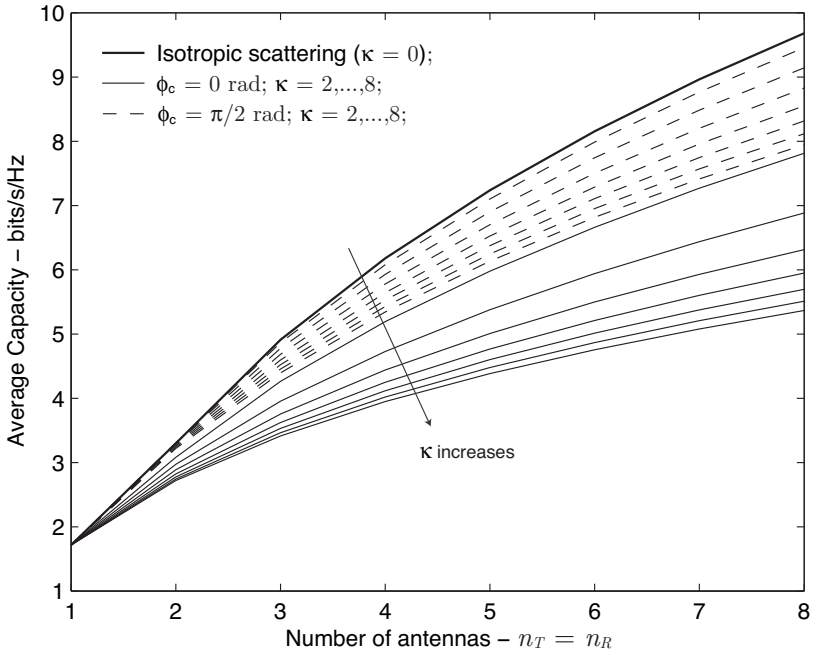

FIGURE 4.16 Average MIMO channel capacity as a function of $n_T = n_R$, for $\gamma = 5$ dB, receiver-sided correlated fading, and nonisotropic scattering modelled by (4.297) with $\phi_c = \{0, \pi/2\}$ and $\kappa = \{2, \ldots, 8\}$.

tering has the effect of decreasing the average capacity, and this feature is more pronounced for $\phi_c = 0$ rad (or equivalently $\phi_c = \pi$), that is, when the receiver moves toward or away from the mean direction of multipath arrival. This condition might have been expected since the multipath components that arrive nearly parallel to the axis of the linear array (which, in this case, coincides with the direction of motion), are the ones that most contribute to the correlation between the antennas, and hence most likely contribute to reduce the capacity. From Figure 4.16 we also infer a substantial capacity reduction as the width of the AOA narrows, i.e. as κ increases, which is consistent with the fact that the antennas become more and more correlated.

4.6.7 CONCLUSION

Section 4.6 has presented the derivation of the closed-form expression for the average capacity of correlated MIMO channels when the number of receiver antennas exceeds the number of transmitter antennas. In addition, closed-form formulas for the capacity under transmitter-sided correlation and two-sided independent fading were obtained. Numerical

results quantified a significant capacity reduction when correlation was present, and predominantly when the scattering was not isotropic. In particular, it was shown that capacity stagnates very easily when transmitter-sided correlation is present, and that fixed-length correlated antenna arrays greatly refrain the unbounded capacity potential of independent fading scenarios. It was also observed that in a typical cellular mobile communication system, the most logical configuration option (in what concerns capacity) is to set the number of mobile unit antennas larger or equal than the number of base station antennas, irrespective of any length constraints and/or correlation applicable to the antenna arrays of mobile units.

CHAPTER 5

OVERSPREAD TRANSMISSION OVER WIRELESS LTV MIMO SYSTEMS

5.1 *Introduction*

In common urban or sub-urban propagation environments, the typical delay spread of the wireless mobile channel is much smaller than its coherence time, which means that the channel may be viewed as almost time-invariant during consecutive time-frames spanning less than the coherence time. Since the coherence time is in the order of milliseconds, one usual approach for designing the receiver is to first consider the channel as time-invariant, devise an equalizer to the received signal (e.g. a linear or decision-feedback equalizer) and then use an adaptive technique (e.g. least mean square (LMS), recursive least squares (RLS), or Kalman) to track the channel variations and update the equalizer parameters (viz. [77], [78], [79],[80], and [81]). This procedure finds its most important application in frequency-selective slowly fading channels, and, under mild conditions, channel tracking may be replaced by blind or decision directed techniques. There are, however, several drawbacks to the adaptive process, namely: 1) it is not a smooth process, as abrupt changes between consecutive intervals are required [78]; 2) no *pilot-aided transmission* (PAT) is exploited; 3) it may be unstable under certain channel variations [82]; 4) it does not take advantage of the time variations in the channel to improve the receiver performance. Introducing "vertical Bell Labs layered space-time" (V-BLAST) [83] architecture into the picture also has its own problems [84], [85].

Recent research has shown that not incorporating the channel variations into the receiver design penalizes its performance. It has been shown that the RAKE receiver is inefficient when the channel is time-varying [86], and attempts have been made to correct the problem by exploiting Doppler diversity [87], [88], [89] and [90], using techniques such as time-frequency (canonical) representations (TFRs) and the basis-expansion model (BEM). These developments encourage the study of wireless channels in a more general setting, as the inherent deficiencies of receiver implementation are the consequence of

considering the channel as interval-wise invariant. The primary barrier for such a study is the analysis of linear time-varying (LTV) filters in a system-wide perspective, which is much more involved than analysing conventional linear time-invariant systems. Designing the receiver structures when all its units are time-varying may seem an overwhelming task, at least without a tractable mathematical description, but if it proves to be possible then it will lead to results of the utmost importance and generality. Aiming at a possible solution, this chapter will dwell into the more practical aspects of MIMO transmission, reception and detection. The idea is to study MIMO systems in a more generic framework, with no constrictive assumptions being presumed from the start except those of a time-varying wireless propagation medium and accurate channel estimation at the receiver (viz. [91], [92], [93] and [94]). This will allow for more general conclusions that may me made more specific by simple deduction.

The wireless channel between each pair of antennas will be considered linear and time-varying. Also, depending on the coherence-time $(\Delta t)_c \approx 1/2\pi f_D$, coherence-bandwidth $(\Delta f)_c \approx 1/2\pi\sigma_\tau$, signal bandwidth W, and a signalling interval T, it can be of two types: 1) *underspread* if

$$[(\Delta f)_c T][W(\Delta t)_c] > 1 \;\sim\; \sigma_\tau f_D < (T/2\pi)(W/2\pi) = (2\pi)^{-2}WT \tag{5.1}$$

or 2) *overspread* if

$$[(\Delta f)_c T][W(\Delta t)_c] < 1 \;\sim\; \sigma_\tau f_D > (T/2\pi)(W/2\pi) = (2\pi)^{-2}WT \tag{5.2}$$

When the signal bandwidth is the reciprocal of the signalling interval (half roll-off signalling), the underspread and overspread conditions are simply $\sigma_\tau f_D < (2\pi)^{-2}$ and $\sigma_\tau f_D > (2\pi)^{-2}$, respectively. Both $(\Delta t)_c$ and $(\Delta f)_c$ depend on the stochastic properties of the wireless channel, which in turn are a function of the relative distance and motion between transmitter and receiver, and also of the propagation conditions within the wireless medium. If one is to devise a slowly fading, frequency non-selective channel, by concurrently choosing a signal bandwidth $W < (\Delta f)_c$ and a signalling interval $T < (\Delta t)_c$, then a necessary condition is that it be underspread. In an underspread channel (such as the typical mobile propagation channel), frequency selectivity ($W > (\Delta f)_c$) induces slow fading (time-flatness, $T < (\Delta t)_c$), and time selectivity (fast-fading, $T > (\Delta t)_c$) induces a frequency-flat response ($W < (\Delta f)_c$). If we are forced to have selectivity in both time and frequency then the channel is necessarily overspread.

The typical wideband channel (e.g. CDMA2000, W-CDMA) is overspread in delay ($\sigma_\tau \gg 1/W \rightarrow W \gg (\Delta f)_c$), hence frequency-selective, yet slowly fading ($T \ll (\Delta t)_c$). On the contrary, to capitalize on capacity, some multicarrier systems (e.g. OFDM) with very low carrier separations (not uncommonly lower than 10 kHz, viz. 3GPP LTE eMBMS, 1 kHz in DAB, a.k.a. Eureka 147) may be frequency-flat ($W < (\Delta f)_c \tilde{\rightarrow} \sigma_\tau < T/2\pi$), hence delay-underspread, but their signalling rate may be too low to permit a slow fading behavior. These narrowband sub-channels may be fast-fading, hence Doppler-overspread ($T > (\Delta t)_c \tilde{\rightarrow} f_D > W/2\pi$), which means that the induced Doppler spread is not negligible from the receiver perspective, as inter-carrier interference (ICI) will be introduced. Channels such as the ones invoked self-dictate the need for a thorough analysis of transmission overspreading, let alone their blending with MIMO technology. Within the gener-

ic framework of this chapter, all types of wireless channels will be studied without preliminary distinction.

5.2 Input/Output Model

The time-varying channel impulse response between the j-th input and the i-th output is denoted by $c_{ij}(\tau,t)$, and it is conveniently chosen as the response at time t of an impulse applied at time $t-\tau$. It can legitimately be thought of as the impulse response at a specified time t, since the time-scale at which the wireless channel varies is typically much longer than the delay spread of the channel. For a time-varying multipath wireless channel, the channel impulse response is given by

$$c_{ij}(\tau,t)=\sum_{k=1}^{N_{ij}(t)}\alpha_k^{(ij)}(t)\delta(\tau-\tau_k^{(ij)}(t)), \tag{5.3}$$

accounting for the attenuations $\alpha_k^{(ij)}(t)$ and delays $\tau_k^{(ij)}(t)$ of $N_{ij}(t)$ multipath replicas. Due to the random nature of the wireless medium, we must interpret the time-varying parameters of (5.3) as sample functions of statistically independent stochastic processes.

Denoting the number of inputs by n_T, the transmitted signal from the j-th input by $s_j(t)$, and the additive noise at the i-th output by $n_i(t)$, the signal received by the i-th output is given by the superposition

$$r_i(t)=\sum_{j=1}^{n_T}\int_{-\infty}^{\infty}c_{ij}(\tau,t)s_j(t-\tau)d\tau+n_i(t). \tag{5.4}$$

In a more compact vector-matrix notation, we can write

$$\mathbf{r}(t)=\int_{-\infty}^{\infty}\mathbf{C}(\tau,t)\mathbf{s}(t-\tau)d\tau+\mathbf{n}(t), \tag{5.5}$$

where we define the column vector-valued signals $\mathbf{r}(t)=[r_i(t)]_{n_R\times 1}$, $\mathbf{n}(t)=[n_i(t)]_{n_R\times 1}$, $\mathbf{s}(t)=[s_j(t)]_{n_T\times 1}$, and the matrix-valued channel response $\mathbf{C}(\tau,t)=[c_{ij}(\tau,t)]_{n_R\times n_T}$ (n_R being the number of outputs).

5.2.1 Baseband MIMO Channel Model

In the model of (5.5), $\mathbf{s}(t)$ is the transmitted vector-valued signal which, in a typical wireless application, is limited to an approximate passband $[f_c-B/2, f_c+B/2]$ of bandwidth $B=2W$ and center frequency f_c. The "up-conversion"/"down-conversion" to/from the carrier frequency f_c are performed solely to permit transmission over the wireless medium, and are always the last and first stages of transmitter and receiver operation, respectively. Since most of the other processing stages are performed at the baseband, it becomes important to derive an equivalent baseband model from (5.5).

Defining the Fourier transform of $\mathbf{s}(t)$ as $\mathbf{s}(f)=[S_j(f)]$ and noting that, since $\mathbf{s}(t)$ is

real-valued, the Fourier transform is conjugate symmetric (i.e. $\mathbf{s}(f) = \mathbf{s}^*(-f)$), we have

$$\begin{aligned}\mathbf{s}(t) &= \int_{-\infty}^{\infty} \mathbf{s}(f)e^{j2\pi ft}df = \int_{f_c-B/2}^{f_c+B/2} \Big[\mathbf{s}(f)e^{j2\pi ft} + \mathbf{s}(-f)e^{-j2\pi ft}\Big]df \\ &= \int_{f_c-B/2}^{f_c+B/2} 2\Re\mathrm{e}\Big\{\mathbf{s}(f)e^{j2\pi ft}\Big\}df = \Re\mathrm{e}\left\{\left[\int_{-B/2}^{+B/2} 2\mathbf{s}(f+f_c)e^{j2\pi ft}df\right]e^{j2\pi f_c t}\right\}.\end{aligned} \tag{5.6}$$

If the baseband signal is denoted by $\underline{\mathbf{s}}(t)$ and its Fourier transform by $\underline{\mathbf{s}}(f)$, then the latter may be defined as two times the positive frequency part of $\mathbf{s}(f)$ shifted to the origin, that is

$$\underline{\mathbf{s}}(f) = \begin{cases} 2\mathbf{s}(f+f_c), & |f| \le B/2, \\ \mathbf{0}, & |f| > B/2, \end{cases} \tag{5.7}$$

which substituted into (5.6) yields

$$\mathbf{s}(t) = \Re\mathrm{e}\left\{\left[\int_{-B/2}^{+B/2} \underline{\mathbf{s}}(f)e^{j2\pi ft}df\right]e^{j2\pi f_c t}\right\} = \Re\mathrm{e}\Big\{\underline{\mathbf{s}}(t)e^{j2\pi f_c t}\Big\}. \tag{5.8}$$

We should have in mind that due to the conjugate symmetry of $\mathbf{s}(f)$, no loss of information occurs if one chooses for the purpose of a signal model the mathematical convenience of $\underline{\mathbf{s}}(t)$ relative to $\mathbf{s}(t)$. In fact, as shown in Figure 5.1, $\mathbf{s}(t)$ is easily obtained from $\underline{\mathbf{s}}(t)$ using carrier multiplication followed by real part extraction, and $\underline{\mathbf{s}}(t)$ is recovered from $\mathbf{s}(t)$ using carrier multiplication followed by ideal lowpass filtering. Although, for simplicity, these operations are best seen in the complex domain, the real domain equivalent has a similar construction.

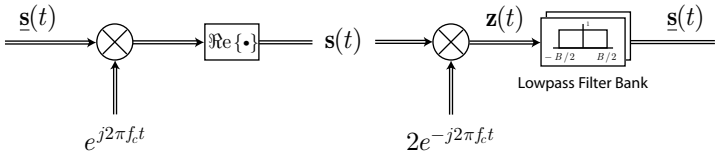

FIGURE 5.1 Illustration of up-conversion and down-conversion operations in the complex domain.

The communication designer can then concentrate its energies on defining the baseband signal $\underline{\mathbf{s}}(t)$ that should be used for transmission.

To unravel the baseband model for the wireless MIMO channel we start by taking the Fourier transform of $\mathbf{C}(\tau,t)$ relative to the response delay τ and denote it by $\mathbf{C}(f,t)$. Clearly

$$\mathbf{C}(\tau,t) = \int_{-\infty}^{\infty} \mathbf{C}(f,t)e^{j2\pi f\tau}df. \tag{5.9}$$

Now, if $\mathbf{u}(t)$ is the output of a noiseless channel, it is related to the input $\mathbf{s}(t)$ by

$$\mathbf{u}(t) = \int_{-\infty}^{\infty} \mathbf{C}(\tau,t)\mathbf{s}(t-\tau)d\tau = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{C}(f,t)e^{j2\pi f\tau}\mathbf{s}(t-\tau)df d\tau, \tag{5.10}$$

which, interchanging the order of integration, easily leads to

$$\begin{aligned}\mathbf{u}(t) &= \int_{-\infty}^{\infty} \mathbf{C}(f,t)\left[\int_{-\infty}^{\infty} e^{j2\pi f\tau}\mathbf{s}(t-\tau)d\tau\right]df \\ &= \int_{-\infty}^{\infty} \mathbf{C}(f,t)\mathbf{s}(f)e^{j2\pi ft}df.\end{aligned} \tag{5.11}$$

It now becomes clear that $\mathbf{u}(t)$ is the inverse Fourier transform of $\mathbf{C}(f,t)\mathbf{s}(f)$, which is a matrix-vector product and, by virtue of the band-limited nature of $\mathbf{s}(f)$, is also band-limited, except that the passband $[f_c - B_u/2, f_c + B_u/2]$ will be slightly larger. This property also follows from the Fourier transform of (5.11), i.e. the convolution

$$\begin{aligned}\mathbf{u}(f) &= \int_{-\infty}^{\infty} \mathbf{C}(f-\nu,\nu)\mathbf{s}(f-\nu)d\nu \\ &= \mathbf{C}(f-\nu,\nu) \underset{\nu}{\otimes} \mathbf{s}(f)\end{aligned} \tag{5.12}$$

where the variable ν is associated with the rate of change of the wireless channel and is called the *Doppler frequency*. One infers that, despite the fact that MIMO channel responses $\mathbf{C}(f,t)$ or $\mathbf{C}(f,\nu)$ are not explicitly band-limited in the f variable, the only frequencies that are required to compute the output $\mathbf{u}(t)$ or $\mathbf{u}(f)$ lie in the passband $[f_c - B/2, f_c + B/2]$, and thus it is possible to define a baseband representation for $\mathbf{C}(f,t)$ that is very similar to (5.7). More specifically, we are able to relax the band-limited condition by letting $B \to \infty$, which leads to

$$\underline{\mathbf{C}}(f,t) = 2\mathbf{C}(f+f_c,t), \tag{5.13}$$

or equivalently

$$\underline{\mathbf{C}}(\tau,t) = 2\mathbf{C}(\tau,t)e^{-j2\pi f_c\tau}. \tag{5.14}$$

It now follows from (5.10) that the noiseless output $\mathbf{u}(t)$ is given by

$$\begin{aligned}\mathbf{u}(t) &= \int_{-\infty}^{\infty} \mathbf{C}(\tau,t)\mathbf{s}(t-\tau)d\tau = \int_{-\infty}^{\infty} \mathbf{C}(\tau,t)\Re\mathrm{e}\left\{\underline{\mathbf{s}}(t-\tau)e^{j2\pi f_c(t-\tau)}\right\}d\tau \\ &= \Re\mathrm{e}\left\{\left[\int_{-\infty}^{\infty} \mathbf{C}(\tau,t)\underline{\mathbf{s}}(t-\tau)e^{-j2\pi f_c\tau}d\tau\right]e^{j2\pi f_c t}\right\}\end{aligned} \tag{5.15}$$

Comparing with (5.8), we find that (5.15) provides the relation between $\mathbf{u}(t)$ and $\underline{\mathbf{u}}(t)$, which using (5.14) is found to be

$$\begin{aligned}\underline{\mathbf{u}}(t) &= \int_{-\infty}^{\infty} \mathbf{C}(\tau,t)\underline{\mathbf{s}}(t-\tau)e^{-j2\pi f_c\tau}d\tau \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \underline{\mathbf{C}}(\tau,t)\underline{\mathbf{s}}(t-\tau)d\tau,\end{aligned} \tag{5.16}$$

which is simply a convolution scaled by $1/2$. From the Fourier transform

$$\begin{aligned}\underline{\mathbf{u}}(f) &= \frac{1}{2}\int_{-\infty}^{\infty} \underline{\mathbf{C}}(f-\nu,\nu)\underline{\mathbf{s}}(f-\nu)d\nu \\ &= \frac{1}{2}\underline{\mathbf{C}}(f-\nu,\nu) \underset{\nu}{\otimes} \underline{\mathbf{s}}(f)\end{aligned} \tag{5.17}$$

we are led to conclude that, due to the convolution in the variable ν the bandwidth of $\underline{\mathbf{u}}(t)$ will be somewhat larger than that of $\underline{\mathbf{s}}(t)$, which is an obvious consequence of the Doppler spread of the channel.

Continuing, the complete (noisy) baseband model for the MIMO wireless channel is

thence given by

$$\begin{aligned}\underline{\mathbf{r}}(t) &= \frac{1}{2}\int_{-\infty}^{\infty} \underline{\mathbf{C}}(\tau,t)\underline{\mathbf{s}}(t-\tau)d\tau + \underline{\mathbf{n}}(t) \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \underline{\mathbf{C}}(f,t)\underline{\mathbf{s}}(f)e^{j2\pi ft}df + \underline{\mathbf{n}}(t).\end{aligned} \tag{5.18}$$

Equation (5.18) is the final channel model because it accounts for the noise that is inevitably present at the channel outputs. In the context of the current presentation, $\mathbf{n}(t)$ is modelled as a wide-sense stationary (WSS) vector random process with zero mean, autocorrelation

$$\mathbf{R}_{\mathbf{nn}}(\tau) = E\big[\mathbf{n}(t+\tau)\mathbf{n}^T(t)\big] = \frac{N_0}{2}\delta(\tau)\mathbf{I}, \tag{5.19}$$

and power density spectrum

$$\mathbf{S}_{\mathbf{nn}}(f) = \frac{N_0}{2}\mathbf{I}. \tag{5.20}$$

In other words, the noise process is both spatially and temporally *white*, hence *white*. Since (in theory) the noise process spans the entire spectrum, one must be more careful on characterizing the baseband representation $\underline{\mathbf{n}}(t)$ of $\mathbf{n}(t)$. To that end, we shall firstly require the following result from the theory of stochastic processes.

LINEAR FILTERING OF VECTOR STOCHASTIC PROCESSES: If $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are, respectively, input and output (complex) WSS vector random processes of a linear time-invariant (LTI) system with transfer function $\mathbf{L}(f)$, then the relation between input and output power density spectrums is

$$\mathbf{S}_{\mathbf{yy}}(f) = \mathbf{L}(f)\mathbf{S}_{\mathbf{xx}}(f)\mathbf{L}^H(f), \tag{5.21}$$

where $\mathbf{L}^H(f)$ is the conjugate transpose of $\mathbf{L}(f)$.This result is easily derived from the correlation functions and the convolution integral. We have

$$\mathbf{y}(t) = \int_{-\infty}^{\infty} \mathbf{L}(\alpha)\mathbf{x}(t-\alpha)d\alpha, \tag{5.22}$$

and thus

$$\begin{aligned}\mathbf{R}_{\mathbf{yy}}(\tau) &= \frac{1}{2}E[\mathbf{y}(t+\tau)\mathbf{y}^H(t)] = \frac{1}{2}\int_{-\infty}^{\infty} \mathbf{L}(\alpha)E[\mathbf{x}(t-\alpha+\tau)\mathbf{y}^H(t)]d\tau \\ &= \int_{-\infty}^{\infty} \mathbf{L}(\alpha)\mathbf{R}_{\mathbf{xy}}(\tau-\alpha)d\tau = \mathbf{L}(\tau) * \mathbf{R}_{\mathbf{xy}}(\tau),\end{aligned} \tag{5.23}$$

where $\mathbf{R}_{\mathbf{xy}}(\tau)$ is the matrix-valued cross-correlation function between input and output. $\mathbf{R}_{\mathbf{xy}}(\tau)$ is related to $\mathbf{R}_{\mathbf{xx}}(\tau)$ by

$$\begin{aligned}\mathbf{R}_{\mathbf{xy}}(\tau) &= \frac{1}{2}E[\mathbf{x}(t+\tau)\mathbf{y}^H(t)] = \frac{1}{2}\int_{-\infty}^{\infty} E[\mathbf{x}(t+\tau)\mathbf{x}^H(t-\alpha)]\mathbf{L}^H(\alpha)d\alpha \\ &= \int_{-\infty}^{\infty} \mathbf{R}_{\mathbf{xx}}(\tau+\alpha)\mathbf{L}^H(\alpha)d\alpha = \mathbf{R}_{\mathbf{xx}}(\tau) * \mathbf{L}^H(-\tau).\end{aligned} \tag{5.24}$$

From (5.23) and (5.24) it follows that $\mathbf{R}_{\mathbf{yy}}(\tau) = \mathbf{L}(\tau) * \mathbf{R}_{\mathbf{xx}}(\tau) * \mathbf{L}^H(-\tau)$, and taking the Fou-

rier transform reveals (5.21).

Recalling that, as illustrated in Figure 5.1, $\underline{\mathbf{n}}(t)$ is determined from $\mathbf{n}(t)$ by down-conversion, we first modulate $\mathbf{n}(t)$ by the carrier $2e^{-j2\pi f_c t}$ and obtain a vector-valued signal $\mathbf{z}(t) = 2\mathbf{n}(t)e^{-j2\pi f_c t}$ with autocorrelation

$$\begin{aligned}\mathbf{R}_{\mathbf{zz}}(\tau) &= \frac{1}{2}E[\mathbf{z}(t+\tau)\mathbf{z}^H(t)] = 2E\big[\mathbf{n}(t+\tau)\mathbf{n}^T(t)\big]e^{-j2\pi f_c\tau} \\ &= 2\mathbf{R}_{\mathbf{nn}}(\tau)e^{-j2\pi f_c\tau}.\end{aligned} \tag{5.25}$$

Taking the Fourier transform, the power density spectrum of $\mathbf{z}(t)$ is

$$\mathbf{S}_{\mathbf{zz}}(f) = 2\mathbf{S}_{\mathbf{nn}}(f+f_c) = N_0\mathbf{I}, \tag{5.26}$$

which, for an arbitrary noise bandwidth B_N, corresponds to a translation of the passband $[f_c - B_N/2,\, f_c + B_N/2]$ to the baseband $[-B_N/2,\, +B_N/2]$. Due to the *white* nature of $\mathbf{n}(t)$, this spectral translation has no effect on the noise process, but the same does not happen in the lowpass filtering stage of the down-conversion. In effect, with the aid of (5.21) and the definition of the ideal lowpass filter

$$\mathbf{L}(f) = \begin{cases}\mathbf{I}, & |f| \le B/2, \\ \mathbf{0}, & |f| > B/2,\end{cases} \tag{5.27}$$

we would derive the power density spectrum of $\underline{\mathbf{n}}(t)$ as $\mathbf{S}_{\underline{\mathbf{nn}}}(f) = N_0\mathbf{I},\ |f| \le B/2$, and $\mathbf{0}$ otherwise, which is clearly a lowpass filtered version of the white noise. However, since $\mathbf{n}(t)$ is (in theory) not band-limited in spectrum, we are prompted to choose a baseband noise model with $B \to \infty$ that leads to the following characterization of $\underline{\mathbf{n}}(t)$:

$$\begin{aligned}&\underline{\mathbf{n}}(t) = 2\mathbf{n}(t)e^{-j2\pi f_c t}, \\ &\mathbf{R}_{\underline{\mathbf{nn}}}(\tau) = 2\mathbf{R}_{\mathbf{nn}}(\tau)e^{-j2\pi f_c\tau} = N_0\delta(\tau)\mathbf{I}, \text{and} \\ &\mathbf{S}_{\underline{\mathbf{nn}}}(f) = 2\mathbf{S}_{\mathbf{nn}}(f+f_c) = N_0\mathbf{I}.\end{aligned} \tag{5.28}$$

In short, we characterize $\underline{\mathbf{n}}(t)$ as a WSS process with a white power density spectrum that is two times that of $\mathbf{n}(t)$. It must be ascertained that, even though white noise is not physically realizable, this theoretical characterization is neither unrealistic nor fallacious from a design point of view. In fact, before down-conversion, a receiver is expected to pass the received signal through a bandpass filter with passband much smaller than the center frequency, but wider than the signal bandwidth. As a consequence the noise will have indeed a baseband representation, one that we interpret as spectrum-flat for mere mathematical simplicity. In addition, in the optimum receiver design, the spectral components of noise outside the received signal band will always have to be removed before detection, and hence the "white" assumption does not adulterate the design process.

Figure 5.2 depicts how the passband model of (5.5) reduces to the baseband model of (5.18). Based on this equivalence, the study that will follow will focus on the baseband channel model of the wireless MIMO system.

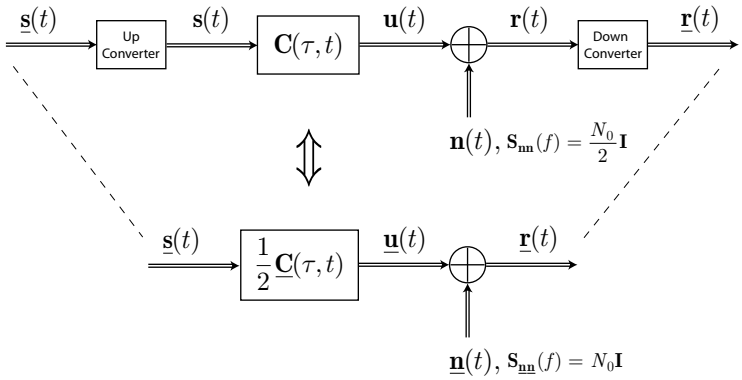

FIGURE 5.2 Equivalent passband and baseband models for the MIMO wireless channel.

5.2.2 BASEBAND MIMO INPUT DESIGN

The goal of a digital communication system is to convey digital information between two distinct points in space, separated by a communication channel. One of the points will act as the transmitter of information and the other as the receiver of information, though both can change roles. This information has usually been *formatted* as a sequence of binary digits (bits) $b_k = \{0,1\}$ that may contain some added redundancy to improve the fidelity of signal transmission. Implicitly, the binary sequence has a *binary rate* R_b and a binary *signalling interval* $T_b = 1/R_b$, which means that the digital communication system shall convey a bit every T_b seconds throughout the channel.

To reduce the effective rate of transmission, the transmitter typically maps each block of $r = \log_2(M)$ bits from the sequence $\{b_k\}$ to a *symbol* of the complex set $\{\underline{a}_i\}_{i=1}^{M}$, yielding the discrete-time sequence $\{\underline{a}_k\}$ with a *symbol rate* R_s and signalling interval $T_s = 1/R_s$. In a MIMO system there are n_T inputs to the channel, meaning that n_T distinct sequences of symbols can be transmitted at the same time. We denote this set of sequences as a vector-valued sequence $\{\underline{\mathbf{a}}_k\}$. To convey $\{\underline{\mathbf{a}}_k\}$ to the receiver, it must be converted in some way to a vector-valued continuous waveform $\underline{\mathbf{s}}(t)$, appropriate for transmission over the baseband channel $\underline{\mathbf{C}}(f,t)$. Recalling the sampling theorem, we notice that $\{\underline{\mathbf{a}}_k\}$ may be regarded as a sequence of vector-valued samples from some signal $\underline{\mathbf{x}}(t)$ with Fourier transform $\underline{\mathbf{x}}(f)$. Ideal sampling with an impulse train leads to

$$\underline{\mathbf{x}}_s(t) = \underline{\mathbf{x}}(t)\sum_{k=-\infty}^{\infty}\delta(t-kT_s) = \sum_{k=-\infty}^{\infty}\delta(t-kT_s)\underline{\mathbf{a}}_k, \tag{5.29}$$

which has the Fourier transform

$$\underline{\mathbf{x}}_s(f) = \underline{\mathbf{a}}\left(e^{j2\pi fT_s}\right) = \sum_{k=-\infty}^{\infty}\underline{\mathbf{a}}_k e^{-j2\pi fkT_s} = \frac{1}{T_s}\sum_{k=-\infty}^{\infty}\underline{\mathbf{x}}\left(f-\frac{k}{T_s}\right). \tag{5.30}$$

Since $\underline{\mathbf{x}}_s(f)$ is periodic with period R_s, in the worst case scenario ($\underline{\mathbf{x}}_s(f) > 0,\ \forall f$) the samples can be conveyed by a signal $\underline{\mathbf{x}}(t)$ with bandwidth $R_s/2$. One concludes at once that the transmitter may pass $\underline{\mathbf{x}}_s(f)$ through a lowpass filter $\underline{\mathbf{G}}(f)$ before transmission without the risk of losing any information required to reconstruct the samples. Letting $\underline{\mathbf{G}}(\tau)$ be the

impulse response of that filter, then $\underline{\mathbf{s}}(t)$ is obtained by convolving with $\underline{\mathbf{x}}_s(t)$, yielding

$$\underline{\mathbf{s}}(t) = \int_{-\infty}^{\infty} \underline{\mathbf{G}}(\tau)\underline{\mathbf{x}}_s(t-\tau)d\tau = \sum_{k=-\infty}^{\infty} \underline{\mathbf{G}}(t-kT_s)\underline{\mathbf{a}}_k. \tag{5.31}$$

The channel output is given by

$$\begin{aligned}\underline{\mathbf{r}}(t) &= \frac{1}{2}\int_0^{\infty} \underline{\mathbf{C}}(\alpha,t)\underline{\mathbf{s}}(t-\alpha)d\alpha + \underline{\mathbf{n}}(t) = \int_{-\infty}^{\infty} \underline{\mathbf{H}}(\tau,t)\underline{\mathbf{x}}_s(t-\tau)d\tau + \underline{\mathbf{n}}(t) \\ &= \sum_{k=-\infty}^{\infty} \underline{\mathbf{H}}(t-kT_s,t)\underline{\mathbf{a}}_k + \underline{\mathbf{n}}(t),\end{aligned} \tag{5.32}$$

where $\underline{\mathbf{H}}(\tau,t)$ is the convolution between $\underline{\mathbf{C}}(\alpha,t)$ and $\underline{\mathbf{G}}(\tau)$, i.e.

$$\underline{\mathbf{H}}(\tau,t) = \frac{1}{2}\underline{\mathbf{C}}(\alpha,t) \underset{\alpha}{\otimes} \underline{\mathbf{G}}(\tau) = \frac{1}{2}\int_0^{\infty} \underline{\mathbf{C}}(\alpha,t)\underline{\mathbf{G}}(\tau-\alpha)d\alpha, \tag{5.33}$$

which induces the frequency response

$$\underline{\mathbf{H}}(f,t) = \frac{1}{2}\underline{\mathbf{C}}(f,t)\underline{\mathbf{G}}(f). \tag{5.34}$$

For mathematical tractability, neither $\underline{\mathbf{G}}(\tau)$ nor $\underline{\mathbf{H}}(\tau,t)$ are yet being restricted to causal impulse responses. Based on the MIMO channel model, the input model of the MIMO system is now completely defined, as shown in Figure 5.3. If no noise was present and the

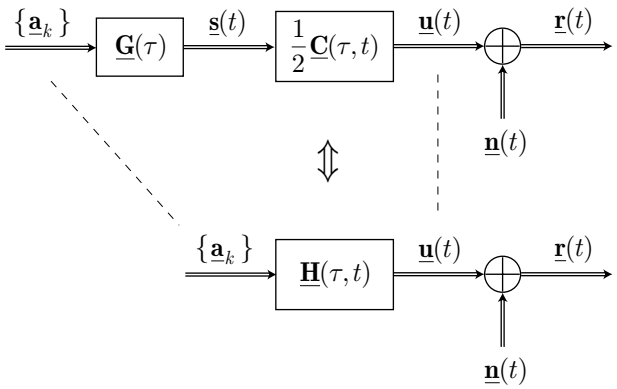

FIGURE 5.3 Transmission model with discrete-time vector-valued input sequence.

condition $n_R \geq n_T$ was met, it is apparent that the input sequence could be easily recovered by sampling the noiseless output. The sampling operation would yield

$$\underline{\mathbf{u}}(lT_s) = \sum_{k=-\infty}^{\infty} \underline{\mathbf{H}}((l-k)T_s, lT_s)\underline{\mathbf{a}}_k, \tag{5.35}$$

which, for a delay of D symbol intervals, would require

$$\underline{\mathbf{H}}(kT_s, lT_s) = \delta_{k-D}\mathbf{I} \qquad \text{for all } k \tag{5.36}$$

or equivalently

$$\underline{\mathbf{H}}\left(e^{j2\pi fT_s}, lT_s\right) = \frac{1}{T_s}\sum_{k=-\infty}^{\infty} \underline{\mathbf{H}}\left(f - \frac{k}{T_s}, lT_s\right) = e^{-j2\pi fDT_s}\mathbf{I}. \tag{5.37}$$

These equations define the generalization of the *Nyquist criterion* to MIMO time-varying channels. In other words, they mean that to achieve *linear separability* of the samples in the vector-valued input sequence, a $\underline{\mathbf{G}}(f)$ should be chosen in (5.34) so that $\underline{\mathbf{H}}(f,t)$ satisfies (5.37). Such $\underline{\mathbf{G}}(f)$ will exist only when $n_R \geq n_T$ ($\underline{\mathbf{H}}(\tau,t)$ is a column matrix) and the upper $n_T \times n_T$ matrix of $\underline{\mathbf{C}}(f,t)$ is nonsingular, hence invertible ($\underline{\mathbf{C}}(f,t)$ is full rank). Besides the corruption by channel noise, there are two naked-eye difficulties with this approach. The first is that $\underline{\mathbf{s}}(t)$ may (and generally will) be distorted by a frequency-selective $\underline{\mathbf{C}}(f,t)$ that the transmitter does not know, as it typically has no channel state information (CSI) available. The second is that, even if the transmitter knew $\underline{\mathbf{C}}(f,t)$, since the latter is time-dependent, the spectral shaping function $\underline{\mathbf{G}}(f)$ would also have to exhibit a time dependence (i.e. $\underline{\mathbf{G}}(f,t)$), with the inconvenience of (5.34) being no longer valid. In fact, by the same reasoning one concludes that

$$\underline{\mathbf{H}}(\tau,t) = \frac{1}{2}\int_0^{\tau_{\max}(t)} \underline{\mathbf{C}}(\alpha,t)\underline{\mathbf{G}}(\tau-\alpha,t-\alpha)d\alpha, \tag{5.38}$$

and

$$\begin{aligned}\underline{\mathbf{H}}(f,t) &= \frac{1}{2}\int_0^{\tau_{\max}(t)} \underline{\mathbf{C}}(\tau,t)\underline{\mathbf{G}}(f,t-\tau)e^{-j2\pi f\tau}d\tau \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \underline{\mathbf{C}}(f+\nu,t)\underline{\mathbf{G}}(f,\nu)_t e^{-j2\pi\nu t}d\nu,\end{aligned} \tag{5.39}$$

where $\tau_{\max}(t)$ is the maximum delay spread of the wireless channel at time t and

$$\underline{\mathbf{G}}(f,\nu)_t = \int_{t-\tau_{\max}(t)}^{t} \underline{\mathbf{G}}(f,\xi)e^{-j2\pi\nu\xi}d\xi. \tag{5.40}$$

It becomes self-evident that, because $\underline{\mathbf{H}}(f,t)$ depends on the past of $\underline{\mathbf{G}}(f,t)$ by a convolution (either in time or in Doppler frequency), it is much harder to design a transmission filter that satisfies (5.37) at every sampling instant, especially if the channel varies very rapidly. Yet sub-optimal, a practical alternative to achieve (near-)separability of samples would be to consider the channel response $\underline{\mathbf{C}}(f,t)$ interval-wise invariant. More specifically, one could consider $\underline{\mathbf{C}}(f,t)$ as approximately invariant during (at most) the coherence-time $(\Delta t)_c$ of the wireless channel, divide the time frame into equispaced intervals, and then specify a different function $\underline{\mathbf{G}}(f,t)$ for every set of samples lT_s within each interval. If each interval was $(\Delta t)_c$ seconds in duration we could roughly write

$$\underline{\mathbf{H}}(f)_m = \frac{1}{2}\underline{\mathbf{C}}(f)_m\underline{\mathbf{G}}(f)_m, \qquad m.(\Delta t)_c + \tau_{\max}(m.(\Delta t)_c) < t \leq (m+1).(\Delta t)_c, \tag{5.41}$$

and (5.37) would become

$$\frac{1}{T_s}\sum_{k=-\infty}^{\infty} \underline{\mathbf{H}}\left(f-\frac{k}{T_s}\right)_m = e^{-j2\pi fDT_s}\mathbf{I}, \tag{5.42}$$

for all $m = \lfloor t/(\Delta t)_c \rfloor$. The feasibility of this approach will naturally depend on the rate of variation of the channel response, and may be practical only for slowly-fading channels that satisfy the condition $(\Delta t)_c \gg \tau_{\max}(t)$, which is typically the case.

5.2.3 Baseband MIMO Output Design

After defining the vector-valued signal $\underline{\mathbf{s}}(t)$ to be transmitted based on the sequence of samples $\{\underline{\mathbf{a}}_k\}$, our interest diverges to signal reception and output MIMO design. Irrespective of the shaping filter $\underline{\mathbf{G}}(f,t)$ used at the transmitter, the discussion shall focus primarily on the global transfer function $\underline{\mathbf{H}}(f,t)$ and the discrete-input transmission model of (5.32). The optimum criterion for designing the MIMO output is the minimization of the error probability of symbol detection, that is, given the received signal

$$\underline{\mathbf{r}}(t) = \sum_{k=-\infty}^{\infty} \underline{\mathbf{H}}(t - kT_s, t)\underline{\mathbf{a}}_k + \underline{\mathbf{n}}(t) = \underline{\mathbf{u}}(t) + \underline{\mathbf{n}}(t), \tag{5.43}$$

the receiver should make a decision about which sequence $\{\underline{\mathbf{a}}_k\}$ of symbols was effectively transmitted, and this decision should be performed with minimum probability of error.

Since the input is discrete, it makes sense to search for a model with a discrete output that lacks no relevant information for the decision process. The first model that may come to mind is based on the sampling theorem. Due to the lowpass characteristic of $\underline{\mathbf{G}}(f)$, the noiseless output $\underline{\mathbf{u}}(t)$ will be bandlimited and have a finite energy given by

$$\int_{-\infty}^{\infty} \|\underline{\mathbf{u}}(t)\|^2 \, dt = \int_{-B_u/2}^{B_u/2} \|\underline{\mathbf{U}}(f)\|^2 \, df. \tag{5.44}$$

Its bandwidth B_u is (finitely) larger than that of $\underline{\mathbf{s}}(t)$ because of the (limited) time-variability of the channel response. From the sampling theorem it follows that $\underline{\mathbf{u}}(t)$ is uniquely determined by its samples $\underline{\mathbf{u}}(kT_u)$ taken in intervals of $T_u = 1/B_u$ seconds, and hence may be expanded into a series of orthogonal functions $\{\mathrm{sinc}((t - kT_u)/T_u)\mathbf{I}\}$ as

$$\underline{\mathbf{u}}(t) = \sum_{k=-\infty}^{\infty} [\mathrm{sinc}((t - kT_u)/T_u)\mathbf{I}]\underline{\mathbf{u}}(kT_u). \tag{5.45}$$

The samples $\underline{\mathbf{u}}(kT_u)$ may be recovered either by pure sampling or by projecting $\underline{\mathbf{u}}(t)$ onto the basis $\{\mathrm{sinc}((t - kT_u)/T_u)\mathbf{I}\}$ in the usual manner:

$$\underline{\mathbf{u}}(lT_u) = \frac{1}{T_u} \int_{-\infty}^{\infty} [\mathrm{sinc}((\tau - lT_u)/T_u)\mathbf{I}]\underline{\mathbf{u}}(\tau)d\tau. \tag{5.46}$$

Recalling from (5.32) that

$$\underline{\mathbf{u}}(t) = \sum_{k=-\infty}^{\infty} \underline{\mathbf{H}}(t - kT_s, t)\underline{\mathbf{a}}_k, \tag{5.47}$$

we conclude that the orthogonal set $\{\mathrm{sinc}((t - kT_u)/T_u)\mathbf{I}\}$ spans the signal space of the set $\{\underline{\mathbf{H}}(t - kT_s, t)\}$.

Trying to perform detection using the samples in (5.43) is obviously possible because they carry all the information present in $\underline{\mathbf{u}}(t)$, but brute estimation of the original symbol vector samples $\{\underline{\mathbf{a}}_k\}$ from the samples $\underline{\mathbf{r}}(kT_u)$ is another issue. Unfortunately, there are some serious drawbacks: 1) signal processing at a sampling rate higher than the input symbol rate is required, something that even the optimal detector may not need; 2) pre-filtering is necessary to reject out-of-band noise and, since no filter is ideal, may provoke signal damage; 3) the samples possess redundant information that is not important for detection,

so further filtering/processing is required to avoid overburdening the detector; 4) the physically unrealizable ideal functions $\mathrm{sinc}((t-kT_u)/T_u)$ (i.e. the Nyquist Kernel functions) lack full-orthogonality

$$\frac{1}{T_u}\sum_{k=-\infty}^{\infty}\mathrm{sinc}((t-kT_u)/T_u)\,\mathrm{sinc}((t'-kT_u)/T_u) \neq \delta(t'-t) \tag{5.48}$$

a fact that may impair detection for certain types of channels; and 5) they fail to capture the time-varying nature of the MIMO channel, which is not our intention from the very beginning.

We thus search for an alternative orthogonal set to discretize the output, and to this end we shall resort to: 1) the theory of general linear time-varying operators (see [95], [96] and [97] for insight); and 2) the classical method of Maximum Likelihood Sequence Estimation (MLSE) extended to a time-varying environment.

5.3 *Discretizing the MIMO Input-Output Model*

We have seen that the output of the MIMO channel is given by

$$\begin{aligned}\underline{\mathbf{r}}(t) &= \underline{\mathbf{u}}(t)+\underline{\mathbf{n}}(t) = \underline{\mathbf{H}}(\tau,t)\underset{\tau}{\otimes}\underline{\mathbf{x}}(t)+\underline{\mathbf{n}}(t)\\ &= \underline{\mathbf{H}}(\tau,t)\underset{\tau}{\otimes}\underline{\mathbf{a}}_l+\underline{\mathbf{n}}(t) = \sum_l \underline{\mathbf{H}}(t-lT_s,t)\underline{\mathbf{a}}_l+\underline{\mathbf{n}}(t)\end{aligned} \tag{5.49}$$

where we remark that the matrix valued functions $\underline{\mathbf{H}}(t-lT_s,t)$ are withdrawn from parallel lines in the domain of the graph of $\underline{\mathbf{H}}(\tau,t)$, as depicted in Figure 5.4. This observation con-

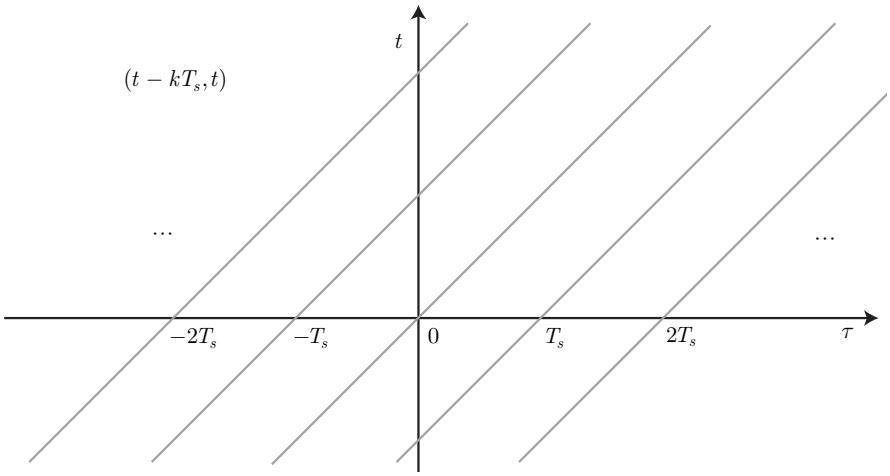

FIGURE 5.4 Transmission domains of a time-varying linear channel as seen by the receiver.

trasts with the time-invariant scenario, where the shaping functions are simply shifted replicas of one another. Here, each symbol is "carried over" to the output by a different function, making it harder for the receiver to understand what has happened since transmission. The channel is fast-fading if the response $\underline{\mathbf{H}}(\tau,t)$ varies between consecutive signalling intervals lT_s, which defaults to a significantly different $\underline{\mathbf{H}}(t-lT_s,t)$ being used every lT_s seconds.

Moving on, we try to devise a model for the system that avoids the intricate convolution notation. After some thought, we may realize that one possibility that makes reason-

able sense is to use operator theory to describe the channel. In fact, the expression for $\underline{\mathbf{u}}(t)$ can be rewritten as

$$\begin{aligned}\underline{\mathbf{u}}(t) = \underline{\mathbf{H}}(\tau,t) \underset{\tau}{\otimes} \underline{\mathbf{x}}_s(t) &= \int_{\tau} \underline{\mathbf{H}}(\tau,t)\underline{\mathbf{x}}_s(t-\tau)d\tau \\ &\underset{\tau=t-t'}{=} \int_{t'} \widehat{\underline{\mathbf{H}}}(t,t')\underline{\mathbf{x}}_s(t')dt'\end{aligned} \tag{5.50}$$

where

$$\widehat{\underline{\mathbf{H}}}(t,t') = \underline{\mathbf{H}}(t-t',t) \Leftrightarrow \underline{\mathbf{H}}(t,t') = \widehat{\underline{\mathbf{H}}}(t',t'-t) \tag{5.51}$$

denotes a two-sided-continuous time-varying matrix operator (or a matrix of operators). The operator $\widehat{\underline{\mathbf{H}}}(t,t')$ can be interpreted as a matrix of continuous-time matrices (the coefficients) and $\underline{\mathbf{x}}_s(t')$ as a vector of continuous-time vectors (functions). Indeed, we are prompted to use the standard matrix notation and algebraic conventions with no limitations, as long as we remind (5.51) and acknowledge that operator multiplication is an infinite sum of infinitesimal parts (the integral). The product of operators $\widehat{\underline{\mathbf{A}}}$ and $\widehat{\underline{\mathbf{B}}}$ is defined as

$$\widehat{\underline{\mathbf{C}}} = \widehat{\underline{\mathbf{A}}}\widehat{\underline{\mathbf{B}}} \triangleq \widehat{\underline{\mathbf{C}}}(t,t') = \int_{t''} \widehat{\underline{\mathbf{A}}}(t,t'')\widehat{\underline{\mathbf{B}}}(t'',t')dt'' \tag{5.52}$$

and the hermitian adjoint product given by

$$\widehat{\underline{\mathbf{C}}} = \widehat{\underline{\mathbf{A}}}^H\widehat{\underline{\mathbf{B}}} \triangleq \widehat{\underline{\mathbf{C}}}(t,t') = \int_{t''} \widehat{\underline{\mathbf{A}}}^H(t'',t)\widehat{\underline{\mathbf{B}}}(t'',t')dt'' \tag{5.53}$$

Other rules may follow and will be noted further ahead. We thus rewrite (5.49) as simply as

$$\underline{\mathbf{r}} = \underline{\mathbf{u}} + \underline{\mathbf{n}} = \widehat{\underline{\mathbf{H}}}\underline{\mathbf{x}}_s + \underline{\mathbf{n}} \tag{5.54}$$

Moreover, we can also give life to a one-sided-discrete time-varying matrix operator and reassemble (5.49) as

$$\begin{aligned}\underline{\mathbf{r}}(t) &= \sum_l \underline{\mathbf{H}}(t-lT_s,t)\underline{\mathbf{a}}_l + \underline{\mathbf{n}}(t) \\ &\triangleq \sum_l \dot{\underline{\mathbf{H}}}(t,lT_s)\underline{\mathbf{a}}_l + \underline{\mathbf{n}}(t) \rightarrow \underline{\mathbf{r}} = \dot{\underline{\mathbf{H}}}\dot{\underline{\mathbf{a}}} + \underline{\mathbf{n}}\end{aligned} \tag{5.55}$$

where the upper dot above the operator informs that it is discrete on one of its sides (implicit from the construction) and continuous on the other, and the dot above the vector informs that it behaves as a vector of discrete-time vectors. Although the output is still continuous, the integral in (5.50) has been replaced by an infinite sum.

5.3.1 Defining a Basis for MIMO Response Expansion

Now, what we would like to do is to approximate $\underline{\mathbf{u}}(t)$ as function of generic orthonormal basis functions $\underline{\mathbf{\Phi}}(t-lT_s,t)$, i.e.

$$\hat{\underline{\mathbf{u}}} = \dot{\underline{\mathbf{\Phi}}}\dot{\mathbf{b}} \triangleq \hat{\underline{\mathbf{u}}}(t) = \sum_l \underline{\mathbf{\Phi}}(t-lT_s,t)\underline{\mathbf{b}}_l = \underline{\mathbf{\Phi}}(\tau,t) \underset{\tau}{\otimes} \underline{\mathbf{b}}_l \tag{5.56}$$

such that the following semi-unitarity condition is satisfied

$$\dot{\underline{\Phi}}^H \dot{\underline{\Phi}} = \ddot{\underline{I}} \triangleq \int_t \underline{\Phi}^H(t - kT_s, t)\underline{\Phi}(t - lT_s, t)dt = \delta_{k-l}\mathbf{I} \tag{5.57}$$

We start by setting a finite energy requirement on $\underline{\mathbf{u}}(t)$, that is

$$\|\underline{\mathbf{u}}\|^2 = \underline{\mathbf{u}}^H \underline{\mathbf{u}} < \infty \triangleq \|\underline{\mathbf{u}}(t)\|^2 = \int_t \underline{\mathbf{u}}^H(t)\underline{\mathbf{u}}(t)dt < \infty \tag{5.58}$$

The energy of the estimator $\hat{\underline{\mathbf{u}}}$ equals the energy of $\dot{\underline{\mathbf{b}}}$, as shown by

$$\begin{aligned} \|\hat{\underline{\mathbf{u}}}\|^2 &= \hat{\underline{\mathbf{u}}}^H \hat{\underline{\mathbf{u}}} = \dot{\underline{\mathbf{b}}}^H \dot{\underline{\Phi}}^H \dot{\underline{\Phi}} \dot{\underline{\mathbf{b}}} = \dot{\underline{\mathbf{b}}}^H \dot{\underline{\mathbf{b}}} = \left\|\dot{\underline{\mathbf{b}}}\right\|^2 \\ &\triangleq \int_t \hat{\underline{\mathbf{u}}}^H(t)\hat{\underline{\mathbf{u}}}(t)dt = \sum_l \sum_m \underline{\mathbf{b}}_l^H \left[\int_t \underline{\Phi}^H(t - lT_s, t)\underline{\Phi}(t - mT_s, t)dt\right]\underline{\mathbf{b}}_m \\ &= \sum_l \underline{\mathbf{b}}_l^H \underline{\mathbf{b}}_l = \|\underline{\mathbf{b}}_l\|^2 \end{aligned} \tag{5.59}$$

At this point, we would like to determine the optimal vector $\dot{\underline{\mathbf{b}}}$. To do so, and for simplicity, we shall use the widely accepted *minimum squared error (MSE)* criterion, i.e.

$$\dot{\underline{\mathbf{b}}}_{opt} = \arg\min_{\dot{\underline{\mathbf{b}}}} \|\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}\|^2 = \arg\min_{\dot{\underline{\mathbf{b}}}} \left\|\underline{\mathbf{u}} - \dot{\underline{\Phi}}\dot{\underline{\mathbf{b}}}\right\|^2 \tag{5.60}$$

which will guarantee that the energy in the difference between $\hat{\underline{\mathbf{u}}}$ and $\underline{\mathbf{u}}$ will be minimum. We let $\hat{\underline{\mathbf{u}}}_o$ be the orthogonal projection of $\hat{\underline{\mathbf{u}}}$ onto the range of $\dot{\underline{\Phi}}$, $\mathcal{R}\left(\dot{\underline{\Phi}}\right)$, and write

$$\begin{aligned} \dot{\underline{\mathbf{b}}}_{opt} &= \arg\min_{\dot{\underline{\mathbf{b}}}} \left\|\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}_o + \hat{\underline{\mathbf{u}}}_o - \dot{\underline{\Phi}}\dot{\underline{\mathbf{b}}}\right\|^2 \\ [\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}_o] &\in \mathcal{R}^{\perp}\left(\dot{\underline{\Phi}}\right) = \mathcal{N}\left(\dot{\underline{\Phi}}^H\right) \\ \left[\hat{\underline{\mathbf{u}}}_o - \dot{\underline{\Phi}}\dot{\underline{\mathbf{b}}}\right] &\in \mathcal{R}\left(\dot{\underline{\Phi}}\right) \end{aligned} \tag{5.61}$$

where $\mathcal{N}\left(\dot{\underline{\Phi}}^H\right)$ is the nullspace of the hermitian adjoint $\dot{\underline{\Phi}}^H$. Hence, applying the Pythagorean theorem, we have

$$\dot{\underline{\mathbf{b}}}_{opt} = \arg\min_{\dot{\underline{\mathbf{b}}}} \left(\|\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}_o\|^2 + \left\|\hat{\underline{\mathbf{u}}}_o - \dot{\underline{\Phi}}\dot{\underline{\mathbf{b}}}\right\|^2\right) = \arg\min_{\dot{\underline{\mathbf{b}}}} \left\|\hat{\underline{\mathbf{u}}}_o - \dot{\underline{\Phi}}\dot{\underline{\mathbf{b}}}\right\|^2 \tag{5.62}$$

which leads to the conclusion that setting $\hat{\underline{\mathbf{u}}} = \dot{\underline{\Phi}}\dot{\underline{\mathbf{b}}} = \hat{\underline{\mathbf{u}}}_o$ minimizes the energy of the error vector. Letting $\underline{\mathbf{u}} = \hat{\underline{\mathbf{u}}}_o + [\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}_o]$, from (5.49) we know that

$$\begin{aligned} [\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}_o] \in \mathcal{N}\left(\dot{\underline{\Phi}}^H\right) &\rightarrow \dot{\underline{\Phi}}^H[\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}_o] = \mathbf{0} \\ &\triangleq \int_t \underline{\Phi}^H(t - lT_s, t)[\underline{\mathbf{u}}(t) - \hat{\underline{\mathbf{u}}}_o(t)]dt = \mathbf{0} \end{aligned} \tag{5.63}$$

and consequently

$$\begin{aligned} \dot{\underline{\Phi}}^H \hat{\underline{\mathbf{u}}}_o &= \dot{\underline{\Phi}}^H \underline{\mathbf{u}} \\ \dot{\underline{\Phi}}^H \dot{\underline{\Phi}} \dot{\underline{\mathbf{b}}}_{opt} &= \dot{\underline{\Phi}}^H \underline{\mathbf{u}} \rightarrow \dot{\underline{\mathbf{b}}}_{opt} = \dot{\underline{\Phi}}^H \underline{\mathbf{u}} = \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Phi}}} \\ &\triangleq \underline{\mathbf{b}}_k^{(opt)} = \int_t \dot{\underline{\Phi}}^H(t, lT_s)\underline{\mathbf{u}}(t)dt = \int_t \underline{\Phi}^H(t - lT_s, t)\underline{\mathbf{u}}(t)dt \end{aligned} \tag{5.64}$$

As expected, the optimal coefficients are obtained by orthogonally projecting $\underline{\mathbf{u}}(t)$ onto

the basis functions $\underline{\mathbf{\Phi}}(t-lT_s,t)$, or equivalently, onto $\mathcal{R}\left(\dot{\underline{\mathbf{\Phi}}}\right)$, and as the second line establishes, are *least square solutions* of $\dot{\underline{\mathbf{\Phi}}}\dot{\underline{\mathbf{b}}} = \underline{\mathbf{u}}$.

We realize that after selecting $\dot{\underline{\mathbf{\Phi}}}$ we automatically have $\dot{\underline{\mathbf{b}}}$, but we would also like to know if there is a choice of $\dot{\underline{\mathbf{\Phi}}}$ that sets the error energy to zero. From (5.49) we know that

$$\|\hat{\underline{\mathbf{u}}}_o\|^2 = \left\|\dot{\underline{\mathbf{b}}}_{opt}\right\|^2 = \dot{\underline{\mathbf{b}}}_{opt}^H \dot{\underline{\mathbf{b}}}_{opt} = \underline{\mathbf{u}}^H \dot{\underline{\mathbf{\Phi}}}\dot{\underline{\mathbf{\Phi}}}^H \underline{\mathbf{u}} \tag{5.65}$$

leading us to the conclusion that imposing the condition $\dot{\underline{\mathbf{\Phi}}}\dot{\underline{\mathbf{\Phi}}}^H = \ddot{\underline{\mathbf{I}}}$, $\hat{\underline{\mathbf{u}}}_o$ and $\underline{\mathbf{u}}$ will have the same energy. More interestingly, and a consequence of orthogonal projection, when they both have the same energy, they will also have zero energy difference, as can be seen by

$$\begin{aligned}\|\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}_o\|^2 &= \|\underline{\mathbf{u}}\|^2 - \underline{\mathbf{u}}^H \hat{\underline{\mathbf{u}}}_o - \underline{\mathbf{u}}\hat{\underline{\mathbf{u}}}_o^H + \|\hat{\underline{\mathbf{u}}}_o\|^2 \\ &= \|\underline{\mathbf{u}}\|^2 - 2\dot{\underline{\mathbf{b}}}_{opt}^H \dot{\underline{\mathbf{b}}}_{opt} + \underline{\mathbf{u}}^H \dot{\underline{\mathbf{\Phi}}}\dot{\underline{\mathbf{\Phi}}}^H \underline{\mathbf{u}} \underbrace{=}_{\dot{\underline{\mathbf{\Phi}}}\dot{\underline{\mathbf{\Phi}}}^H = \ddot{\underline{\mathbf{I}}}} 0\end{aligned} \tag{5.66}$$

We conclude that, besides the condition (5.57), the orthonormal set $\underline{\mathbf{\Phi}}(t-lT_s,t)$ should also satisfy the additional sufficient restriction

$$\dot{\underline{\mathbf{\Phi}}}\dot{\underline{\mathbf{\Phi}}}^H = \ddot{\underline{\mathbf{I}}} \triangleq \left[\sum_{l=-\infty}^{\infty} \underline{\mathbf{\Phi}}(t-lT_s,t)\underline{\mathbf{\Phi}}^H(t'-lT_s,t')\right] = \delta(t'-t)\mathbf{I} \tag{5.67}$$

which renders $\dot{\underline{\mathbf{\Phi}}}$ as a full-orthonormal (i.e. full-unitary) matrix operator and the set $\underline{\mathbf{\Phi}}(t-lT_s,t)$ a complete one. Moreover,

$$\left\|\dot{\underline{\mathbf{b}}}_{opt}\right\|^2 = \underline{\mathbf{u}}^H \dot{\underline{\mathbf{\Phi}}}\dot{\underline{\mathbf{\Phi}}}^H \underline{\mathbf{u}} = \|\underline{\mathbf{u}}\|^2 < \infty \tag{5.68}$$

and thus the estimate of $\underline{\mathbf{u}}(t)$ has finite energy, $\|\hat{\underline{\mathbf{u}}}_o\|^2 < \infty$.

In short, conditions (5.57), (5.58) and (5.67) guarantee an approximation almost everywhere (a.e.) between $\hat{\underline{\mathbf{u}}}(t)$ and $\underline{\mathbf{u}}(t)$

$$\|\underline{\mathbf{u}} - \hat{\underline{\mathbf{u}}}\|^2 \triangleq \int_t (\underline{\mathbf{u}}(t) - \hat{\underline{\mathbf{u}}}(t))^H (\underline{\mathbf{u}}(t) - \hat{\underline{\mathbf{u}}}(t))dt = 0 \quad \rightarrow \underline{\mathbf{u}}(t) = \hat{\underline{\mathbf{u}}}(t) \quad a.e. \tag{5.69}$$

and since in theory the optimal detection is always energy-based, and in practice $\underline{\mathbf{u}}(t)$ is always a "well-behaved" real-world vector-valued signal, one shall assume that $\underline{\mathbf{u}}(t) = \hat{\underline{\mathbf{u}}}(t)$ everywhere. The optimal coefficients, $\dot{\underline{\mathbf{b}}}_{opt}$, can be rewritten in terms of the original coefficients (the sequence of transmitted symbols), $\dot{\underline{\mathbf{a}}}$, as follows

$$\begin{aligned}\dot{\underline{\mathbf{b}}}_{opt} &= \dot{\underline{\mathbf{\Phi}}}^H \underline{\mathbf{u}} = \dot{\underline{\mathbf{\Phi}}}^H \dot{\underline{\mathbf{H}}}\dot{\underline{\mathbf{a}}} = \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}\dot{\underline{\mathbf{a}}} \\ &\triangleq \underline{\mathbf{b}}_k^{(opt)} = \sum_l \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}k,l}\underline{\mathbf{a}}_l = \sum_l \underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\underline{\mathbf{a}}_{k-l}\end{aligned} \tag{5.70}$$

where $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}} = \dot{\underline{\mathbf{\Phi}}}^H \dot{\underline{\mathbf{H}}}$ is a two-sided-discrete (two upper dots) matrix operator that denotes the projection of $\dot{\underline{\mathbf{H}}}$ onto $\dot{\underline{\mathbf{\Phi}}}$, and

$$\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}k,l} = \int_t \dot{\underline{\mathbf{\Phi}}}^H(t,kT_s)\dot{\underline{\mathbf{H}}}(t,lT_s)dt = \underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-l,k} = \int_t \underline{\mathbf{\Phi}}^H(t-kT_s,t)\underline{\mathbf{H}}(t-lT_s,t)dt \tag{5.71}$$

is the correlation (inner-product) between $\underline{\mathbf{H}}(t-lT_s,t)$ and $\underline{\mathbf{\Phi}}(t-kT_s,t)$. In turn, this allows

us to express $\underline{\dot{\mathbf{H}}}$ in terms of $\underline{\dot{\mathbf{\Phi}}}$ and $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}$ using (5.67), as

$$\begin{aligned}\underline{\dot{\mathbf{H}}} &= \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}} \\ &\triangleq \underline{\dot{\mathbf{H}}}(t, lT_s) = \sum_k \underline{\dot{\mathbf{\Phi}}}(t, kT_s)\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l} \\ \underline{\mathbf{H}}(t - lT_s, t) &= \sum_k \underline{\mathbf{\Phi}}(t - kT_s, t)\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-l,k}\end{aligned} \tag{5.72}$$

Using (5.70) and the Cauchy-Schwarz inequality, the finite energy condition in (5.68) can also be expressed in terms of the energy of the input vector $\underline{\dot{\mathbf{a}}}$ and the Frobenius norm of the operator $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}$, $\left\|\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}\right\|_F^2 = tr\left(\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}^H \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}\right)$, as follows

$$\begin{aligned}\left(\underline{\mathbf{b}}_k^{(i)}\right)^2 &= \left(\sum_l \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l}^{(i,*)} \underline{\mathbf{a}}_l\right)^2 \le \left(\sum_l \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l}^{(i,*)} \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l}^{(i,*)H}\right)\left(\sum_m \underline{\mathbf{a}}_m^H \underline{\mathbf{a}}_m\right) = \left(\sum_l \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l}^{(i,*)} \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l}^{(i,*)H}\right)\|\underline{\dot{\mathbf{a}}}\|^2 \\ \left\|\underline{\dot{\mathbf{b}}}_{opt}\right\|^2 &= \sum_i \sum_k \left(\underline{\mathbf{b}}_k^{(i)}\right)^2 \le \sum_i \left(\sum_k \sum_l \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l}^{(i,*)} \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l}^{(i,*)H}\right)\|\underline{\dot{\mathbf{a}}}\|^2 = \left\|\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}\right\|_F^2 \|\underline{\dot{\mathbf{a}}}\|^2\end{aligned} \tag{5.73}$$

which means that for the expansion to converge in energy it suffices that we guarantee

$$\|\underline{\hat{\mathbf{u}}}_o\|^2 = \left\|\underline{\dot{\mathbf{b}}}_{opt}\right\|^2 \le \left\|\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}\right\|_F^2 \|\underline{\dot{\mathbf{a}}}\|^2 < \infty \tag{5.74}$$

The first equality in (5.72) resembles the familiar QR factorization from linear algebra with matrices replaced by linear operators, except that, in order to be causal, $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}}$ is necessarily a lower-triangular operator, that is $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l} = \mathbf{0}, \forall k < l$. To determine $\underline{\dot{\mathbf{\Phi}}}$ we shall assume, with no loss of generality, that the first symbol to be transmitted was $\underline{\mathbf{a}}_0$ and the last $\underline{\mathbf{a}}_L$, so that we have

$$\underline{\dot{\mathbf{H}}}(t, lT_s) = \sum_{k \ge l}^{L} \underline{\dot{\mathbf{\Phi}}}(t, kT_s)\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l} \tag{5.75}$$

The concept behind the factorization is a backward, block-wise Gram-Schmidt orthogonalization procedure. After $\underline{\mathbf{a}}_L$ is received (and $\underline{\dot{\mathbf{H}}}(t, LT_s)$ is known) we compute the L$^{\text{th}}$ normal basis matrix as follows

$$\begin{aligned}\underline{\dot{\mathbf{H}}}(t, LT_s) &= \underline{\dot{\mathbf{\Phi}}}(t, LT_s)\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}L,L} \rightarrow \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\Phi}}L,L} = \left\|\underline{\dot{\mathbf{H}}}(t, LT_s)\right\| \\ \underline{\dot{\mathbf{\Phi}}}(t, LT_s) &= \underline{\dot{\mathbf{H}}}(t, LT_s)\left\|\underline{\dot{\mathbf{H}}}(t, LT_s)\right\|^{-1}\end{aligned} \tag{5.76}$$

where

$$\left\|\underline{\dot{\mathbf{H}}}(t, LT_s)\right\| = \left[\underline{\dot{\mathbf{H}}}^H \underline{\dot{\mathbf{H}}}\right]_{LL}^{1/2} = \left[\int_t \underline{\dot{\mathbf{H}}}^H(t, LT_s)\underline{\dot{\mathbf{H}}}(t, LT_s)dt\right]^{1/2} \tag{5.77}$$

The subsequent basis matrix, $\underline{\dot{\mathbf{\Phi}}}(t, (L-1)T_s)$, is found by projecting onto the previous one, computing the difference and normalizing the outcome, which yields

$$\begin{aligned}\underline{\dot{\mathbf{\Phi}}}(t, (L-1)T_s) = &\left[\underline{\dot{\mathbf{H}}}(t, (L-1)T_s) - \underline{\dot{\mathbf{\Phi}}}(t, LT_s)\int_t \underline{\dot{\mathbf{\Phi}}}^H(t, LT_s)\underline{\dot{\mathbf{H}}}(t, (L-1)T_s)dt\right] \cdot \\ &\cdot \left\|\underline{\dot{\mathbf{H}}}(t, (L-1)T_s) - \underline{\dot{\mathbf{\Phi}}}(t, LT_s)\int_t \underline{\dot{\mathbf{\Phi}}}^H(t, LT_s)\underline{\dot{\mathbf{H}}}(t, (L-1)T_s)dt\right\|^{-1}\end{aligned} \tag{5.78}$$

and

$$\begin{aligned}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}L-1,L-1} &= \left\| \dot{\underline{\mathbf{H}}}(t,(L-1)T_s) - \dot{\underline{\mathbf{\Phi}}}(t,LT_s)\int_t \dot{\underline{\mathbf{\Phi}}}^H(t,LT_s)\dot{\underline{\mathbf{H}}}(t,(L-1)T_s)dt \right\| \\ \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}L,L-1} &= \int_t \dot{\underline{\mathbf{\Phi}}}^H(t,LT_s)\dot{\underline{\mathbf{H}}}(t,(L-1)T_s)dt \end{aligned} \tag{5.79}$$

The same reasoning applies to any general $\dot{\underline{\mathbf{\Phi}}}(t,(L-m)T_s)$, which is derived by projecting onto to the previous space, finding the difference and normalizing, i.e.

$$\begin{aligned}\dot{\underline{\mathbf{\Phi}}}(t,(L-m)T_s) = &\left[\dot{\underline{\mathbf{H}}}(t,(L-m)T_s) - \sum_{i=L-m+1}^{L} \dot{\underline{\mathbf{\Phi}}}(t,iT_s)\int_t \dot{\underline{\mathbf{\Phi}}}^H(t,iT_s)\dot{\underline{\mathbf{H}}}(t,(L-m)T_s)dt\right]\cdot \\ &\cdot\left\|\dot{\underline{\mathbf{H}}}(t,(L-m)T_s) - \sum_{i=L-m+1}^{L} \dot{\underline{\mathbf{\Phi}}}(t,iT_s)\int_t \dot{\underline{\mathbf{\Phi}}}^H(t,iT_s)\dot{\underline{\mathbf{H}}}(t,(L-m)T_s)dt\right\|^{-1}\end{aligned} \tag{5.80}$$

and also

$$\begin{aligned}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}L-m,L-m} &= \left\| \dot{\underline{\mathbf{H}}}(t,(L-m)T_s) - \sum_{i=L-m+1}^{L} \dot{\underline{\mathbf{\Phi}}}(t,iT_s)\int_t \dot{\underline{\mathbf{\Phi}}}^H(t,iT_s)\dot{\underline{\mathbf{H}}}(t,(L-m)T_s)dt \right\| \\ \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}i,L-m} &= \int_t \dot{\underline{\mathbf{\Phi}}}^H(t,iT_s)\dot{\underline{\mathbf{H}}}(t,(L-m)T_s)dt \qquad\qquad L-m<i\leq L \end{aligned} \tag{5.81}$$

As it is, the process of computing the basis $\dot{\underline{\mathbf{\Phi}}}(t,(L-m)T_s)$ and $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}i,L-m}$ involves a great deal of processing, but provided that the matrix-valued functions $\left\{\dot{\underline{\mathbf{H}}}(t,lT_s)\right\}$ are independent of one another, it truly constructs an orthonormal basis for the received signal and, of no less importance, guarantees a causal $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}$. Another interesting fact is that if the channel varies slowly so that the coherence time spans several signalling intervals, the factorization in (5.72) is always possible, even if it results in a trivial solution as in (5.76). Whenever the channel remains constant, the next basis function that is computed will be related to the previous one by $\dot{\underline{\mathbf{H}}}(t,(l+1)T_s) = \dot{\underline{\mathbf{H}}}(t+lT_s,lT_s)$, showing that it should still be independent or otherwise the channel would not be causal. Even if the channel varies in such a manner that forces dependence between some channel functions, that is not an issue because the only difference is that some basis functions may be derived as null matrices.

More importantly, since in practice the channel has finite memory, it follows that $\dot{\underline{\mathbf{H}}}(t,(L-m)T_s)$ and $\dot{\underline{\mathbf{\Phi}}}(t,iT_s)$ are non-overlapping for $|i-(L-m)|>D$ (where $D=\lceil \tau_{\max}/T_s \rceil$ and $\tau_{\max}$ is the maximum delay spread of $\underline{\mathbf{H}}(\tau,t)$), and hence it will only be necessary to compute $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}i,L-m}$ for $i\leq L-m+D$. Accordingly, the expression for the optimal coefficients $\dot{\underline{\mathbf{b}}}_{opt}$ is simplified as

$$\underline{\mathbf{b}}_k^{(opt)} = \sum_{k-D<l\leq k} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}k,l}\underline{\mathbf{a}}_l \tag{5.82}$$

Also, the upper limit of the sums in (5.80) and (5.81) is in practice $L-m+D$ and not L, thus greatly streamlining the orthonormalization procedure. This is better seen by expand-

ing the decomposition onto its matrix form

$$\begin{pmatrix} \underline{\dot{\mathbf{H}}}(t,0T_s) \\ \underline{\dot{\mathbf{H}}}(t,1T_s) \\ \vdots \\ \underline{\dot{\mathbf{H}}}(t,(L-D)T_s) \\ \vdots \\ \underline{\dot{\mathbf{H}}}(t,(L-1)T_s) \\ \underline{\dot{\mathbf{H}}}(t,LT_s) \end{pmatrix}^T = \begin{pmatrix} \underline{\dot{\mathbf{\Phi}}}(t,0T_s) \\ \underline{\dot{\mathbf{\Phi}}}(t,1T_s) \\ \vdots \\ \underline{\dot{\mathbf{\Phi}}}(t,(L-D)T_s) \\ \vdots \\ \underline{\dot{\mathbf{\Phi}}}(t,(L-1)T_s) \\ \underline{\dot{\mathbf{\Phi}}}(t,LT_s) \end{pmatrix}^T \begin{pmatrix} \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}0,0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \iddots & \vdots & \ddots & \vdots \\ \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}D,0} & \cdots & \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}D,D} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \iddots & \vdots & \ddots & \vdots & \iddots & \vdots \\ \mathbf{0} & \cdots & \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}L-D,L-2D} & \cdots & \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}L-D,L-D} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \iddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}L,L-D} & \cdots & \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}L,L} \end{pmatrix} \tag{5.83}$$

from which it is clear that $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$ is not only lower triangular but band-based sparse with lower bandwidth D. Additionally, if $\tau_{\max} < T_s$ $(D = 1)$ then the channel is underspread in time and flat in frequency, so that there is no need for operator orthonormalization because (5.75) becomes $\underline{\dot{\mathbf{H}}}(t,lT_s) = \underline{\dot{\mathbf{\Phi}}}(t,lT_s)\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}l,l}$, so that $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$ is diagonal and a plain QR factorization (or simply (5.76)) is plenty. The main drawback for achieving causality is that we have to start the orthonormalization backwards from $\underline{\dot{\mathbf{H}}}(t,LT_s)$, something that may not be feasible if L is too large. An alternative approach may be to keep L sufficiently low and perform the orthonormalization for every L vector symbols, even though this will have some impact on the detection stages of the receiver.

All in all, we have decomposed the original operator $\underline{\dot{\mathbf{H}}}$ into terms $\underline{\dot{\mathbf{\Phi}}}$ and $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$, as depicted in Figure 5.5, where $\underline{\dot{\mathbf{\Phi}}}^H\underline{\dot{\mathbf{\Phi}}} = \underline{\ddot{\mathbf{I}}}$ and $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}} = \underline{\dot{\mathbf{\Phi}}}^H\underline{\dot{\mathbf{H}}}$. $\underline{\mathbf{\Phi}}(\tau,t)$ is any function that satisfies

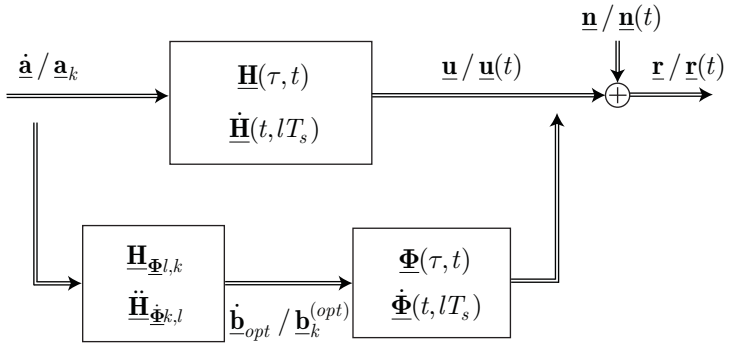

FIGURE 5.5 Decomposition of the MIMO transfer matrix into discrete and orthonormal operators.

the definition

$$\underline{\mathbf{\Phi}}(t - kT_s, t) = \underline{\dot{\mathbf{\Phi}}}(t, kT_s) \tag{5.84}$$

The output of $\underline{\dot{\mathbf{H}}}$ is given by

$$\underline{\mathbf{u}} = \underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{b}}}_{opt} = \hat{\underline{\mathbf{u}}}_o \tag{5.85}$$

indicating that the system $\underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{b}}} = \underline{\mathbf{u}}$ is in fact consistent ($\underline{\dot{\mathbf{\Phi}}}$ has $\underline{\mathbf{u}}(t)$ in its range), and its exact solutions are the least squares solutions. Additionally, since $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}} = \underline{\dot{\mathbf{\Phi}}}^H\underline{\dot{\mathbf{H}}}$, we have

$$\underline{\dot{\mathbf{b}}}_{opt} = \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}} = \underline{\dot{\mathbf{\Phi}}}^H\underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} \tag{5.86}$$

which must imply the following

$$\hat{\underline{\mathbf{u}}}_o = \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{b}}}_{opt} = \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{\Phi}}}^H\underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{\Phi}}}^H\underline{\mathbf{u}} \not\Rightarrow \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{\Phi}}}^H = \hat{\underline{\mathbf{I}}} \tag{5.87}$$

We conclude that an exact approximation to $\underline{\mathbf{u}}(t)$ has been found without requiring that $\underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{\Phi}}}^H = \hat{\mathbf{I}}$, that is, without the need of a full-orthonormal $\underline{\dot{\mathbf{\Phi}}}$.

Our intention at the beginning of this section was to discretize the MIMO input-output model in (5.49) or, equivalently, (5.53). Up to this point we have determined a "lossless" representation of $\underline{\mathbf{u}}(t)$ in terms of a semi-orthonormal $\underline{\dot{\mathbf{\Phi}}}$ extracted from the channel transfer operator $\underline{\dot{\mathbf{H}}}$. The input-output relation is rewritten as

$$\begin{aligned} \underline{\mathbf{r}} = \underline{\mathbf{u}} + \underline{\mathbf{n}} = \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{b}}}_{opt} + \underline{\mathbf{n}} = \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{u}}}_{\underline{\dot{\mathbf{\Phi}}}} + \underline{\mathbf{n}} \\ \triangleq \underline{\mathbf{r}}(t) = \sum_l \underline{\mathbf{\Phi}}(t - lT_s, t)\underline{\dot{\mathbf{u}}}_{\underline{\dot{\mathbf{\Phi}}}l} + \underline{\mathbf{n}}(t) \end{aligned} \tag{5.88}$$

and now we apply the adjoint operator on the left, yielding

$$\begin{aligned} \underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} &= \underline{\dot{\mathbf{\Phi}}}^H \underline{\mathbf{r}} = \underline{\dot{\mathbf{\Phi}}}^H \underline{\mathbf{u}} + \underline{\dot{\mathbf{\Phi}}}^H \underline{\mathbf{n}} = \underline{\dot{\mathbf{b}}}_{opt} + \underline{\dot{\mathbf{\Phi}}}^H \underline{\mathbf{n}} \\ &= \underline{\dot{\mathbf{u}}}_{\underline{\dot{\mathbf{\Phi}}}} + \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}} \\ &\triangleq \underline{\mathbf{r}}_{\underline{\mathbf{\Phi}}k} = \int_t \underline{\mathbf{\Phi}}^H(t - kT_s, t)\underline{\mathbf{r}}(t)dt = \underline{\mathbf{b}}_k^{(opt)} + \underline{\mathbf{n}}_{\underline{\mathbf{\Phi}}k} \end{aligned} \tag{5.89}$$

One should note that applying the adjoint operator is equivalent to passing the received signal $\underline{\mathbf{r}}(t)$ through a filter $\underline{\mathbf{\Phi}}^H(-\tau, t-\tau)$ (one that satisfies (5.84)) and then sampling at $t = kT_s$, because

$$\begin{aligned} \underline{\mathbf{r}}_{\underline{\mathbf{\Phi}}k} &= \int_t \underline{\mathbf{\Phi}}^H(t - kT_s, t)\underline{\mathbf{r}}(t)dt \underset{t=kT_s-\tau}{=} \int_\tau \underline{\mathbf{\Phi}}^H(-\tau, kT_s - \tau)\underline{\mathbf{r}}(kT_s - \tau)dt \\ &= \int_\tau \underline{\mathbf{\Phi}}^H(-\tau, t-\tau)\underline{\mathbf{r}}(t-\tau)dt\Big|_{t=kT_s} = \underline{\mathbf{\Phi}}^H(-\tau, t-\tau) \underset{\tau}{\otimes} \underline{\mathbf{r}}(t)\Big|_{t=kT_s} \end{aligned} \tag{5.90}$$

so that the model in Figure 5.5 is extended as shown in Figure 5.6. By the cascading, we

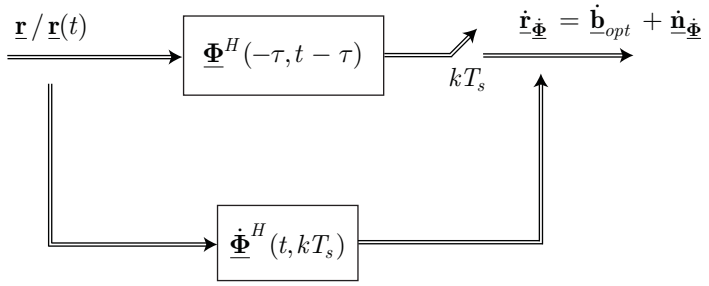

FIGURE 5.6 Discretizing the MIMO output with the semi-orthonormal channel operator $\underline{\dot{\mathbf{\Phi}}}$.

find that a discrete output has been recovered from the discrete input by applying the two-sided-discrete operator $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$ or discrete matrix filter $\underline{\mathbf{H}}_{\underline{\dot{\mathbf{\Phi}}}}$, as illustrated in Figure 5.7.

Even though this discrete model is obtained by simple projection onto the space spanned by the channel operator $\underline{\dot{\mathbf{\Phi}}}$, one does not know yet if it conveys all the information that the original continuous model conveyed. If it does, we are in good company, but if it does not, then some important information that could be useful to the detector's decisions may have been lost. We already know that $\underline{\mathbf{u}}(t)$ can be constructed from $\underline{\dot{\mathbf{b}}}_{opt}$ because $\underline{\dot{\mathbf{\Phi}}}$ spans the entire space of the channel's noiseless output, so there should be no problem here. However, $\underline{\dot{\mathbf{\Phi}}}$ does not span the entire noise space because the condition $\underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{\Phi}}}^H = \hat{\mathbf{I}}$ is not met, and as a result some information from the output $\underline{\mathbf{r}}(t)$ is lost after

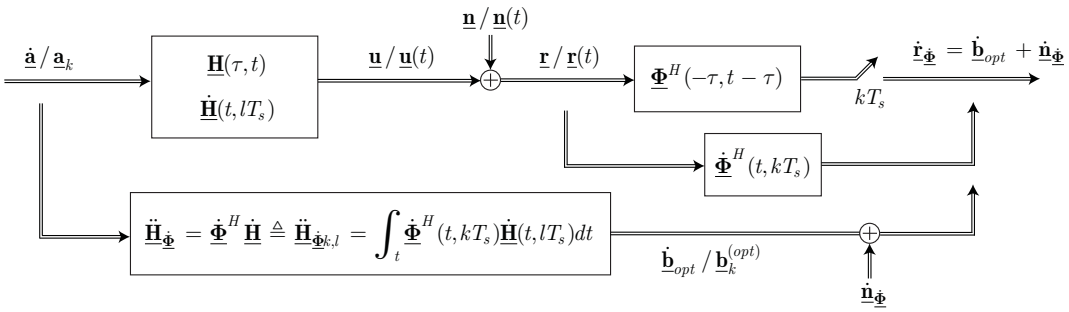

FIGURE 5.7 Discrete MIMO model after projecting the output onto $\mathcal{R}\left(\dot{\underline{\mathbf{\Phi}}}\right)$.

the projection. We shall now determine whether or not this information is relevant from an optimal input estimation perspective and, in case it is deemed irrelevant, the model of Figure 5.7 can be regarded as equivalent to the original.

5.3.2 TIME-VARYING MAXIMUM LIKELIHOOD SEQUENCE ESTIMATION (TV-MLSE)

Let us idealize a full-orthonormal matrix operator $\dot{\underline{\mathbf{\Psi}}}$ that satisfies both conditions $\dot{\underline{\mathbf{\Psi}}}^H\dot{\underline{\mathbf{\Psi}}} = \ddot{\mathbf{I}}$ and $\dot{\underline{\mathbf{\Psi}}}\dot{\underline{\mathbf{\Psi}}}^H = \ddot{\mathbf{I}}$. As discussed in the previous section, this operator spans the entire space of the MIMO noisy output $\underline{\mathbf{r}}(t)$, because if its estimate is

$$\hat{\underline{\mathbf{r}}} = \dot{\underline{\mathbf{\Psi}}}\dot{\underline{\mathbf{r}}}_{\dot{\underline{\mathbf{\Psi}}}} \qquad \dot{\underline{\mathbf{r}}}_{\dot{\underline{\mathbf{\Psi}}}} = \dot{\underline{\mathbf{\Psi}}}^H\underline{\mathbf{r}} \tag{5.91}$$

then it equals $\underline{\mathbf{r}}(t)$ at least almost everywhere

$$\hat{\underline{\mathbf{r}}} = \dot{\underline{\mathbf{\Psi}}}\dot{\underline{\mathbf{\Psi}}}^H\underline{\mathbf{r}} = \underline{\mathbf{r}} \qquad a.e. \tag{5.92}$$

Developing on this, we orthogonally project $\underline{\mathbf{r}}(t)$ onto the range of $\dot{\underline{\mathbf{\Psi}}}$ and obtain an equivalent discrete representation that losslessly describes $\underline{\mathbf{r}}(t)$. It is given by

$$\begin{aligned}\dot{\underline{\mathbf{r}}}_{\dot{\underline{\mathbf{\Phi}}}} &= \dot{\underline{\mathbf{\Psi}}}^H\underline{\mathbf{r}} = \dot{\underline{\mathbf{\Psi}}}^H\underline{\mathbf{u}} + \dot{\underline{\mathbf{\Psi}}}^H\underline{\mathbf{n}} = \dot{\underline{\mathbf{u}}}_{\dot{\underline{\mathbf{\Psi}}}} + \dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Psi}}}} \\ &\triangleq \underline{\mathbf{r}}_{\underline{\mathbf{\Psi}}k} = \int_t \underline{\mathbf{\Psi}}^H(t-kT_s,t)\underline{\mathbf{u}}(t)dt + \int_t \underline{\mathbf{\Psi}}^H(t-kT_s,t)\underline{\mathbf{n}}(t)dt = \underline{\mathbf{u}}_{\underline{\mathbf{\Psi}}k} + \underline{\mathbf{n}}_{\underline{\mathbf{\Psi}}k}\end{aligned} \tag{5.93}$$

The sole purpose of doing this is to simplify the mathematical treatment that will follow. From (5.28) we statistically model the input baseband noise as zero-mean space-time white gaussian noise with autocorrelation

$$\begin{aligned}\widehat{\underline{\mathbf{R}}}_{\underline{\mathbf{n}}\underline{\mathbf{n}}} &= \frac{1}{2}E[\,\underline{\mathbf{n}}\underline{\mathbf{n}}^H\,] = N_0\widehat{\underline{\mathbf{I}}} \\ &\triangleq \mathbf{R}_{\underline{\mathbf{n}}(t)\underline{\mathbf{n}}(t')}(t-t') = \frac{1}{2}E[\,\underline{\mathbf{n}}(t)\underline{\mathbf{n}}^H(t')] = N_0\delta(t-t')\mathbf{I}\end{aligned} \tag{5.94}$$

which leads to the following autocorrelation of the discretized noise

$$\begin{aligned}\ddot{\underline{\mathbf{R}}}_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Psi}}}}\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Psi}}}}} &= \frac{1}{2}E\left[\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Psi}}}}\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Psi}}}}^H\right] = \dot{\underline{\mathbf{\Psi}}}^H\frac{1}{2}E[\,\underline{\mathbf{n}}\underline{\mathbf{n}}^H\,]\dot{\underline{\mathbf{\Psi}}} = N_0\ddot{\mathbf{I}} \\ &\triangleq \mathbf{R}_{\underline{\mathbf{n}}_{\underline{\mathbf{\Psi}}k}\underline{\mathbf{n}}_{\underline{\mathbf{\Psi}}l}}(k-l) = N_0\delta_{k-l}\mathbf{I}\end{aligned} \tag{5.95}$$

and its multivariate normal distribution

$$\underline{\mathbf{n}}_{\underline{\mathbf{\Psi}}k} \sim \widetilde{\mathcal{N}}\left(E[\underline{\mathbf{n}}_{\underline{\mathbf{\Psi}}k}], \mathbf{R}_{\underline{\mathbf{n}}_{\underline{\mathbf{\Psi}}k}\underline{\mathbf{n}}_{\underline{\mathbf{\Psi}}k}}\right) \equiv \widetilde{\mathcal{N}}(\mathbf{0}, N_0\mathbf{I}) \tag{5.96}$$

Denoting the number of MIMO outputs as n_R, its pdf is swiftly derived as

$$\begin{aligned} f_{\underline{\mathbf{n}}_{\underline{\Psi}k}}\left(\underline{\mathbf{n}}_{\underline{\Psi}k}\right) &= (2\pi)^{-n_R}\left[\det(N_0\mathbf{I})\right]^{-1} e^{-\frac{1}{2}tr\left((N_0\mathbf{I})^{-1}\underline{\mathbf{n}}_{\underline{\Psi}k}\underline{\mathbf{n}}_{\underline{\Psi}k}^H\right)} \\ &= (2\pi N_0)^{-n_R}\, e^{-\frac{1}{2}N_0^{-1}\underline{\mathbf{n}}_{\underline{\Psi}k}^H\underline{\mathbf{n}}_{\underline{\Psi}k}} \end{aligned} \tag{5.97}$$

We also recognize that, due to its gaussian nature, the discrete vector components $\underline{\mathbf{n}}_{\underline{\Psi}k}$ and $\underline{\mathbf{n}}_{\underline{\Psi}l}$ are independent for $k \neq l$, which means that the pdf of $\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}$ is the product of (5.97) for every k, that is

$$f_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}}\left(\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}\right) = \prod_k f_{\underline{\mathbf{n}}_{\underline{\Psi}k}}\left(\underline{\mathbf{n}}_{\underline{\Psi}k}\right) = \left[\prod_k (2\pi N_0)^{-n_R}\right] e^{-\frac{1}{2}N_0^{-1}\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}^H\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}} \tag{5.98}$$

The importance of the pdf in (5.98) lies in the fact that it sets the ground to the maximum likelihood (ML) detection of the transmitted symbols. ML detection guarantees the lowest average error probability of the input symbol sequence, so it is optimal from this perspective. In fact, if the symbols of the input set are equally likely to be transmitted, then the classical Bayes rule permits the following

$$\dot{\underline{\mathbf{a}}}_{ML} = \arg\max_{\dot{\underline{\mathbf{a}}}} P\left(\dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}} \mid \dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}}\right) = \arg\max_{\dot{\underline{\mathbf{a}}}} f\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} \mid \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right) \tag{5.99}$$

where the conditional pdf can be retrieved from the noise pdf $f_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}}\left(\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}\right)$ by a simple transformation $\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}} = \dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}$, yielding

$$\begin{aligned} f\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} \mid \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right) &= f_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}}\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right)\left|\frac{\left(d\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}\right)^{\wedge}}{\left(d\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} \mid \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right)^{\wedge}}\right| = f_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Psi}}}}\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right) \\ &= \left[\prod_k (2\pi N_0)^{-n_R}\right] e^{-\frac{1}{2}N_0^{-1}\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right)^H\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right)} \end{aligned} \tag{5.100}$$

which is minimized when the quadratic form $\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right)^H\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right)$ is minimized, and thus (5.99) is rewritten in the form

$$\dot{\underline{\mathbf{a}}}_{ML} = \arg\max_{\dot{\underline{\mathbf{a}}}} f\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} \mid \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right) = \arg\min_{\dot{\underline{\mathbf{a}}}}\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right)^H\left(\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right) = \arg\min_{\dot{\underline{\mathbf{a}}}}\left\|\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}}\right\|^2 \tag{5.101}$$

But, from (5.93), $\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Psi}}} - \dot{\underline{\mathbf{u}}}_{\dot{\underline{\Psi}}} = \dot{\underline{\Psi}}^H(\underline{\mathbf{r}} - \underline{\mathbf{u}})$, and accordingly (5.101) becomes

$$\begin{aligned} \dot{\underline{\mathbf{a}}}_{ML} &= \arg\min_{\dot{\underline{\mathbf{a}}}}(\underline{\mathbf{r}} - \underline{\mathbf{u}})^H(\underline{\mathbf{r}} - \underline{\mathbf{u}}) = \arg\min_{\dot{\underline{\mathbf{a}}}}\left(\underline{\mathbf{r}} - \dot{\underline{\mathbf{H}}}\dot{\mathbf{a}}\right)^H\left(\underline{\mathbf{r}} - \dot{\underline{\mathbf{H}}}\dot{\mathbf{a}}\right) = \arg\min_{\dot{\underline{\mathbf{a}}}}\left\|\underline{\mathbf{r}} - \dot{\underline{\mathbf{H}}}\dot{\mathbf{a}}\right\|^2 \\ \triangleq \{\underline{\mathbf{a}}_k\}^{(ML)} &= \arg\min_{\{\underline{\mathbf{a}}_k\}}\int_t\left(\underline{\mathbf{r}}(t) - \sum_l \underline{\mathbf{H}}(t - lT_s, t)\underline{\mathbf{a}}_l\right)^H\left(\underline{\mathbf{r}}(t) - \sum_l \underline{\mathbf{H}}(t - lT_s, t)\underline{\mathbf{a}}_l\right)dt \\ &= \arg\min_{\{\underline{\mathbf{a}}_k\}}\int_t\left(\underline{\mathbf{r}}(t) - \underline{\mathbf{H}}(\tau, t)\underset{\tau}{\otimes}\underline{\mathbf{a}}_l\right)^H\left(\underline{\mathbf{r}}(t) - \underline{\mathbf{H}}(\tau, t)\underset{\tau}{\otimes}\underline{\mathbf{a}}_l\right)dt \end{aligned} \tag{5.102}$$

In words, the input vector sequence of maximum likelihood is the one that minimizes the energy between the noisy and noiseless outputs of the MIMO channel, a conclusion that extends the SISO time-invariant result with evident similarity.

5.3.3 TIME-SHEAR MATRIX MATCHED FILTER (TS-MMF)

Expanding the norm in (5.102), we reach

$$\begin{aligned}\dot{\underline{\mathbf{a}}}_{ML} &= \arg\min_{\dot{\underline{\mathbf{a}}}} \left(\underline{\mathbf{r}} - \dot{\underline{\underline{\mathbf{H}}}}\dot{\underline{\mathbf{a}}}\right)^H \left(\underline{\mathbf{r}} - \dot{\underline{\underline{\mathbf{H}}}}\dot{\underline{\mathbf{a}}}\right) = \arg\min_{\dot{\underline{\mathbf{a}}}} \left[-2\,\mathrm{Re}\left\{\dot{\underline{\mathbf{a}}}^H \dot{\underline{\underline{\mathbf{H}}}}^H \underline{\mathbf{r}}\right\} + \left\|\dot{\underline{\underline{\mathbf{H}}}}\dot{\underline{\mathbf{a}}}\right\|^2\right] \\ &= \arg\min_{\dot{\underline{\mathbf{a}}}} \left[-2\,\mathrm{Re}\left\{\dot{\underline{\mathbf{a}}}^H \dot{\underline{\mathbf{y}}}\right\} + \left\|\dot{\underline{\underline{\mathbf{H}}}}\dot{\underline{\mathbf{a}}}\right\|^2\right]\end{aligned} \tag{5.103}$$

where $\dot{\underline{\mathbf{y}}} = \dot{\underline{\underline{\mathbf{H}}}}^H \underline{\mathbf{r}}$ is the output of the adjoint operator and the only output information required for ML decision. Again, recalling the definition in (5.51), the adjoint channel operator expands as

$$\begin{aligned}\dot{\underline{\mathbf{y}}} = \dot{\underline{\underline{\mathbf{H}}}}^H \underline{\mathbf{r}} \triangleq \underline{\mathbf{y}}(kT_s) &= \int_t \dot{\underline{\underline{\mathbf{H}}}}^H(t, kT_s)\underline{\mathbf{r}}(t)dt = \int_t \underline{\underline{\mathbf{H}}}^H(t - kT_s, t)\underline{\mathbf{r}}(t)dt \\ &= \underline{\underline{\mathbf{H}}}^H(-\tau, t-\tau) \underset{\tau}{\otimes} \underline{\mathbf{r}}(t)\Big]_{t=kT_s}\end{aligned} \tag{5.104}$$

yielding $\underline{\underline{\mathbf{H}}}^H(-\tau, t-\tau)$ as the matrix matched filter (MMF) for the time-varying MIMO system. This is the optimal filter for signal detection. It is non-causal, non-predictive, and in effect translates to the adjoint operator $\widehat{\underline{\underline{\mathbf{H}}}}^H(t', t)$ in the operator domain, because

$$\begin{aligned}\underline{\mathbf{y}}(t) &= \int_\tau \underline{\underline{\mathbf{H}}}^H(-\tau, t-\tau)\underline{\mathbf{r}}(t-\tau)d\tau \underset{\tau=t-t'}{=} \int_{t'} \underline{\underline{\mathbf{H}}}^H(t'-t, t')\underline{\mathbf{r}}(t')dt' \\ &= \int_{t'} \widehat{\underline{\underline{\mathbf{H}}}}^H(t', t)\underline{\mathbf{r}}(t')dt'\end{aligned} \tag{5.105}$$

so that the optimum receiver operator is always the hermitian adjoint operator of the transmitter operator. It is important to note that $\underline{\underline{\mathbf{H}}}^H(-\tau, t-\tau)$ is not necessarily obtained from the entire $\underline{\underline{\mathbf{H}}}(\tau, t)$ as long as it satisfies

$$\underline{\underline{\mathbf{H}}}^H(-\tau, kT_s - \tau) = \dot{\underline{\underline{\mathbf{H}}}}^H(kT_s - \tau, kT_s) \tag{5.106}$$

or, in other words, equals the operator $\dot{\underline{\underline{\mathbf{H}}}}^H$ in the transmission domains of Figure 5.4. The matched filter is solely required to be matched to the transmission domains, not the entire $\underline{\underline{\mathbf{H}}}(\tau, t)$. This fact comes as no surprise in that, since the input symbols are carried by matrix-valued functions $\underline{\underline{\mathbf{H}}}(t - lT_s, t)$, the best receiver should correlate with each of them, and it indeed does, as shown by

$$\begin{aligned}\underline{\mathbf{y}}(kT_s) &= \underline{\underline{\mathbf{H}}}^H(-\tau, t-\tau) \underset{\tau}{\otimes} \underline{\mathbf{r}}(t)\Big]_{t=kT_s} = \int_t \underline{\underline{\mathbf{H}}}^H(t - kT_s, t)\underline{\mathbf{r}}(t)dt \\ &= \sum_l \left[\int_t \underline{\underline{\mathbf{H}}}^H(t - kT_s, t)\underline{\underline{\mathbf{H}}}(t - lT_s, t)dt\right]\underline{\mathbf{a}}_l + \int_t \underline{\underline{\mathbf{H}}}^H(t - kT_s, t)\underline{\mathbf{n}}(t)dt\end{aligned} \tag{5.107}$$

or, more simply, as

$$\begin{aligned}\underline{\mathbf{y}}(kT_s) &= \sum_l \underline{\mathbf{S}}((k-l)T_s, kT_s)\,\underline{\mathbf{a}}_l + \underline{\mathbf{n}}_{\underline{\mathbf{H}}k} = \sum_l \underline{\mathbf{S}}(lT_s, kT_s)\,\underline{\mathbf{a}}_{k-l} + \underline{\mathbf{n}}_{\underline{\mathbf{H}}k} \\ &\triangleq \dot{\underline{\mathbf{y}}} = \ddot{\underline{\mathbf{S}}}\,\dot{\underline{\mathbf{a}}} + \dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{H}}}} = \dot{\underline{\mathbf{u}}}_{\dot{\underline{\mathbf{H}}}} + \dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{H}}}}\end{aligned} \tag{5.108}$$

where

$$\begin{aligned}\underline{\mathbf{S}}((k-l)T_s,kT_s) &= \underline{\mathbf{H}}_{\underline{\mathbf{H}}k-l,k} = \int_t \underline{\mathbf{H}}^H(t-kT_s,t)\underline{\mathbf{H}}(t-lT_s,t)dt \\ \triangleq \ddot{\underline{\mathbf{S}}}(kT_s,lT_s) &= \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{H}}}k,l} = \int_t \dot{\underline{\mathbf{H}}}^H(t,kT_s)\dot{\underline{\mathbf{H}}}(t,lT_s)dt \triangleq \ddot{\underline{\mathbf{S}}} = \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{H}}}} = \dot{\underline{\mathbf{H}}}^H\dot{\underline{\mathbf{H}}}\end{aligned} \tag{5.109}$$

is the correlation between the different $\underline{\mathbf{H}}(t-lT_s,t)$. Contrary to the LTI scenario, $\underline{\mathbf{S}}(lT_s,kT_s)$ is non-stationary because it depends on each particular k and the difference $k-l$, hence not Toeplitz. One may also check that

$$\begin{aligned}\|\underline{\mathbf{u}}\|^2 &= \underline{\mathbf{u}}^H\underline{\mathbf{u}} = \dot{\underline{\mathbf{a}}}^H\dot{\underline{\mathbf{H}}}^H\dot{\underline{\mathbf{H}}}\dot{\underline{\mathbf{a}}} = \dot{\underline{\mathbf{a}}}^H\ddot{\underline{\mathbf{S}}}\dot{\underline{\mathbf{a}}} \geq 0 \\ &\triangleq \|\underline{\mathbf{u}}(t)\|^2 = \sum_k \underline{\mathbf{a}}_k^H \sum_l \left[\int_t \underline{\mathbf{H}}^H(t-kT_s,t)\underline{\mathbf{H}}(t-lT_s,t)dt\right]\underline{\mathbf{a}}_l \\ &= \sum_k \underline{\mathbf{a}}_k^H \sum_l \underline{\mathbf{S}}((k-l)T_s,kT_s)\underline{\mathbf{a}}_l \geq 0\end{aligned} \tag{5.110}$$

meaning that $\ddot{\underline{\mathbf{S}}}$ is non-negative definite hermitian, hence a valid covariance operator. The continuous MMF output before sampling is given by

$$\begin{aligned}\underline{\mathbf{y}} &= \dot{\underline{\mathbf{S}}}(t,lT_s)\dot{\underline{\mathbf{a}}} + \widehat{\underline{\mathbf{H}}}^H\underline{\mathbf{n}} \\ &\triangleq \underline{\mathbf{y}}(t) = \sum_l \underline{\mathbf{S}}(t-lT_s,t)\,\underline{\mathbf{a}}_l + \underline{\mathbf{n}}_{\underline{\mathbf{H}}}(t) = \underline{\mathbf{S}}(\tau,t)\underset{\tau}{\otimes}\underline{\mathbf{a}}_l + \underline{\mathbf{H}}^H(-\tau,t-\tau)\underset{\tau}{\otimes}\underline{\mathbf{n}}(t)\end{aligned} \tag{5.111}$$

and as a result the transfer function between the MIMO input and the matched filter output is the continuous correlation function

$$\begin{aligned}\underline{\mathbf{S}}(\tau,t) &= \underline{\mathbf{H}}^H(-\alpha,t-\alpha)\underset{\alpha}{\otimes}\underline{\mathbf{H}}(\tau,t) \\ &= \int_\alpha \underline{\mathbf{H}}^H(-\alpha,t-\alpha)\underline{\mathbf{H}}(\tau-\alpha,t-\alpha)d\alpha = \int_\alpha \underline{\mathbf{H}}^H(\alpha-t,\alpha)\underline{\mathbf{H}}(\alpha-(t-\tau),\alpha)d\alpha\end{aligned} \tag{5.112}$$

This matrix valued bi-dimensional function is not causal due to the non-causality of the matrix matched filter, and gathers a lot of redundant information that can be safely ignored (by sampling) without penalty.

The sampled correlation matrix $\underline{\mathbf{S}}(lT_s,kT_s)$ can be seen as the correlation between the different transmission domains in Figure 5.4, a fact that leads us to another interpretation of the matrix matched filter. Writing $\underline{\mathbf{MMF}}(\tau,t) = \underline{\mathbf{H}}^H(-\tau,t-\tau)$, we see that there is an implicit linear transformation of coordinates

$$(\tau,t) \rightarrow (-\tau,t-\tau) \quad \Leftrightarrow \quad \begin{pmatrix}\tau \\ t\end{pmatrix} = \begin{pmatrix}-1 & 0 \\ -1 & 1\end{pmatrix}\begin{pmatrix}\tau' \\ t'\end{pmatrix} \tag{5.113}$$

where (τ',t') are the old coordinates. This transformation is visualized in Figure 5.8, from which it can be perceived that the matched filter has the form of a $\pi/4$-skewed mirroring of the channel transfer-function (TF) $\underline{\mathbf{H}}(\tau,t)$. In other words, $\underline{\mathbf{H}}(\tau,t)$ experiences a vertical (time) shear of unit slope followed by an horizontal (delay) reflection. The delay reflection is already present in the classical time-invariant matched filter ($\underline{\mathbf{MMF}}(\tau) = \underline{\mathbf{H}}^H(-\tau)$), so the peculiar aspect in this case is the time shearing of the MIMO time-varying response. In light of this, we shall distinguish this new filter with the name *time-shear matrix matched filter* (TS-MMF). The system diagram of the TS-MMF is shown in Figure 5.9.

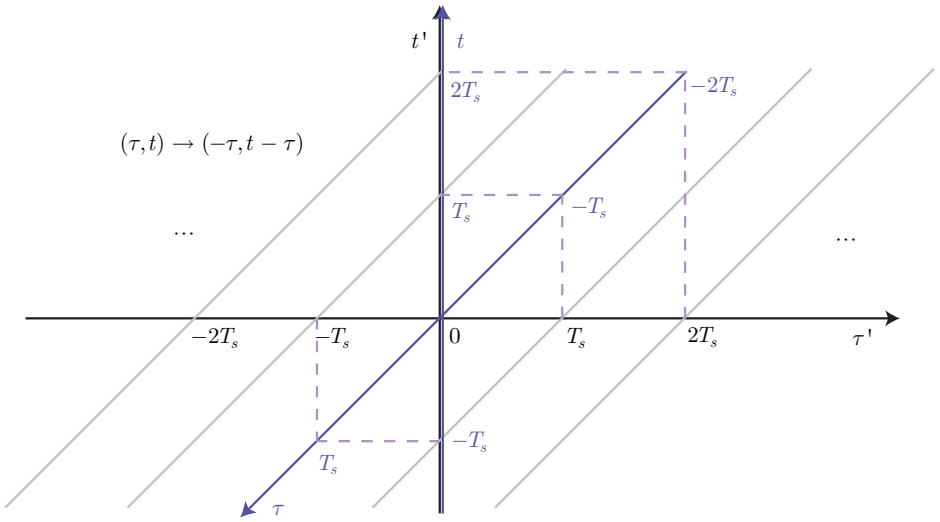

FIGURE 5.8 Mapping of the MIMO channel response onto its associated matched filter.

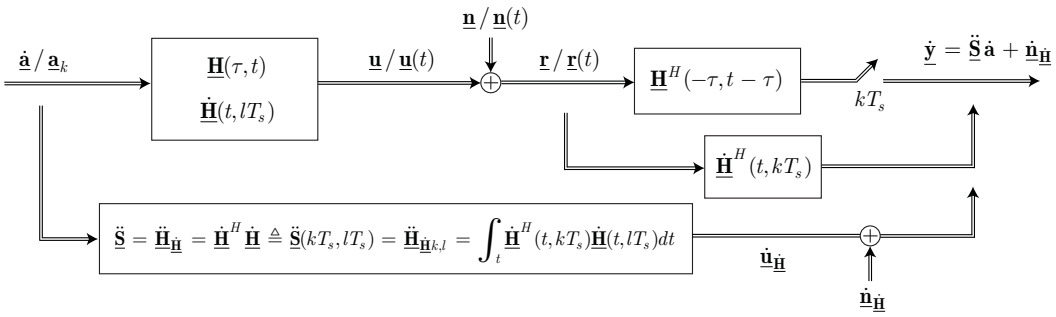

FIGURE 5.9 Matched filtering of the MIMO time-varying channel model.

5.3.4 DRAWBACKS OF THE TS-MMF

Despite being the filter that minimizes the error rate at the detector output, the TS-MMF introduces some difficulties into the system, mainly implementation related. First of all, the implementation of the time domain samples $\underline{\mathbf{H}}^H(-\tau, kT_s - \tau)$ of the TS-MMF derived from the TF $\underline{\mathbf{H}}(\tau,t)$ requires a perfect knowledge of the time-varying nature of the channel. While this is possible for slowly varying channels, it is totally prohibitive for channels where $(\Delta t)_c < T_s < \tau_{\max}$ because $\underline{\mathbf{H}}(\tau,t)$ effectively changes before it has the chance to be measured. The only solution in such an adverse scenario may be channel prediction. On the other hand, whenever the channel is fast-fading such that $\tau_{\max} < (\Delta t)_c < T_s$, a possible candidate to fill estimation gaps may be interpolation [98], [99].

Another difficulty with the TS-MMF is its non-causality, due to the mirroring effect applied to $\underline{\mathbf{H}}(\tau,t)$. Consequently, implementation of $\underline{\mathbf{H}}^H(-\tau, t-\tau)$ is only possible if a sufficient delay $\tau_0 > \tau_{\max}$ is introduced into the filter, resulting

$$\underline{\mathbf{MMF}}(\tau - \tau_0, t) = \underline{\mathbf{H}}^H(\tau_0 - \tau, t + \tau_0 - \tau) \tag{5.114}$$

as depicted in Figure 5.10. Inconveniently, this means that the output at time t requires knowledge of the channel at time $t + \tau_0$, so the TS-MMF trades its causality by predictivity. This knowledge is possible by sending a pilot at time $t = t'$ and waiting τ_0 seconds

to get the full response $H(t-t',t)$, and then perform some variable changes

$$H(t-t',t) \overset{+t'}{\rightarrow} H(t,t'+t) \overset{-}{\rightarrow} H(-t,t'-t) \overset{-\tau_0}{\rightarrow} H(\tau_0 - t, t'+\tau_0 - t) \tag{5.115}$$

Again, this is a problem when $(\Delta t)_c < T_s < \tau_{\max}$ because even if one decided to wait τ_0

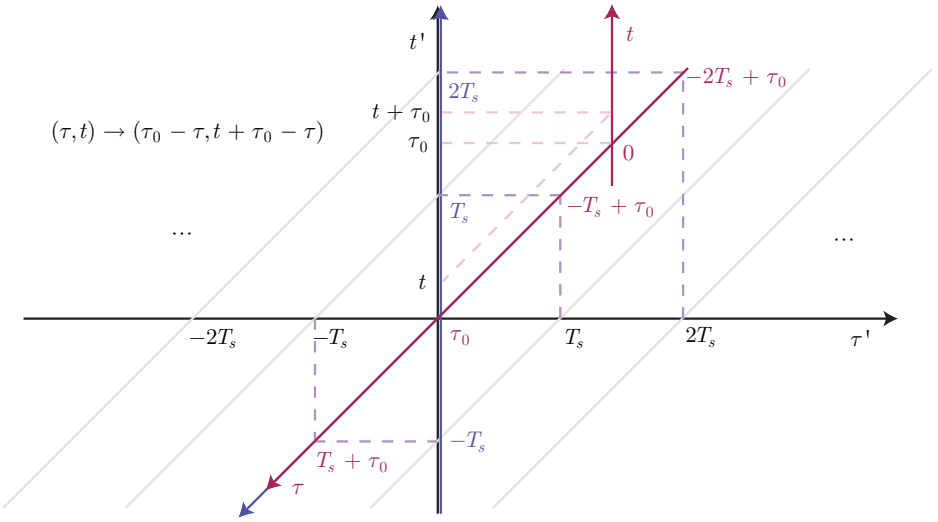

FIGURE 5.10 Delayed matched filtering of the MIMO time-varying channel model.

seconds for the response, it would vary during that time, meaning that it would not be possible to do the same for a $t = t''$ if $t''-t' < \tau_{\max}$.

Finally, another important problem of the TS-MMF concerns to detection. From Figure 5.9 we see that the noise output after sampling is $\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}} = \underline{\dot{\mathbf{H}}}^H \underline{\mathbf{n}}$, which leads to the following non-stationary autocorrelation

$$\begin{aligned} \underline{\ddot{\mathbf{R}}}_{\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}} \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}} &= \frac{1}{2} E\left[\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}} \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}^H\right] = \underline{\dot{\mathbf{H}}}^H \frac{1}{2} E[\underline{\mathbf{n}}\underline{\mathbf{n}}^H] \underline{\dot{\mathbf{H}}} = N_0 \underline{\dot{\mathbf{H}}}^H \underline{\dot{\mathbf{H}}} = N_0 \underline{\ddot{\mathbf{S}}} \\ &\triangleq \underline{\mathbf{R}}_{\underline{\mathbf{n}}_{\underline{\mathbf{H}}'} \underline{\mathbf{n}}_{\underline{\mathbf{H}}'}}(k-l,k) = N_0 \int_t \underline{\mathbf{H}}^H(t-kT_s,t) \underline{\mathbf{H}}(t-lT_s,t) dt = N_0 \underline{\mathbf{S}}((k-l)T_s, kT_s) \end{aligned} \tag{5.116}$$

Thus, the output noise vector samples are not uncorrelated, hence the noise is not white and the decision process is more complicated because the distribution

$$f_{\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}}\left(\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}\right) = \left[\det\left(2\pi \underline{\ddot{\mathbf{R}}}_{\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}} \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}}\right)\right]^{-1} e^{-\frac{1}{2} tr\left(\left(\underline{\ddot{\mathbf{R}}}_{\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}} \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}}\right)^{-1} \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}} \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}^H\right)} \tag{5.117}$$

is much harder to work with. Since $\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{H}}}}$ is a sufficient statistic, the ML criterion would be

$$\begin{aligned} \underline{\dot{\mathbf{a}}}_{ML} &= \arg\min_{\underline{\dot{\mathbf{a}}}} \left(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{H}}}} - \underline{\dot{\mathbf{u}}}_{\underline{\dot{\mathbf{H}}}}\right)^H \underline{\ddot{\mathbf{R}}}^{-1}_{\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}} \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{H}}}}} \left(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{H}}}} - \underline{\dot{\mathbf{u}}}_{\underline{\dot{\mathbf{H}}}}\right) \\ &= \arg\min_{\underline{\dot{\mathbf{a}}}} \left(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{H}}}} - \underline{\ddot{\mathbf{S}}}\,\underline{\dot{\mathbf{a}}}\right)^H \underline{\ddot{\mathbf{S}}}^{-1} \left(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{H}}}} - \underline{\ddot{\mathbf{S}}}\,\underline{\dot{\mathbf{a}}}\right) \end{aligned} \tag{5.118}$$

This is caused by the lack of semi-orthonormality of the channel operator $\underline{\dot{\mathbf{H}}}$, so adds up as an extra motivation for (losslessly) orthogonalizing the channel.

5.3.5 Optimality of the Discrete MIMO Model in $\mathcal{R}\left(\underline{\dot{\mathbf{\Phi}}}\right)$

At this point we no longer need the operator $\underline{\dot{\mathbf{\Psi}}}$, introduced in Section 5.3.2, so we return

to the semi-orthonormal channel operator $\underline{\dot{\mathbf{\Phi}}}$ from Section 5.3.1 and try to express the ML criterion in (5.102) within $\mathcal{R}\big(\underline{\dot{\mathbf{\Phi}}}\big)$, checking for its equivalence. The projection of the noise $\underline{\mathbf{n}}(t)$ onto $\mathcal{R}\big(\underline{\dot{\mathbf{\Phi}}}\big)$ is denoted by $\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}} = \underline{\dot{\mathbf{\Phi}}}^H \underline{\mathbf{n}}$, so the MIMO input-output relation after the channel decomposition $\underline{\dot{\mathbf{H}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$ equals

$$\underline{\mathbf{r}} = \underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} + \underline{\mathbf{n}} = \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}} + \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}} + \underline{\boldsymbol{\varepsilon}} = \underline{\dot{\mathbf{\Phi}}}\big(\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}} + \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}}\big) + \underline{\boldsymbol{\varepsilon}} \tag{5.119}$$

where $\underline{\boldsymbol{\varepsilon}} \triangleq \underline{\boldsymbol{\varepsilon}}(t)$ is the part of the noise out of $\mathcal{R}\big(\underline{\dot{\mathbf{\Phi}}}\big)$. Furthermore, because of orthogonal projection we note that it is orthogonal to $\mathcal{R}\big(\underline{\dot{\mathbf{\Phi}}}\big)$, i.e.

$$\underline{\boldsymbol{\varepsilon}} = \underline{\mathbf{n}} - \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}} = \underline{\mathbf{r}} - \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} \qquad \rightarrow \underline{\dot{\mathbf{\Phi}}}^H \underline{\boldsymbol{\varepsilon}} = \mathbf{0} \rightarrow \underline{\boldsymbol{\varepsilon}} \in \mathcal{N}\Big(\underline{\dot{\mathbf{\Phi}}}^H\Big) = \mathcal{R}^{\perp}\big(\underline{\dot{\mathbf{\Phi}}}\big) \tag{5.120}$$

Based on these facts, one may revise (5.102) by adding and subtracting $\underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}}$,

$$\begin{aligned}\underline{\dot{\mathbf{a}}}_{ML} &= \arg\min_{\underline{\dot{\mathbf{a}}}} \big\|\underline{\mathbf{r}} - \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big\|^2 = \arg\min_{\underline{\dot{\mathbf{a}}}} \big\|\big(\underline{\mathbf{r}} - \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}}\big) + \underline{\dot{\mathbf{\Phi}}}\big(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)\big\|^2 \\ &= \arg\min_{\underline{\dot{\mathbf{a}}}} \Big[\big\|\underline{\mathbf{r}} - \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}}\big\|^2 + \big\|\underline{\dot{\mathbf{\Phi}}}\big(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)\big\|^2\Big] = \arg\min_{\underline{\dot{\mathbf{a}}}} \big\|\underline{\dot{\mathbf{\Phi}}}\big(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)\big\|^2\end{aligned} \tag{5.121}$$

because $\big(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)^H \underline{\dot{\mathbf{\Phi}}}^H \underline{\boldsymbol{\varepsilon}} = \mathbf{0}$ and $\big\|\underline{\mathbf{r}} - \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}}\big\|$ is independent of $\underline{\dot{\mathbf{a}}}$. And, finally, we express the maximum likelihood estimates $\underline{\dot{\mathbf{a}}}_{ML}$ as a sole function of the discretized MIMO model of Figure 5.7 or, equivalently, the space $\mathcal{R}\big(\underline{\dot{\mathbf{\Phi}}}\big)$,

$$\begin{aligned}\underline{\dot{\mathbf{a}}}_{ML} &= \arg\min_{\underline{\dot{\mathbf{a}}}} \big\|\underline{\dot{\mathbf{\Phi}}}\big(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)\big\|^2 = \arg\min_{\underline{\dot{\mathbf{a}}}} \Big[\big(\underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)^H \big(\underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)\Big] \\ &= \arg\min_{\underline{\dot{\mathbf{a}}}} \Big[\big(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)^H \underline{\dot{\mathbf{\Phi}}}^H \underline{\dot{\mathbf{\Phi}}}\big(\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big)\Big] \\ &= \arg\min_{\underline{\dot{\mathbf{a}}}} \big\|\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}}\big\|^2\end{aligned} \tag{5.122}$$

The exclusive use of the identity $\underline{\dot{\mathbf{\Phi}}}^H \underline{\dot{\mathbf{\Phi}}} = \underline{\ddot{\mathbf{I}}}$ should be noticed from this deduction, and also how elegantly the latter has been accomplished due to its operator-based nature. One concludes that, after projecting the MIMO output onto the operator $\underline{\dot{\mathbf{\Phi}}}$ extracted from the channel transfer matrix $\underline{\mathbf{H}}(\tau, t)$, all the information required for optimal detection is still available, so one can confidently rely on the time-varying causal discrete model

$$\begin{aligned}\underline{\dot{\mathbf{r}}}_{\underline{\dot{\mathbf{\Phi}}}} &= \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}} + \underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}} \\ &\triangleq \underline{\mathbf{r}}_{\underline{\mathbf{\Phi}}k} = \sum_{l=k-D}^{k} \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}k,l}\underline{\mathbf{a}}_l + \underline{\mathbf{n}}_{\underline{\mathbf{\Phi}}k} = \sum_{l=0}^{D} \underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\underline{\mathbf{a}}_{k-l} + \underline{\mathbf{n}}_{\underline{\mathbf{\Phi}}k}\end{aligned} \tag{5.123}$$

where $D = \lceil \tau_{\max} / T_s \rceil$, for any ensuing analysis. $\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k} = \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}k,k-l}$ is taken from (5.81). Not only the model is attractively discrete and causal, but also the noise is still space-time white because

$$\underline{\ddot{\mathbf{R}}}_{\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}}} = \frac{1}{2}E\Big[\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{n}}}_{\underline{\dot{\mathbf{\Phi}}}}^H\Big] = \underline{\dot{\mathbf{\Phi}}}^H \frac{1}{2}E[\underline{\mathbf{n}}\underline{\mathbf{n}}^H]\underline{\dot{\mathbf{\Phi}}} = N_0\underline{\ddot{\mathbf{I}}} \tag{5.124}$$

which is a highly desirable characteristic since it simplifies the statistical treatment of the model. Capitalizing on this, (5.122) has the simple form

$$\underline{\dot{\mathbf{a}}}_{ML} = \arg\min_{\underline{\dot{\mathbf{a}}}} \sum_{k} \left\| \underline{\mathbf{r}}_{\underline{\mathbf{\Phi}}k} - \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}k,k}\underline{\mathbf{a}}_k - \sum_{l=k-D}^{k-1} \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}k,l}\underline{\mathbf{a}}_l \right\|^2 \tag{5.125}$$

where $\underline{\mathbf{a}}_k$ is the current channel input and $\{\underline{\mathbf{a}}_{k-D},...,\underline{\mathbf{a}}_{k-1}\}$ is the set of previous inputs (or current state) that accounts for the channel memory. Thus a trellis of states can be built and the Viterbi algorithm applied with confluent path elimination at each stage m and surviving distance metrics

$$DM_{surv}\{\underline{\mathbf{a}}_m,...,\underline{\mathbf{a}}_{m-D+1}\} = \min_{\underline{\mathbf{a}}_{m-D}} \left\| \underline{\mathbf{r}}_{\underline{\mathbf{\Phi}}m} - \sum_{l=m-D}^{m} \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}m,l}\underline{\mathbf{a}}_l \right\|^2 + DM_{surv}\{\underline{\mathbf{a}}_{m-1},...,\underline{\mathbf{a}}_{m-D}\} \quad (5.126)$$

5.3.6 Decomposing the TS-MMF into Discrete and Semi-Orthonormal Factors

We gather from Figure 5.5 that $\underline{\mathbf{H}}(\tau,t)$ can be successfully decomposed into factors $\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}$ and $\underline{\mathbf{\Phi}}(\tau,t)$, still, up to this point, we only have the expression that decomposes $\underline{\dot{\mathbf{H}}}$ into terms $\underline{\dot{\mathbf{\Phi}}}$ and $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$, $\underline{\dot{\mathbf{H}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$

$$\begin{aligned}\underline{\dot{\mathbf{H}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}} \triangleq \underline{\dot{\mathbf{H}}}(t,lT_s) &= \sum_k \underline{\dot{\mathbf{\Phi}}}(t,kT_s)\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}k,l} \\ &= \underline{\mathbf{H}}(t-lT_s,t) = \sum_k \underline{\mathbf{\Phi}}(t-kT_s,t)\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-l,k}\end{aligned} \quad (5.127)$$

It is not possible to retrieve $\underline{\mathbf{H}}(\tau,t)$ from $\underline{\dot{\mathbf{\Phi}}}$ and $\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$, because they only possess essential information that is conveyed from the channel, that is, the transmission domains of Figure 5.4. However, it is possible to build a simple function that equals the operator $\underline{\dot{\mathbf{H}}}$ on these domains and is null otherwise

$$\underline{\mathbf{H}}_{(td)}(\tau,t) = \underline{\mathbf{H}}(\tau,t)\sum_k \delta_{(\tau-(t-kT_s))} = \sum_k \underline{\mathbf{H}}(t-kT_s,t)\delta_{(\tau-(t-kT_s))} \quad (5.128)$$

where δ_{a-b} is the usual Kronecker delta. Its sampled version is

$$\underline{\mathbf{H}}_{(tds)}(\tau,t) = \underline{\mathbf{H}}(\tau,t)\sum_k \delta(\tau-(t-kT_s)) = \sum_k \underline{\mathbf{H}}(t-kT_s,t)\delta(\tau-(t-kT_s)) \quad (5.129)$$

The function in (5.128) can adequately replace $\underline{\mathbf{H}}(\tau,t)$ as

$$\underline{\mathbf{u}}(t) = \underline{\mathbf{\Phi}}(\alpha,t)\underset{\alpha}{\otimes}\underline{\mathbf{b}}_k^{(opt)} = \underline{\mathbf{\Phi}}(\alpha,t)\underset{\alpha}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\underset{l}{\otimes}\underline{\mathbf{a}}_k = \underline{\mathbf{H}}(\tau,t)\underset{\tau}{\otimes}\underline{\mathbf{a}}_m = \underline{\mathbf{H}}_{(td)}(\tau,t)\underset{\tau}{\otimes}\underline{\mathbf{a}}_m \quad (5.130)$$

so it must be the case that

$$\underline{\mathbf{H}}_{(tds)}(\tau,t) = \underline{\mathbf{H}}_{(td)}(\tau,t)\sum_k \delta(\tau-(t-kT_s)) = \underline{\mathbf{\Phi}}(\alpha,t)\underset{\alpha}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k} \quad (5.131)$$

which must be interpreted using the two-dimensional Dirac impulse function $\delta^2(x,y)$ in the following manner

$$\underline{\mathbf{H}}_{(tds)}(\tau,t) = \underline{\mathbf{\Phi}}(\alpha,t)\underset{\alpha}{\otimes}\sum_{l,k}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\delta^2(\tau-lT_s,t-kT_s) \quad (5.132)$$

Convolving, we find

$$\begin{aligned}\underline{\mathbf{H}}_{(tds)}(\tau,t) &= \sum_{l,k}\left[\int_{\alpha}\underline{\mathbf{\Phi}}(\alpha,t)\delta^2(\tau-\alpha-lT_s,t-\alpha-kT_s)d\alpha\right]\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\\ &= \sum_{l}\left[\sum_{k}\underline{\mathbf{\Phi}}(t-kT_s,t)\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-l,k}\right]\delta(\tau-(t-lT_s))\end{aligned} \tag{5.133}$$

and, naturally

$$\begin{aligned}\underline{\mathbf{H}}(t-mT_s,t) = \underline{\mathbf{H}}_{(td)}(t-mT_s,t) &= \sum_{l}\left[\sum_{k}\underline{\mathbf{\Phi}}(t-kT_s,t)\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-l,k}\right]\delta_{l-m}\\ &\overset{l=m}{=}\sum_{k}\underline{\mathbf{\Phi}}(t-kT_s,t)\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-m,k}\end{aligned} \tag{5.134}$$

matching (5.127). So, (5.133) seems to be the expression we were looking for. To dispel any remaining doubts we double-check by substitution into (5.130)

$$\begin{aligned}\underline{\mathbf{u}}(t) &= \underline{\mathbf{H}}_{(td)}(\tau,t)\underset{\tau}{\otimes}\underline{\mathbf{a}}_m = \underline{\mathbf{H}}_{(td)}(\tau,t)\underset{\tau}{\otimes}\sum_{m}\underline{\mathbf{a}}_m\delta(t-mT_s)\\ &= \sum_{l}\left[\sum_{k}\underline{\mathbf{\Phi}}(t-kT_s,t)\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-l,k}\right]\sum_{m}\underline{\mathbf{a}}_m\int_{\tau}\delta_{(\tau-(t-lT_s))}\delta(t-\tau-mT_s)d\tau\end{aligned} \tag{5.135}$$

and set $\tau = t - mT_s$ and $m = l$, yielding

$$\begin{aligned}\underline{\mathbf{u}}(t) &= \sum_{l}\left[\sum_{k}\underline{\mathbf{\Phi}}(t-kT_s,t)\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}k-l,k}\right]\underline{\mathbf{a}}_l\\ &\overset{l\to k-l}{=}\sum_{k}\underline{\mathbf{\Phi}}(t-kT_s,t)\sum_{l}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\underline{\mathbf{a}}_{k-l} = \sum_{k}\underline{\mathbf{\Phi}}(t-kT_s,t)\underline{\mathbf{b}}_k^{(opt)}\end{aligned} \tag{5.136}$$

which is the expected expansion of $\underline{\mathbf{u}}(t)$ in $\mathcal{R}\left(\dot{\underline{\mathbf{\Phi}}}\right)$, utterly confirming our conjecture.

Now, one may wonder if the TS-MMF $\underline{\mathbf{H}}^H(-\tau,t-\tau)$ can be decomposed similarly to (5.131). With operator conventions we simply find the hermitian adjoint of (5.127) and get

$$\dot{\underline{\mathbf{H}}}^H = \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}}\dot{\underline{\mathbf{\Phi}}}^H \triangleq \dot{\underline{\mathbf{H}}}^H(t,lT_s) = \sum_{k}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}k,l}\dot{\underline{\mathbf{\Phi}}}^H(t,kT_s) \tag{5.137}$$

but its relation with $\underline{\mathbf{H}}^H(-\tau,t-\tau)$ may seem obscure. So what we do is to start from (5.132) as follows

$$\begin{aligned}&\underline{\mathbf{H}}_{(tds)}(\tau,t) = \underline{\mathbf{\Phi}}(\alpha,t)\underset{\alpha}{\otimes}\sum_{l,k}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\delta^2(\tau-lT_s,t-kT_s)\\ &\rightarrow \underline{\mathbf{H}}^H_{(tds)}(\tau,t) = \sum_{l,k}\int_{\alpha}\underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}l,k}\delta^2(\tau-\alpha-lT_s,t-\alpha-kT_s)\underline{\mathbf{\Phi}}^H(\alpha,t)d\alpha\end{aligned} \tag{5.138}$$

and now perform appropriate variable changes to reach at

$$\begin{aligned}\underline{\underline{\mathbf{H}}}^H_{(tds)}(-\tau,t-\tau) &= \sum_{l,k}\int_\alpha \underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}l,k}\delta^2(-\tau-\alpha-lT_s,t-\tau-\alpha-kT_s)\underline{\underline{\mathbf{\Phi}}}^H(\alpha,t-\tau)d\alpha \\ &= \sum_{l,k}\int_\alpha \underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,k-l}\delta^2(\alpha-lT_s,t-\alpha-kT_s+lT_s)\underline{\underline{\mathbf{\Phi}}}^H(-\tau+\alpha,t-\tau)d\alpha \\ &= \sum_{l,k}\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,k-l}\delta^2(\alpha-lT_s,t-kT_s)\underset{\alpha}{\otimes}\underline{\underline{\mathbf{\Phi}}}^H(-\tau,t-\tau)\end{aligned} \tag{5.139}$$

Comparing this expression with (5.132) we deduce that the TS-MMF is decomposed as

$$\underline{\underline{\mathbf{H}}}^H(-\tau,t-\tau)\equiv\underline{\underline{\mathbf{H}}}^H_{(tds)}(-\tau,t-\tau)=\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,k-l}\underset{\alpha}{\otimes}\underline{\underline{\mathbf{\Phi}}}^H(-\tau,t-\tau) \tag{5.140}$$

which is the convolution between a discrete matched filter $\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,k-l}$ and a continuous orthonormalized matched filter $\underline{\underline{\mathbf{\Phi}}}^H(-\tau,t-\tau)$ (from this point on to be designated ORTHO-TS-MMF). Indeed the decomposition is possible and it is strongly consistent with (5.131).

To eliminate the doubts that the TS-MMF decomposition could have been inferred from (5.137) we recall (5.104), where it is found that $\dot{\underline{\underline{\mathbf{H}}}}^H$ behaves as a matched filter followed by sampling

$$\begin{aligned}\underline{\mathbf{y}}(kT_s)=\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{H}}}}k} &= \int_t \dot{\underline{\underline{\mathbf{H}}}}^H(t,kT_s)\underline{\mathbf{r}}(t)dt \\ &= \underline{\underline{\mathbf{H}}}^H(-\tau,t-\tau)\underset{\tau}{\otimes}\underline{\mathbf{r}}(t)\Big]_{t=kT_s}\equiv\underline{\underline{\mathbf{H}}}^H_{(td)}(-\tau,t-\tau)\underset{\tau}{\otimes}\underline{\mathbf{r}}(t)\Big]_{t=kT_s}\end{aligned} \tag{5.141}$$

Equivalently,

$$\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{\Phi}}}}k}=\int_t\dot{\underline{\underline{\mathbf{\Phi}}}}^H(t,kT_s)\underline{\mathbf{r}}(t)dt=\underline{\underline{\mathbf{\Phi}}}^H(-\tau,t-\tau)\underset{\tau}{\otimes}\underline{\mathbf{r}}(t)\Big]_{t=kT_s} \tag{5.142}$$

and

$$\begin{aligned}\underline{\mathbf{y}}(kT_s)=\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{H}}}}k} &= \sum_l\ddot{\underline{\underline{\mathbf{H}}}}^H_{\dot{\underline{\underline{\mathbf{\Phi}}}}l,k}\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{\Phi}}}}l}=\sum_l\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}l-k,l}\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{\Phi}}}}l}=\sum_l\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,k-l}\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{\Phi}}}}k-l} \\ &= \underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,k-l}\underset{l}{\otimes}\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{\Phi}}}}k}=\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,r-l}\underset{\alpha}{\otimes}\dot{\underline{\mathbf{r}}}_{\dot{\underline{\underline{\mathbf{\Phi}}}}k}=\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-l,r-l}\underset{\alpha}{\otimes}\underline{\underline{\mathbf{\Phi}}}^H(-\tau,t-\tau)\underset{\tau}{\otimes}\underline{\mathbf{r}}(t)\Big]_{t=kT_s}\end{aligned} \tag{5.143}$$

from which we confirm (5.140).

Moreover, it is straightforward to show that the full transfer function given by (5.112) can also be decomposed into discrete and orthogonal terms. Using the identity

$$\begin{aligned}\ddot{\underline{\underline{\mathbf{\Omega}}}}&=\dot{\underline{\underline{\mathbf{\Phi}}}}^H\dot{\underline{\underline{\mathbf{\Phi}}}}=\ddot{\underline{\mathbf{I}}} \\ &\overset{\rho=t-t'}{\triangleq}\underline{\underline{\mathbf{\Omega}}}(\rho,t)\big]_{t=kT_s}\overset{t"\to t-\alpha}{=}\underline{\underline{\mathbf{\Phi}}}^H(-\alpha,t-\alpha)\underset{\alpha}{\otimes}\underline{\underline{\mathbf{\Phi}}}(\rho,t)\Big]_{t=kT_s}=\delta_\rho\mathbf{I}\end{aligned} \tag{5.144}$$

we are led to

$$\begin{aligned}\sum_k\underline{\underline{\mathbf{S}}}_{(tds)}(\tau,t)\Big]_{t=kT_s}\delta(t-kT_s)=\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-r,p-r}\underset{\beta}{\otimes}\Bigg[\sum_k\underline{\underline{\mathbf{\Phi}}}^H(-\alpha,t-\alpha)\underset{\alpha}{\otimes}\underline{\underline{\mathbf{\Phi}}}(\rho,t)\Big]_{t=kT_s}\cdot \\ \cdot\delta(t-kT_s)\Bigg]\underset{\rho}{\otimes}\underline{\underline{\mathbf{H}}}_{\underline{\underline{\mathbf{\Phi}}}l,m}=\sum_k\underline{\underline{\mathbf{H}}}^H_{\underline{\underline{\mathbf{\Phi}}}-r,p-r}\underset{\beta}{\otimes}\underline{\underline{\mathbf{H}}}_{\underline{\underline{\mathbf{\Phi}}}l,m}\delta(t-kT_s)\end{aligned} \tag{5.145}$$

and the last member of (5.145) can be simplified as follows

$$\begin{aligned}\sum_k \underline{\mathbf{S}}_{(tds)}(\tau,t)\big]_{t=kT_s}\delta(t-kT_s) &= \sum_{r,p}\underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,p-r}\sum_{l,m}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,m}\sum_k \delta(t-kT_s)\delta(t-pT_s)\cdot\\ \cdot\delta^2(\tau-rT_s-lT_s,t-rT_s-mT_s) &\overset{k=p\wedge m=k-r}{=} \sum_{k,l}\left[\sum_r \underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,k-r}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l-r,k-r}\right]\cdot\\ \cdot\delta^2(\tau-lT_s,t-kT_s) &= \sum_{k,l}\Big[\underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,k-r}\underset{r}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\Big]\delta^2(\tau-lT_s,t-kT_s)\end{aligned} \tag{5.146}$$

Finally, from the definition in (5.129) we have the result

$$\begin{aligned}\sum_k \underline{\mathbf{S}}_{(tds)}(\tau,t)\big]_{t=kT_s}\delta(t-kT_s) &= \underline{\mathbf{S}}(\tau,t)\sum_{k,l}\delta(\tau-(t-lT_s))\delta(t-kT_s)\\ &= \sum_{k,l}\underline{\mathbf{S}}(lT_s,kT_s)\delta^2(\tau-lT_s,t-kT_s)\end{aligned} \tag{5.147}$$

which equated with (5.146) shows that the time-sampled version of $\underline{\mathbf{S}}_{(tds)}(\tau,t)$ behaves as a discrete matrix filter with response

$$\underline{\mathbf{S}}(lT_s,kT_s) = \underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,k-r}\underset{r}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k} \tag{5.148}$$

We should observe the remarkable resemblance of (5.148) with (5.112). Since the role of $\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}$ is to discretize the domains of transmission, the role of $\underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,k-r}$ is to match the receiver to these discrete domains. This result could also have been easily determined using matrix operators. From (5.72) and (5.109) we have

$$\begin{aligned}\ddot{\underline{\mathbf{S}}} = \dot{\underline{\mathbf{H}}}^H\dot{\underline{\mathbf{H}}} &= \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}}\dot{\underline{\mathbf{\Phi}}}^H\dot{\underline{\mathbf{\Phi}}}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}} = \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}}\dot{\underline{\mathbf{\Omega}}}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}} = \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}\\ &\triangleq \ddot{\underline{\mathbf{S}}}(kT_s,lT_s) = \sum_r \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}r,k}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}r,l}\end{aligned} \tag{5.149}$$

which after some clear-cut manipulations yields

$$\begin{aligned}&\ddot{\underline{\mathbf{S}}}(kT_s,lT_s) = \sum_r \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}r,k}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}r,l} \Rightarrow \underline{\mathbf{S}}((k-l)T_s,kT_s) = \sum_r \underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}r-k,r}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}r-l,r}\\ &\overset{l\to k-l}{\Rightarrow} \underline{\mathbf{S}}(lT_s,kT_s) = \sum_r \underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,k-r}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l-r,k-r} = \underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,k-r}\underset{r}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k}\end{aligned} \tag{5.150}$$

One realizes once more that working with operators is much more intuitive and lean than working with integrals, sums and convolutions. After completing the deductions in the operator "world", the time domain results are derived with the utmost simplicity. The system diagram pertaining to these decompositions is shown in Figure 5.11.

5.4 *Frequency-Domain Description*

Describing the MIMO LTI system in the frequency domain is as easy as taking Fourier transforms of all vectors and matrices and then performing simple matrix-vector multiplications. The inherent simplicity results from three facts: 1) MIMO LTI systems are de-

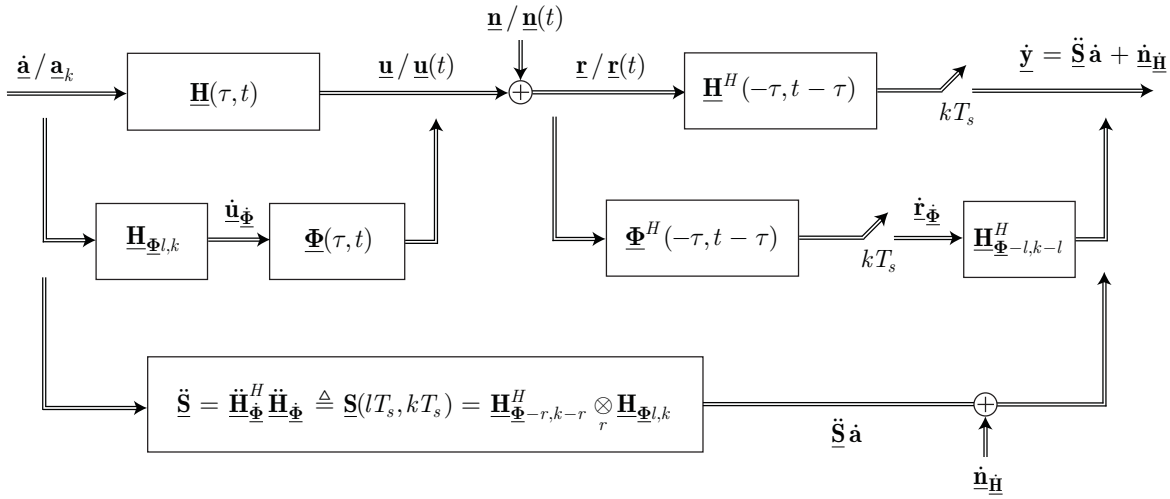

FIGURE 5.11 Decomposition of the TS-MMF and global function into discrete and semi-orthonormal filters.

scribed by Toeplitz matrix operators, i.e.

$$\begin{aligned}\underline{\mathbf{y}} &= \widehat{\underline{\mathbf{H}}}\underline{\mathbf{x}} = \widehat{\underline{\mathbf{H}}}_{Toeplitz}\underline{\mathbf{x}} \\ &\triangleq \underline{\mathbf{y}}(t) = \int_{t'} \widehat{\underline{\mathbf{H}}}(t,t')\underline{\mathbf{x}}(t')dt' = \int_{t'} \widehat{\underline{\mathbf{H}}}(t-t')\underline{\mathbf{x}}(t')dt'\end{aligned} \tag{5.151}$$

2) the matrix valued exponentials $\mathbf{K}(t,f) = e^{j2\pi ft}\mathbf{I}$ are eigenfunctions of Toeplitz operators, i.e.

$$\begin{aligned}\int_{t'} \widehat{\underline{\mathbf{H}}}(t-t')\widehat{\mathbf{K}}(t',f)dt' &= \left[\int_{t'} \widehat{\underline{\mathbf{H}}}(t')\widehat{\mathbf{K}}^H(t',f)dt'\right]\widehat{\mathbf{K}}(t,f) = \mathbf{\Lambda}(f)\widehat{\mathbf{K}}(t,f) \\ &\triangleq \widehat{\underline{\mathbf{H}}}_{Toeplitz}\widehat{\mathbf{K}} = \left[\widehat{\underline{\mathbf{H}}}_{col-const}\widehat{\mathbf{K}}^H\right]\cdot\widehat{\mathbf{K}} = \mathbf{\Lambda}(f)\cdot\widehat{\mathbf{K}}\end{aligned} \tag{5.152}$$

where the following identities have been used,

$$\begin{aligned}\widehat{\mathbf{K}}(t',f) &= \widehat{\mathbf{K}}(t'-t,f)\widehat{\mathbf{K}}(t,f) \\ \widehat{\mathbf{K}}(-t',f) &= \widehat{\mathbf{K}}^H(t',f)\end{aligned} \tag{5.153}$$

and 3) these eigenfunctions are fully orthonormal, i.e.

$$\begin{aligned}\widehat{\mathbf{K}}^H\widehat{\mathbf{K}} &= \hat{\mathbf{I}} \triangleq \int_t \widehat{\mathbf{K}}^H(t,f)\widehat{\mathbf{K}}(t,f')dt = \delta(f'-f)\mathbf{I} \\ \widehat{\mathbf{K}}\widehat{\mathbf{K}}^H &= \hat{\mathbf{I}} \triangleq \int_f \widehat{\mathbf{K}}(t,f)\widehat{\mathbf{K}}^H(t',f)df = \delta(t'-t)\mathbf{I}\end{aligned} \tag{5.154}$$

This last property means that they losslessly expand any finite energy vector signal as

$$\begin{aligned}\underline{\mathbf{x}} = \widehat{\mathbf{K}}\underset{\mathcal{F}}{\underline{\mathbf{x}}} &\triangleq \underline{\mathbf{x}}(t) = \int_f \widehat{\mathbf{K}}(t,f)\underset{\mathcal{F}}{\underline{\mathbf{x}}}(f)df \\ \underset{\mathcal{F}}{\underline{\mathbf{x}}} = \widehat{\mathbf{K}}^H\underline{\mathbf{x}} &\triangleq \underset{\mathcal{F}}{\underline{\mathbf{x}}}(f) = \int_t \widehat{\mathbf{K}}^H(t,f)\underline{\mathbf{x}}(t)dt\end{aligned} \tag{5.155}$$

and provide the means to transform operator multiplication into plain multiplication

$$\begin{aligned}\underline{\mathbf{y}} &= \widehat{\underline{\mathbf{H}}}_{Toeplitz}\underline{\mathbf{x}} = \widehat{\underline{\mathbf{H}}}_{Toeplitz}\widehat{\mathbf{K}}\underset{\mathcal{F}}{\underline{\mathbf{x}}} = \mathbf{\Lambda}(f)\cdot\widehat{\mathbf{K}}\underset{\mathcal{F}}{\underline{\mathbf{x}}} \\ \underset{\mathcal{F}}{\underline{\mathbf{y}}} &= \widehat{\mathbf{K}}^H\underline{\mathbf{y}} = \mathbf{\Lambda}(f)\cdot\widehat{\mathbf{K}}^H\widehat{\mathbf{K}}\underset{\mathcal{F}}{\underline{\mathbf{x}}} = \mathbf{\Lambda}(f)\cdot\underset{\mathcal{F}}{\underline{\mathbf{x}}} \triangleq \underset{\mathcal{F}}{\underline{\mathbf{y}}}(f) = \mathbf{\Lambda}(f)\underset{\mathcal{F}}{\underline{\mathbf{x}}}(f)\end{aligned} \tag{5.156}$$

Unfortunately, and to the best of the author's knowledge, when the operators are not Toeplitz there are no universal eigenfunctions known to guarantee a property equivalent to (5.156). There are some results available in the literature, namely chirp eigenfunctions [100] and time-frequency distributions [101], [102], that try to circumvent this problem but fail to do so in a satisfying manner. In the first case, chirp eigenfunctions are made available to a class of very simple, two-ray flat fading channel models, and in the second case it is found that the time-frequency version of (5.156) is only a good approximation for underspread channels. The difficulty surrounding LTV channels is that they also spread the input in frequency, not only in time, resulting in a convolution in both domains. Accounting for these limitations, the frequency-domain characterization of the LTV MIMO channel will be addressed by trying to extend the operator approach to the two-dimensional channel Fourier transform

$$\underline{\mathbf{H}}(\tau,t) \overset{\mathcal{F}}{\rightarrow} \underset{\mathcal{F}}{\underline{\mathbf{H}}}(f,\nu) = \int_{\tau}\int_{t} \underline{\mathbf{H}}(\tau,t)e^{-j2\pi(f\tau+\nu t)}dtd\tau \tag{5.157}$$

where the underscript $\mathcal{F}$ denotes that we are working in the Fourier domain.

5.4.1 FREQUENCY-DOMAIN OPERATORS

From Section 5.2 it is known that delay convolution with a generic $\underline{\mathbf{H}}(\tau,t)$ has the following frequency-domain counterpart

$$\underset{\mathcal{F}}{\underline{\mathbf{y}}}(f) = \int_{\nu} \underset{\mathcal{F}}{\underline{\mathbf{H}}}(f-\nu,\nu)\underset{\mathcal{F}}{\underline{\mathbf{x}}}(f-\nu)d\nu = \underset{\mathcal{F}}{\underline{\mathbf{H}}}(f-\nu,\nu)\underset{\nu}{\otimes}\underset{\mathcal{F}}{\underline{\mathbf{x}}}(f) \tag{5.158}$$

where the convolution is now in the Doppler variable ν. If the channels' time variations are statistically modelled, then the Fourier transform in the Doppler variable must also be modelled in statistical terms. The random matrix $\underset{\mathcal{F}}{\underline{\mathbf{H}}}(f,\nu)$,

$$\underset{\mathcal{F}}{\underline{\mathbf{H}}}(f,\nu) = \int_{t} \underset{\mathcal{F}}{\underline{\mathbf{H}}}(f,t)e^{-j2\pi\nu t}dt \tag{5.159}$$

is such that $E\left[\left|\underset{\mathcal{F}}{\mathbf{H}}_{ij}(f,t)\right|^2\right] < \infty$ and

$$\underset{\mathcal{F}}{\widehat{\mathbf{H}}}_{ij}(f,t) = \int_{\nu} \underset{\mathcal{F}}{\mathbf{H}}_{ij}(f,\nu)e^{j2\pi\nu t}d\nu \rightarrow E\left[\left|\underset{\mathcal{F}}{\mathbf{H}}_{ij}(f,t) - \underset{\mathcal{F}}{\widehat{\mathbf{H}}}_{ij}(f,t)\right|^2\right] = 0 \tag{5.160}$$

It is informative to graphically depict the convolution in (5.158) to understand how it works. This can be done by noticing that for a fixed f the graph of $\underset{\mathcal{F}}{\underline{\mathbf{H}}}(f-\nu,\nu)$ varies along lines of equation $g(\nu) = -\nu + f$, as shown in Figure 5.12. The interval of convolution is at most $2\sqrt{2}\nu_{\max}$, which is the maximum span of intersection of the lines with the domain of $\underset{\mathcal{F}}{\underline{\mathbf{H}}}(f,\nu)$. We see that whenever $\nu_{\max} > 0$ the output at f will be the contribution of all input frequencies within Doppler range, effectively embedding the input with inter-frequency interference (IFI) that is responsible for time-fading. Due to the skewed nature of the convolution, the bandwidth of the output signal will always be higher than that of the input,

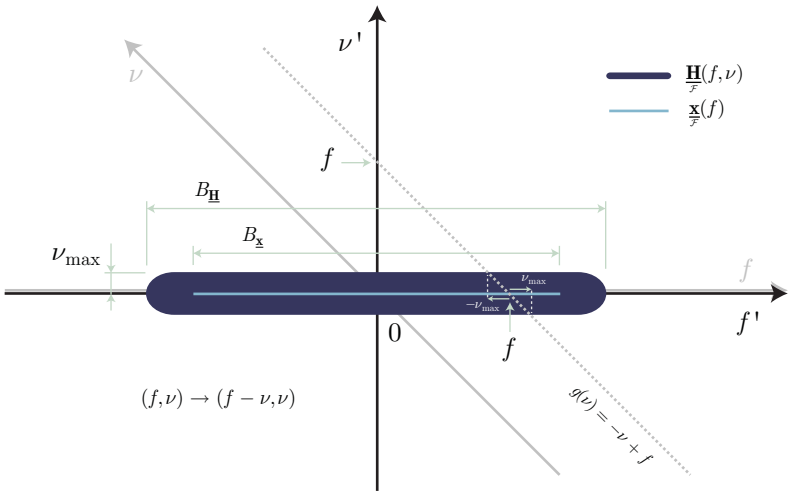

FIGURE 5.12 Skewed Doppler convolution induced by LTV MIMO channels.

and at the most

$$B_{\underline{\mathbf{y}}} \leq \min\left(B_{\underline{\mathbf{x}}}, B_{\underline{\mathbf{H}}}\right) + 2\nu_{\max} \tag{5.161}$$

negatively impacting the fidelity of the original signal (specially if it is narrowband). When the TF is time-invariant ($\nu_{\max} = 0$) the Doppler frequency response is an impulse function

$$\underline{\mathbf{H}}_{\mathcal{F}_{(LTI)}}(f-\nu,\nu) = \underline{\mathbf{H}}_{\mathcal{F}_{(LTI)}}(f-\nu)\delta(\nu) = \underline{\mathbf{H}}_{\mathcal{F}_{(LTI)}}(f)\delta(\nu) \tag{5.162}$$

so that the frequency output is the typical product

$$\underline{\mathbf{y}}_{\mathcal{F}}(f) = \int_{\nu} \underline{\mathbf{H}}_{\mathcal{F}_{LTI}}(f)\delta(\nu)\underline{\mathbf{x}}_{\mathcal{F}}(f-\nu)d\nu = \underline{\mathbf{H}}_{\mathcal{F}_{LTI}}(f)\underline{\mathbf{x}}_{\mathcal{F}}(f) \tag{5.163}$$

Setting $\nu = f - f'$ in (5.158) and defining the counterpart of (5.51) as

$$\widehat{\underline{\mathbf{H}}}_{\mathcal{F}}(f,-f') = \underline{\mathbf{H}}_{\mathcal{F}}(f',f-f') \Leftrightarrow \underline{\mathbf{H}}_{\mathcal{F}}(-f,f') = \widehat{\underline{\mathbf{H}}}_{\mathcal{F}}(f+f',f) \tag{5.164}$$

the first can be restated as follows

$$\begin{aligned} \underline{\mathbf{y}}_{\mathcal{F}}(f) &= \int_{f'} \underline{\mathbf{H}}_{\mathcal{F}}(f',f-f')\underline{\mathbf{x}}_{\mathcal{F}}(f')df' = \int_{f'} \widehat{\underline{\mathbf{H}}}_{\mathcal{F}}(f,-f')\underline{\mathbf{x}}_{\mathcal{F}}(f')df' \\ &\triangleq \underline{\mathbf{y}}_{\mathcal{F}} = \widehat{\underline{\mathbf{H}}}_{\mathcal{F}}\underline{\mathbf{x}}_{\mathcal{F}} \end{aligned} \tag{5.165}$$

Using the time-domain definition in (5.51) one can easily show that $\widehat{\underline{\mathbf{H}}}_{\mathcal{F}}(f,f')$ is the Fourier transform of $\widehat{\underline{\mathbf{H}}}(t,t')$,

$$\begin{aligned} \mathcal{F}\left\{\widehat{\underline{\mathbf{H}}}(t,t')\right\} &= \int_t \int_{t'} \underline{\mathbf{H}}(t-t',t)e^{-j2\pi(ft+f't')}dtdt' = \int_t \int_{t'} \underline{\mathbf{H}}(t,t')e^{-j2\pi((f+f')t'-f't)}dtdt' \\ &= \underline{\mathbf{H}}_{\mathcal{F}}(-f',f+f') = \widehat{\underline{\mathbf{H}}}_{\mathcal{F}}(f,f') \qquad \Rightarrow \qquad \widehat{\underline{\mathbf{H}}}(t,t') \overset{\mathcal{F}}{\to} \widehat{\underline{\mathbf{H}}}_{\mathcal{F}}(f,f') \end{aligned} \tag{5.166}$$

Using the two-sided-continuous frequency operator $\widehat{\underline{\mathbf{H}}}_{\mathcal{F}}$ we have effectively replaced frequency convolutions by linear operator products. In fact, owing to the continuity of the discrete-time fourier transform of discrete-time signals, this approach can be easily ex-

tended to one-sided-discrete operators: 1)

$$\begin{aligned}\underline{\mathbf{y}} = \underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{x}}} \overset{\mathcal{F}}{\triangleq} \underset{\mathcal{F}}{\mathbf{y}} = \underset{\mathcal{F}}{\dot{\mathbf{H}}}\underset{\mathcal{F}}{\dot{\mathbf{x}}} \triangleq \underset{\mathcal{F}}{\mathbf{y}}(f) = f_s^{-1}\int_{f'/f_s} \underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,-f')\underset{\mathcal{F}}{\dot{\mathbf{x}}}(f')df' \\ \underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,-f') = \underset{\mathcal{F}}{\mathbf{H}}\big(f', e^{j2\pi(f-f')T_s}\big) \qquad \underset{\mathcal{F}}{\dot{\mathbf{x}}}(f') = \underset{\mathcal{F}}{\mathbf{x}}\big(e^{j2\pi f'T_s}\big)\end{aligned} \tag{5.167}$$

and 2)

$$\begin{aligned}\underline{\dot{\mathbf{y}}} = \underline{\dot{\mathbf{H}}}\underline{\mathbf{x}} \overset{\mathcal{F}}{\triangleq} \underset{\mathcal{F}}{\mathbf{y}} = \underset{\mathcal{F}}{\dot{\mathbf{H}}}\underset{\mathcal{F}}{\dot{\mathbf{x}}} \triangleq \underset{\mathcal{F}}{\mathbf{y}}(f) = \int_{f'} \underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,-f')\underset{\mathcal{F}}{\mathbf{x}}(f')df' \\ \underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,-f') = \underset{\mathcal{F}}{\mathbf{H}}\big(e^{j2\pi f'T_s}, f-f'\big) \qquad \underset{\mathcal{F}}{\dot{\mathbf{y}}}(f') = \underset{\mathcal{F}}{\mathbf{y}}\big(e^{j2\pi f'T_s}\big)\end{aligned} \tag{5.168}$$

and two-sided-discrete operators

$$\begin{aligned}\underline{\ddot{\mathbf{y}}} = \underline{\ddot{\mathbf{H}}}\underline{\dot{\mathbf{x}}} \overset{\mathcal{F}}{\triangleq} \underset{\mathcal{F}}{\dot{\mathbf{y}}} = \underset{\mathcal{F}}{\ddot{\mathbf{H}}}\underset{\mathcal{F}}{\dot{\mathbf{x}}} \triangleq \underset{\mathcal{F}}{\dot{\mathbf{y}}}(f) = f_s^{-1}\int_{f'/f_s} \underset{\mathcal{F}}{\ddot{\mathbf{H}}}(f,-f')\underset{\mathcal{F}}{\dot{\mathbf{x}}}(f')df' \\ \underset{\mathcal{F}}{\ddot{\mathbf{H}}}(f,-f') = \underset{\mathcal{F}}{\mathbf{H}}\big(e^{j2\pi f'T_s}, e^{j2\pi(f-f')T_s}\big)\end{aligned} \tag{5.169}$$

where the integration in the first and latter case is over one period f_s. The latter case is intuitive because it is the discrete counterpart of two-sided-continuous operators, however the first and second cases are a bit harder because they involve hybrid operators. We shall prove the first. Starting with the system

$$\underline{\mathbf{y}}(t) = \sum_l \underline{\mathbf{H}}(t-lT_s,t)\underline{\mathbf{x}}_l = \sum_l \underline{\dot{\mathbf{H}}}(t,lT_s)\underline{\mathbf{x}}_l \tag{5.170}$$

and the inverse discrete-time Fourier transform (IDTFT)

$$\underline{\mathbf{x}}_l = f_s^{-1}\int_{f'/f_s} \underline{\mathbf{x}}\big(e^{j2\pi\nu T_s}\big)e^{j2\pi\nu lT_s}d\nu \tag{5.171}$$

we have

$$\begin{aligned}\underline{\mathbf{y}}(t) = f_s^{-1}\int_{f'/f_s}\sum_l \underline{\dot{\mathbf{H}}}(t,lT_s)e^{j2\pi f'lT_s}\underline{\mathbf{x}}\big(e^{j2\pi f'T_s}\big)df' = f_s^{-1}\int_{f'/f_s}\underset{\mathcal{F}}{\dot{\mathbf{H}}}(t,-f')\underset{\mathcal{F}}{\dot{\mathbf{x}}}(f')df' \\ \rightarrow \underset{\mathcal{F}}{\mathbf{y}}(f) = f_s^{-1}\int_{f'/f_s}\underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,-f')\underset{\mathcal{F}}{\dot{\mathbf{x}}}(f')df' \triangleq \underset{\mathcal{F}}{\mathbf{y}} = \underset{\mathcal{F}}{\dot{\mathbf{H}}}\underset{\mathcal{F}}{\dot{\mathbf{x}}}\end{aligned} \tag{5.172}$$

Also, from (5.51), (5.166) and the IDTFT, we arrive at

$$\begin{aligned}\underline{\mathbf{H}}(t-t',t) = \underline{\dot{\mathbf{H}}}(t,t') = f_s^{-1}\int_f\int_{f'/f_s}\underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,f')e^{j2\pi(ft+f't')}df'df \\ \rightarrow \underline{\mathbf{H}}(t,t') = \underline{\dot{\mathbf{H}}}(t',t'-t) = f_s^{-1}\int_f\int_{f'/f_s}\underset{\mathcal{F}}{\dot{\mathbf{H}}}(f+f',-f)e^{j2\pi(ft+f't')}df'df\end{aligned} \tag{5.173}$$

which finally leads to the definition in (5.167)

$$\underset{\mathcal{F}}{\dot{\mathbf{H}}}(f+f',-f) = \underset{\mathcal{F}}{\mathbf{H}}\big(f,e^{j2\pi f'T_s}\big) \rightarrow \underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,-f') = \underset{\mathcal{F}}{\mathbf{H}}\big(f',e^{j2\pi(f-f')T_s}\big) \tag{5.174}$$

completing the proof. Again, and since the DTFT of a constant is given by

$$\sum_l e^{-j2\pi\nu lT_s} = f_s\sum_k \delta(\nu - kf_s) \tag{5.175}$$

a time-invariant TF is expressed as

$$\begin{aligned}\underset{\mathcal{F}}{\mathbf{H}}_{(LTI)}\left(f',e^{j2\pi(f-f')T_s}\right) &= \underset{\mathcal{F}}{\mathbf{H}}_{(LTI)}(\zeta)\,f_s\delta(f-\zeta)\underset{\zeta}{\otimes}\sum_k\delta(f'-kf_s)\mathbf{I}\\ &= f_s\sum_k\underset{\mathcal{F}}{\mathbf{H}}_{(LTI)}(f'-kf_s)\,\delta(f-(f'-kf_s))\end{aligned} \tag{5.176}$$

which substituted into (5.167) yields the anticipated frequency product

$$\begin{aligned}&\underset{\mathcal{F}}{\mathbf{y}}(f) = f_s^{-1}\int_{f'/f_s}\underset{\mathcal{F}}{\mathbf{H}}_{(LTI)}\left(f',e^{j2\pi(f-f')T_s}\right)\underset{\mathcal{F}}{\mathbf{x}}\left(e^{j2\pi f'T_s}\right)df'\\ &=\sum_k\int_{f'/f_s}\underset{\mathcal{F}}{\mathbf{H}}_{(LTI)}(f'-kf_s)\,\delta(f'-f-kf_s)\underset{\mathcal{F}}{\mathbf{x}}\left(e^{j2\pi f'T_s}\right)df'\overset{f'-f=kf_s=0}{=}\underset{\mathcal{F}}{\mathbf{H}}_{(LTI)}(f)\underset{\mathcal{F}}{\mathbf{x}}\left(e^{j2\pi fT_s}\right)\end{aligned} \tag{5.177}$$

5.4.2 The Discrete MIMO Model in the Space $\mathcal{R}\left(\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}\right)$

It has been shown in Section 5.3 that the MIMO noiseless output can be losslessly expanded in the space $\mathcal{R}\left(\underline{\dot{\mathbf{\Phi}}}\right)$ by the following time-domain pair

$$\underline{\mathbf{u}} = \underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\dot{\mathbf{b}}}_{opt} \qquad \underline{\dot{\mathbf{b}}}_{opt} = \underline{\dot{\mathbf{\Phi}}}^H\underline{\mathbf{u}} = \underline{\dot{\mathbf{\Phi}}}^H\underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} = \underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}\underline{\dot{\mathbf{a}}} \tag{5.178}$$

which, at the same time, decomposes the MIMO channel operator as $\underline{\dot{\mathbf{H}}} = \underline{\dot{\mathbf{\Phi}}}\underline{\ddot{\mathbf{H}}}_{\underline{\dot{\mathbf{\Phi}}}}$. Even though this operator product is undertaken in the time domain, it may also be carried out in the frequency domain by taking into account the rules established in Section 5.4.1. We thus restate (5.178) as follows

$$\underset{\mathcal{F}}{\mathbf{u}} = \underset{\mathcal{F}}{\dot{\mathbf{H}}}\underset{\mathcal{F}}{\dot{\mathbf{a}}} = \underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}\underset{\mathcal{F}}{\ddot{\mathbf{H}}}_{\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}}\underset{\mathcal{F}}{\dot{\mathbf{a}}} = \underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}\underset{\mathcal{F}}{\dot{\mathbf{b}}}_{opt} \qquad \underset{\mathcal{F}}{\dot{\mathbf{b}}}_{opt} = \underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}^H\underset{\mathcal{F}}{\mathbf{u}} = \underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}^H\underset{\mathcal{F}}{\dot{\mathbf{H}}}\underset{\mathcal{F}}{\dot{\mathbf{a}}} = \underset{\mathcal{F}}{\ddot{\mathbf{H}}}_{\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}}\underset{\mathcal{F}}{\dot{\mathbf{a}}} \tag{5.179}$$

In effect, we are expressing the frequency output $\underset{\mathcal{F}}{\mathbf{u}}$ in the space $\mathcal{R}\left(\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}\right)$. The channel decomposition is expanded as

$$\begin{aligned}\underset{\mathcal{F}}{\dot{\mathbf{H}}} = \underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}\underset{\mathcal{F}}{\ddot{\mathbf{H}}}_{\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}} \qquad &\triangleq \quad \underset{\mathcal{F}}{\dot{\mathbf{H}}}(f,-f') = f_s^{-1}\int_{f''/f_s}\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}(f,-f'')\underset{\mathcal{F}}{\ddot{\mathbf{H}}}_{\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}}(f'',-f')df''\\ \underset{\mathcal{F}}{\ddot{\mathbf{H}}}_{\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}} = \underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}^H\underset{\mathcal{F}}{\dot{\mathbf{H}}} \qquad &\triangleq \quad \underset{\mathcal{F}}{\ddot{\mathbf{H}}}_{\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}}(f,-f') = \int_{f''}\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}^H(f'',-f)\underset{\mathcal{F}}{\dot{\mathbf{H}}}(f'',-f')df''\end{aligned} \tag{5.180}$$

and we must have the orthogonality condition (identity operator)

$$\begin{aligned}\underset{\mathcal{F}}{\ddot{\mathbf{\Omega}}} = \underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}^H\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}} = \underline{\ddot{\mathbf{I}}} \triangleq \underset{\mathcal{F}}{\ddot{\mathbf{\Omega}}}(f,-f') &= \int_{f''}\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}^H(f'',-f)\underset{\mathcal{F}}{\dot{\mathbf{\Phi}}}(f'',-f')df''\\ &= f_s\delta(f-\zeta)\underset{\zeta}{\otimes}\sum_k\delta(f'-kf_s)\mathbf{I} = f_s\sum_k\delta(f-f'-kf_s)\mathbf{I}\end{aligned} \tag{5.181}$$

or, from (5.167), equivalently

$$\begin{aligned}&\underset{\mathcal{F}}{\mathbf{\Omega}}\left(e^{j2\pi f'T_s},e^{j2\pi(f-f')T_s}\right) = \int_{f''}\underset{\mathcal{F}}{\mathbf{\Phi}}^H\left(f,e^{j2\pi(f''-f)T_s}\right)\underset{\mathcal{F}}{\mathbf{\Phi}}\left(f',e^{j2\pi(f''-f')T_s}\right)df''\\ &\overset{f'=f-\nu}{\rightarrow}\underset{\mathcal{F}}{\mathbf{\Omega}}\left(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\right)\overset{f''\rightarrow f-\zeta}{=}\int_{\zeta}\underset{\mathcal{F}}{\mathbf{\Phi}}^H\left(f,e^{-j2\pi\zeta T_s}\right)\underset{\mathcal{F}}{\mathbf{\Phi}}\left(f-\nu,e^{j2\pi(\nu-\zeta)T_s}\right)d\zeta\\ &\qquad = \underset{\mathcal{F}}{\mathbf{\Phi}}^H\left(f,e^{-j2\pi\zeta T_s}\right)\underset{\zeta}{\otimes}\underset{\mathcal{F}}{\mathbf{\Phi}}\left(f-\nu,e^{j2\pi\nu T_s}\right) = f_s\sum_k\delta(\nu-kf_s)\mathbf{I}\end{aligned} \tag{5.182}$$

One infers that if orthonormality is present in the time domain it must also be present in the frequency domain, and operators smooth the path to the orthogonality conditions. It is interesting to narrow (5.182) to the LTI scenario and see if the result is the one expected. Applying (5.176)

$$\begin{aligned}
&\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}\big(f,e^{-j2\pi\zeta T_s}\big)=f_s\sum_k\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}(f-kf_s)\,\delta(-\zeta+kf_s)\\
&\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}\big(f-\nu,e^{j2\pi(\nu-\zeta)T_s}\big)=f_s\sum_m\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}(f-\nu-mf_s)\,\delta(\nu-\zeta+mf_s)\\
&\underset{\mathcal{F}}{\boldsymbol{\Omega}}_{(LTI)}\big(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\big)=f_s\sum_k\underset{\mathcal{F}}{\boldsymbol{\Omega}}_{(LTI)}\big(e^{j2\pi fT_s}\big)\,\delta(\nu-kf_s)
\end{aligned}\tag{5.183}$$

one recovers

$$\begin{aligned}
&f_s\sum_k\underset{\mathcal{F}}{\boldsymbol{\Omega}}_{(LTI)}\big(e^{j2\pi fT_s}\big)\,\delta(\nu-kf_s)=\int_\zeta\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}\big(f,e^{-j2\pi\zeta T_s}\big)\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}\big(f-\nu,e^{j2\pi(\nu-\zeta)T_s}\big)\,d\zeta\\
&=f_s\sum_k f_s\sum_m\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}(f-kf_s)\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}(f-\nu-mf_s)\int_\zeta\delta(\zeta-kf_s)\delta(\nu-\zeta+mf_s)d\zeta\\
&\overset{\zeta=kf_s\,/\,\nu=kf_s-mf_s}{=}f_s\sum_k\left[f_s\sum_m\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}(f-mf_s)\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}(f-mf_s)\right]\delta(\nu-kf_s)
\end{aligned}\tag{5.184}$$

which, comparing with (5.182), implies

$$\underset{\mathcal{F}}{\boldsymbol{\Omega}}_{(LTI)}\big(e^{j2\pi fT_s}\big)=f_s\sum_m\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}(f-mf_s)\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}(f-mf_s)=\mathbf{I}\tag{5.185}$$

This expression is no less than the sampled version of $\underset{\mathcal{F}}{\boldsymbol{\Omega}}_{(LTI)}(f)=\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}(f)\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}(f)$, and resembles the Nyquist condition for transmission without intersymbol interference (ISI). It may also be easily retrieved from the initial orthogonality relationship in (5.57), restricted to time-invariancy

$$\begin{aligned}
&\underline{\boldsymbol{\Omega}}_{(LTI)}(lT_s)=\int_\tau\underline{\boldsymbol{\Phi}}^{H}_{(LTI)}(\tau-kT_s)\underline{\boldsymbol{\Phi}}_{(LTI)}(\tau-(k-l)T_s)d\tau=\delta_l\mathbf{I}\\
&\rightarrow\underline{\boldsymbol{\Omega}}_{(LTI)}(t)=\int_\tau\underline{\boldsymbol{\Phi}}^{H}_{(LTI)}(-\tau)\underline{\boldsymbol{\Phi}}_{(LTI)}(t-\tau)d\tau=\underline{\boldsymbol{\Phi}}^{H}_{(LTI)}(-\tau)\underset{\tau}{\otimes}\underline{\boldsymbol{\Phi}}_{(LTI)}(t)
\end{aligned}\tag{5.186}$$

Now, taking the Fourier transform and sampling we find

$$\begin{aligned}
&\underset{\mathcal{F}}{\boldsymbol{\Omega}}_{(LTI)}(f)=\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}(f)\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}(f)\\
&\rightarrow\underset{\mathcal{F}}{\boldsymbol{\Omega}}_{(LTI)}\big(e^{j2\pi fT_s}\big)=f_s\sum_m\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}_{(LTI)}(f-mf_s)\underset{\mathcal{F}}{\boldsymbol{\Phi}}_{(LTI)}(f-mf_s)=\mathbf{I}
\end{aligned}\tag{5.187}$$

which is exactly (5.185). In light of these deductions one confirms that the time-invariant case can be obtained directly from the time-varying formulas by setting some fundamental restrictions such as (5.176).

It may be finally checked by similar derivations that the channel decomposes as

$$\underset{\mathcal{F}}{\mathbf{H}}\big(f-\nu,e^{j2\pi\nu T_s}\big)=\underset{\mathcal{F}}{\boldsymbol{\Phi}}\big(f-\zeta,e^{j2\pi\zeta T_s}\big)\underset{\zeta\,/\,f_s}{\otimes}\underset{\mathcal{F}\,\underset{\mathcal{F}}{\boldsymbol{\Phi}}}{\mathbf{H}}\big(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\big)\tag{5.188}$$

and its discrete version equals

$$\underset{\mathcal{F}}{\mathbf{H}}_{\underset{\mathcal{F}}{\boldsymbol{\Phi}}}\left(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\right)=\underset{\mathcal{F}}{\boldsymbol{\Phi}}^{H}\left(f,e^{-j2\pi\zeta T_s}\right)\underset{\zeta}{\otimes}\underset{\mathcal{F}}{\mathbf{H}}\left(f-\nu,e^{j2\pi\nu T_s}\right) \tag{5.189}$$

supplying us with sufficient information to construct a diagram (Figure 5.13) for the time-varying (Doppler-spread) MIMO discretization in the frequency domain.

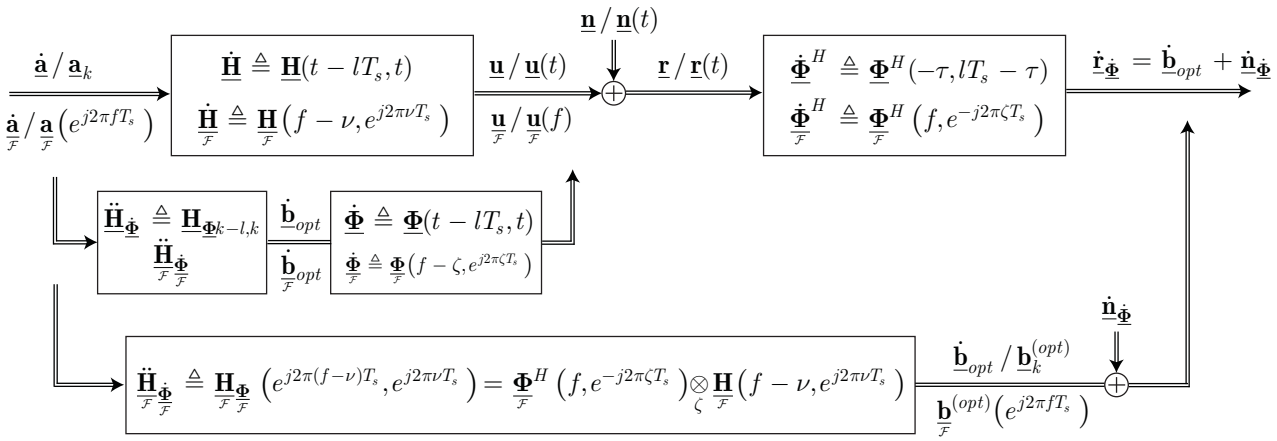

FIGURE 5.13 Frequency perspective of the discrete MIMO model after channel decomposition and output projection onto $\mathcal{R}\left(\underset{\mathcal{F}}{\dot{\boldsymbol{\Phi}}}\right)$.

5.4.3 THE TS-MMF IN THE FREQUENCY DOMAIN

It was found in Section 5.3.3 that the continuous matrix matched filter to the channel $\underline{\mathbf{H}}(\tau,t)$ was the filter $\underline{\mathbf{MMF}}(\tau,t)=\underline{\mathbf{H}}^{H}(-\tau,t-\tau)$. To find its frequency-domain counterpart $\underset{\mathcal{F}}{\underline{\mathbf{MMF}}}(f-\nu,\nu)$ one can proceed in one of two manners. The first is to begin with the two-dimensional inverse Fourier transform

$$\underline{\mathbf{H}}(\tau,t)=\int_f\int_\nu\underset{\mathcal{F}}{\underline{\mathbf{H}}}(f,\nu)e^{j2\pi(f\tau+\nu t)}d\nu df \tag{5.190}$$

and make some variable changes

$$\begin{aligned}\underline{\mathbf{H}}^{H}(-\tau,t-\tau)&=\int_f\int_\nu\underset{\mathcal{F}}{\underline{\mathbf{H}}}^{H}(f,\nu)e^{j2\pi(f\tau-\nu t+\nu\tau)}d\nu df\\&=\int_f\int_\nu\underset{\mathcal{F}}{\underline{\mathbf{H}}}^{H}(f+\nu,-\nu)e^{j2\pi(f\tau+\nu t)}d\nu df\end{aligned} \tag{5.191}$$

from which we retrieve

$$\underset{\mathcal{F}}{\underline{\mathbf{MMF}}}(f-\nu,\nu)=\underset{\mathcal{F}}{\underline{\mathbf{H}}}^{H}(f,-\nu) \tag{5.192}$$

The alternative method is to remember that the MMF in the time-domain corresponds to the Hermitian adjoint operator in the operator domain, and the same happens in the frequency domain. If $\underset{\mathcal{F}}{\widehat{\underline{\mathbf{H}}}}$ denotes the two-sided-continuous channel frequency operator, then

$$\underset{\mathcal{F}}{\mathbf{y}}=\underset{\mathcal{F}}{\widehat{\underline{\mathbf{H}}}}^{H}\underset{\mathcal{F}}{\mathbf{r}}\triangleq\underset{\mathcal{F}}{\mathbf{y}}(f)=\int_{f'}\underset{\mathcal{F}}{\widehat{\underline{\mathbf{H}}}}^{H}(f',-f)\underset{\mathcal{F}}{\underline{\mathbf{r}}}(f')df' \tag{5.193}$$

Now we set $f' = f - \nu$ and recall (5.164) to get

$$\begin{aligned}\underline{\underline{\mathbf{y}}}_{\mathcal{F}}(f) &= \int_{f'} \underline{\underline{\widehat{\mathbf{H}}}}^{H}_{\mathcal{F}}(f',-f)\,\underline{\underline{\mathbf{r}}}_{\mathcal{F}}(f')df' = \int_{f'} \underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}(f,f'-f)\,\underline{\underline{\mathbf{r}}}_{\mathcal{F}}(f')df' \\ &\overset{f'\to f-\nu}{=} \int_{\nu} \underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}(f,-\nu)\,\underline{\underline{\mathbf{r}}}_{\mathcal{F}}(f-\nu)d\nu = \underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}(f,-\nu)\underset{\nu}{\otimes}\underline{\underline{\mathbf{r}}}_{\mathcal{F}}(f)\end{aligned} \tag{5.194}$$

which, in light of (5.158), gives the frequency expression for the TS-MMF as in (5.192), and yields a (skewless) convolution that is very similar to the delay convolution in (5.50). This means that the TS-MMF has the desirable properties of not introducing IFI and not expanding the system bandwidth (either input or channel limited). After sampling the output the TS-MMF behaves as

$$\underline{\underline{\mathbf{H}}}^{H}(-\tau,lT_s-\tau) \to f_s\sum_{k}\underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}(f,-(\nu-kf_s)) = \underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}\left(f,e^{-j2\pi\nu T_s}\right) \tag{5.195}$$

a result that might also have been directly retrieved from

$$\underline{\underline{\dot{\mathbf{y}}}}_{\mathcal{F}}(f) = \int_{f'} \underline{\underline{\dot{\mathbf{H}}}}^{H}_{\mathcal{F}}(f',-f)\,\underline{\underline{\mathbf{r}}}_{\mathcal{F}}(f')df' = \int_{\nu} \underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}\left(f,e^{-j2\pi\nu T_s}\right)\underline{\underline{\mathbf{r}}}_{\mathcal{F}}(f-\nu)d\nu \tag{5.196}$$

To decompose the TS-MMF we start from

$$\underline{\underline{\dot{\mathbf{H}}}}^{H}_{\mathcal{F}} = \underline{\underline{\ddot{\mathbf{H}}}}^{H}_{\mathcal{F}}\underline{\underline{\dot{\mathbf{\Phi}}}}^{H}_{\mathcal{F}} \triangleq \underline{\underline{\dot{\mathbf{H}}}}^{H}_{\mathcal{F}}(f',-f) = f_s^{-1}\int_{f''/f_s}\underline{\underline{\ddot{\mathbf{H}}}}^{H}_{\mathcal{F}}(f'',-f)\,\underline{\underline{\dot{\mathbf{\Phi}}}}^{H}_{\mathcal{F}}(f',-f'')df'' \tag{5.197}$$

apply (5.167) and (5.169), and set $f' = f - \nu$ and $f'' = f - \zeta$, finding

$$\begin{aligned}\underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}\left(f,e^{-j2\pi\nu T_s}\right) &= f_s^{-1}\int_{\zeta/f_s}\underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}\left(e^{j2\pi fT_s},e^{-j2\pi\zeta T_s}\right)\underline{\underline{\mathbf{\Phi}}}^{H}_{\mathcal{F}}\left(f-\zeta,e^{-j2\pi(\nu-\zeta)T_s}\right)d\zeta \\ &= \underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}\left(e^{j2\pi fT_s},e^{-j2\pi\zeta T_s}\right)\underset{\zeta/f_s}{\otimes}\underline{\underline{\mathbf{\Phi}}}^{H}_{\mathcal{F}}\left(f,e^{-j2\pi\nu T_s}\right)\end{aligned} \tag{5.198}$$

From (5.198) one realizes that the frequency-domain TS-MMF decomposes into two matched filters, one matched to the discrete filter $\underline{\underline{\mathbf{H}}}_{\mathcal{F}}$ and the other matched to the orthonormalized version of the channel $\underline{\underline{\mathbf{\Phi}}}_{\mathcal{F}}$, as already obtained in (5.189). This is in consonance with the time-domain results.

To conclude the frequency-domain interpretation of the TS-MMF approach to discretization, we set the aim of finding the global input-output transfer function. Again, working with operators and using the orthogonality condition in (5.181) one easily attains

$$\underline{\underline{\ddot{\mathbf{S}}}}_{\mathcal{F}} = \underline{\underline{\dot{\mathbf{H}}}}^{H}_{\mathcal{F}}\,\underline{\underline{\dot{\mathbf{H}}}}_{\mathcal{F}} = \underline{\underline{\ddot{\mathbf{H}}}}^{H}_{\mathcal{F}}\underline{\underline{\dot{\mathbf{\Phi}}}}^{H}_{\mathcal{F}}\,\underline{\underline{\dot{\mathbf{\Phi}}}}_{\mathcal{F}}\underline{\underline{\ddot{\mathbf{H}}}}_{\mathcal{F}} = \underline{\underline{\ddot{\mathbf{H}}}}^{H}_{\mathcal{F}}\underline{\underline{\ddot{\mathbf{\Omega}}}}_{\mathcal{F}}\underline{\underline{\ddot{\mathbf{H}}}}_{\mathcal{F}} = \underline{\underline{\ddot{\mathbf{H}}}}^{H}_{\mathcal{F}}\underline{\underline{\ddot{\mathbf{H}}}}_{\mathcal{F}} \tag{5.199}$$

which expands as

$$\underline{\underline{\ddot{\mathbf{S}}}}_{\mathcal{F}}(f,-f') = f_s^{-1}\int_{f''/f_s}\underline{\underline{\ddot{\mathbf{H}}}}^{H}_{\mathcal{F}}(f'',-f)\,\underline{\underline{\ddot{\mathbf{H}}}}_{\mathcal{F}}(f'',-f')df'' \tag{5.200}$$

and leaves the operator domain in the following form

$$\begin{aligned}\underline{\underline{\mathbf{S}}}_{\mathcal{F}}\left(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\right) &= f_s^{-1}\int_{\zeta/f_s}\underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}\left(e^{j2\pi fT_s},e^{-j2\pi\zeta T_s}\right)\underline{\underline{\mathbf{H}}}_{\mathcal{F}}\left(e^{j2\pi(f-\nu)T_s},e^{j2\pi(\nu-\zeta)T_s}\right)d\zeta \\ &= \underline{\underline{\mathbf{H}}}^{H}_{\mathcal{F}}\left(e^{j2\pi fT_s},e^{-j2\pi\zeta T_s}\right)\underset{\zeta/f_s}{\otimes}\underline{\underline{\mathbf{H}}}_{\mathcal{F}}\left(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\right)\end{aligned} \tag{5.201}$$

The MIMO system after matched filtering and sampling is thus decomposed into a discrete, Doppler-spread filter and its respective discrete matrix matched filter, noticeably recalling the identical time-domain counterpart

$$\underline{\mathbf{S}}(lT_s,kT_s)=\underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}-r,k-r}\underset{r}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}l,k} \tag{5.202}$$

Whenever the condition

$$\underline{\mathbf{S}}(lT_s,kT_s)=\delta_l\mathbf{I}\ \sim\ \underset{\mathcal{F}}{\underline{\mathbf{S}}}\big(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\big)=f_s\sum_k\delta(\nu-kf_s)\mathbf{I} \tag{5.203}$$

is met, then there is linear separability of vector samples, so that there is neither ISI nor Doppler spread levied on the input symbols, only additive noise.

Let us inquire what happens when the channel is time-invariant. In light of (5.176) one infers

$$\begin{aligned}
\underset{\mathcal{F}}{\underline{\mathbf{S}}}_{(LTI)}\big(e^{j2\pi(f-\nu)T_s},e^{j2\pi\nu T_s}\big)&=f_s\sum_m\underset{\mathcal{F}}{\underline{\mathbf{S}}}_{(LTI)}\big(e^{j2\pi fT_s}\big)\delta(\nu-mf_s)\\
\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi fT_s},e^{-j2\pi\zeta T_s}\big)&=f_s\sum_k\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi(f-kf_s)T_s}\big)\delta(-\zeta+kf_s)\\
\underset{\mathcal{F}}{\underline{\mathbf{H}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi(f-\nu)T_s},e^{j2\pi(\nu-\zeta)T_s}\big)&=f_s\sum_m\underset{\mathcal{F}}{\underline{\mathbf{H}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi(f-\nu-mf_s)T_s}\big)\delta(\nu-\zeta+mf_s)
\end{aligned} \tag{5.204}$$

so that (5.201) is reexpressed as

$$\begin{aligned}
\sum_m\underset{\mathcal{F}}{\underline{\mathbf{S}}}_{(LTI)}\big(e^{j2\pi fT_s}\big)\delta(\nu-mf_s)&=\sum_k\sum_m\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi(f-kf_s)T_s}\big)\underset{\mathcal{F}}{\underline{\mathbf{H}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi(f-\nu-mf_s)T_s}\big)\\
&\qquad\cdot\int_{\zeta/f_s}\delta(-\zeta+kf_s)\delta(\nu-\zeta+mf_s)d\zeta\\
&\overset{\zeta=kf_s=0}{=}\sum_m\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi fT_s}\big)\underset{\mathcal{F}}{\underline{\mathbf{H}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi fT_s}\big)\delta(\nu-mf_s)
\end{aligned} \tag{5.205}$$

leading to the spectral product

$$\underset{\mathcal{F}}{\underline{\mathbf{S}}}_{(LTI)}\big(e^{j2\pi fT_s}\big)=\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi fT_s}\big)\underset{\mathcal{F}}{\underline{\mathbf{H}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi fT_s}\big) \tag{5.206}$$

which is exactly the DTFT of the time-invariant instance of (5.202)

$$\underline{\mathbf{S}}_{(LTI)}(lT_s)=\underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}(LTI)-r}\underset{r}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}(LTI)l}\ \overset{\mathcal{F}}{\rightarrow}\ \underset{\mathcal{F}}{\underline{\mathbf{S}}}_{(LTI)}\big(e^{j2\pi fT_s}\big)=\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi fT_s}\big)\underset{\mathcal{F}}{\underline{\mathbf{H}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}\big(e^{j2\pi fT_s}\big) \tag{5.207}$$

or, from the z-Transform,

$$\underline{\mathbf{S}}_{(LTI)}(lT_s)=\underline{\mathbf{H}}^H_{\underline{\mathbf{\Phi}}(LTI)-r}\underset{r}{\otimes}\underline{\mathbf{H}}_{\underline{\mathbf{\Phi}}(LTI)l}\ \overset{\mathcal{Z}}{\rightarrow}\ \underset{\mathcal{F}}{\underline{\mathbf{S}}}_{(LTI)}(z)=\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}(1/z^*)\underset{\mathcal{F}}{\underline{\mathbf{H}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}(LTI)}(z) \tag{5.208}$$

From the previous discussion in Section 5.3.3 it was acknowledged that $\ddot{\underline{\mathbf{S}}}$ is hermitian and non-negative definite, yet non-stationary. For the stationary case, multivariate spectral decompositions such as (5.206) (or (5.208)) are not really unexpected, and in fact have been around for quite a while since the time of Norbert Wiener [103], [104]. The accomplishment has been to extend the decomposition to the realm of non-stationary discrete operators, and the result is (5.201), not a product, but a convolution. The system diagram after these decompositions is depicted in Figure 5.14, where $\underset{\mathcal{F}}{\underline{\mathbf{r}}}(f)$, $\underset{\mathcal{F}}{\underline{\mathbf{y}}}(e^{j2\pi fT_s})$ and $\underset{\mathcal{F}}{\underline{\mathbf{r}}}_{\underset{\mathcal{F}}{\underline{\mathbf{\Phi}}}}(f)$ are nec-

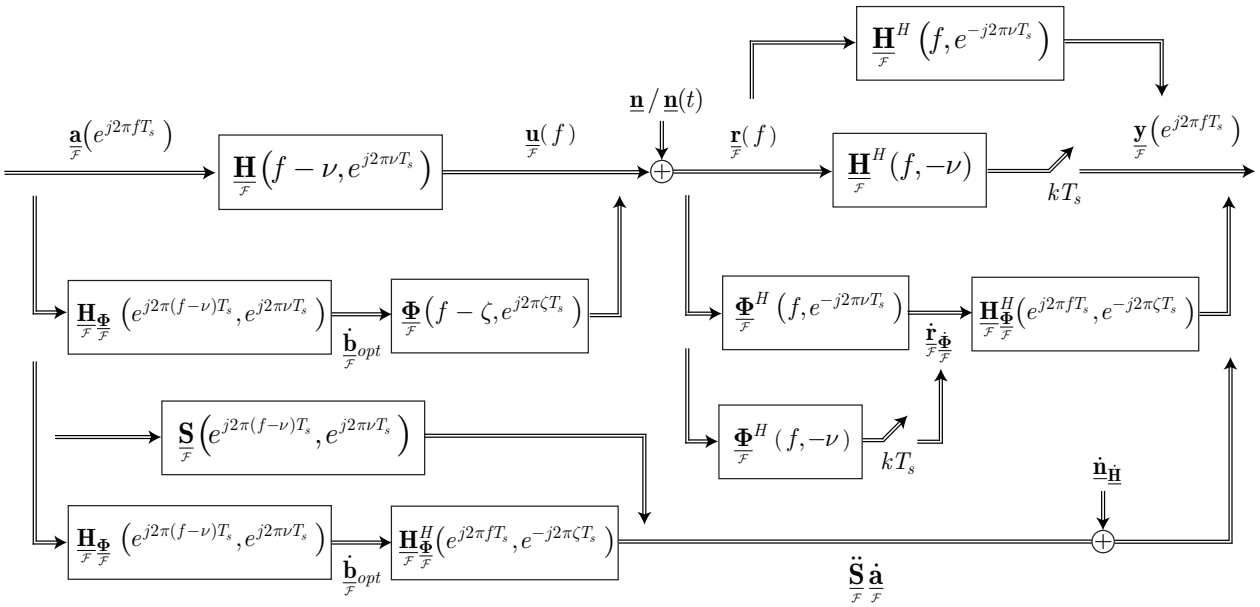

FIGURE 5.14 Discrete MIMO model after matched filtering and spectral decomposition.

essarily stochastic spectral representations of their time-domain counterparts.

5.5 *Noise Whitening and Full Discretization*

Two approaches to optimal detection in LTV MIMO, namely 1) the TS-MMF and 2) the orthonormalized TS-MMF, have been presented in Section 5.3, having the latter been favoured over the first. In the following sections we shall study the possibilities of 1) orthonormalizing the TS-MMF using a fully discrete matrix filter and 2) fully discretizing both the TS-MMF and the ORTHO-TS-MMF.

5.5.1 WHITENING THE NOISE AT THE TS-MMF OUTPUT

We have seen that the advantage of the ORTHO-TS-MMF was the uncorrelatedness of the noise samples at its output, because this meant a reduction in detection complexity. It was also shown that the ORTHO-TS-MMF $\dot{\underline{\mathbf{\Phi}}}^H$ originated from the channel decomposition into a causal $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}$ (that may or may not be invertible) and a semi-orthonormal $\dot{\underline{\mathbf{\Phi}}}$ that assured the white noise condition

$$\ddot{\underline{\mathbf{R}}}_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Phi}}}}\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Psi}}}}} = N_0 \dot{\underline{\mathbf{\Phi}}}^H \dot{\underline{\mathbf{\Phi}}} = N_0 \ddot{\underline{\mathbf{I}}} \tag{5.209}$$

At the output of the TS-MMF $\dot{\underline{\mathbf{H}}}^H$ the autocorrelation is

$$\ddot{\mathbf{R}}_{\dot{\underline{\mathbf{n}}}_{\ddot{\mathbf{H}}}\dot{\underline{\mathbf{n}}}_{\ddot{\mathbf{H}}}} = N_0 \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}} \dot{\underline{\mathbf{\Phi}}}^H \dot{\underline{\mathbf{\Phi}}} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}} = N_0 \ddot{\underline{\mathbf{S}}} \tag{5.210}$$

so it becomes natural to ask if there exists a filter $\ddot{\underline{\mathbf{\Upsilon}}}$ that, applied at the TS-MMF output, renders the noise white, i.e.

$$\ddot{\underline{\mathbf{R}}}_{\dot{\underline{\mathbf{n}}}_{\ddot{\mathbf{H}}}\ddot{\underline{\mathbf{\Upsilon}}}^H \dot{\underline{\mathbf{n}}}_{\ddot{\mathbf{H}}}\ddot{\underline{\mathbf{\Upsilon}}}^H} = N_0 \ddot{\underline{\mathbf{\Upsilon}}} \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}} \dot{\underline{\mathbf{\Phi}}}^H \dot{\underline{\mathbf{\Phi}}} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}} \ddot{\underline{\mathbf{\Upsilon}}}^H = N_0 \ddot{\underline{\mathbf{\Upsilon}}} \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}} \ddot{\underline{\mathbf{\Upsilon}}}^H = N_0 \ddot{\underline{\mathbf{I}}} \tag{5.211}$$

If such a filter $\ddot{\underline{\Upsilon}}$ indeed exists, it is fully discrete and must be such that

$$\left(\dot{\underline{\mathbf{H}}}\,\ddot{\underline{\Upsilon}}^H\right)^H \dot{\underline{\mathbf{H}}}\,\ddot{\underline{\Upsilon}}^H = \left(\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\,\ddot{\underline{\Upsilon}}^H\right)^H \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\,\ddot{\underline{\Upsilon}}^H = \dot{\underline{\Theta}}^H\dot{\underline{\Theta}} = \mathbf{I} \qquad \dot{\underline{\Theta}}^H = \ddot{\underline{\Upsilon}}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\dot{\underline{\Phi}}^H \tag{5.212}$$

We know, however, that, whenever $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$ is defined, the solution of minimum Frobenius norm of the system $\dot{\underline{\mathbf{H}}}^H = \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\dot{\underline{\Phi}}^H$ is obtained from the Moore-Penrose pseudoinverse of $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$, $\ddot{\underline{\mathbf{H}}}^{\dagger}_{\dot{\underline{\Phi}}}$, so we may test the choice of a filter that satisfies

$$\dot{\underline{\Theta}}^H = \ddot{\underline{\Upsilon}}\dot{\underline{\mathbf{H}}}^H \underset{\ddot{\underline{\Upsilon}}=\ddot{\underline{\mathbf{H}}}^{\dagger H}_{\dot{\underline{\Phi}}}}{=} \ddot{\underline{\mathbf{H}}}^{\dagger H}_{\dot{\underline{\Phi}}}\dot{\underline{\mathbf{H}}}^H = \ddot{\underline{\mathbf{H}}}^{\dagger H}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\dot{\underline{\Phi}}^H \tag{5.213}$$

To make things clearer, we now expand $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$ in a URV factorization, for instance, its singular value decomposition (SVD)

$$\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}} = \ddot{\underline{\mathbf{U}}}\ddot{\underline{\mathbf{D}}}\ddot{\underline{\mathbf{V}}}^H \quad \ddot{\underline{\mathbf{U}}}^H\ddot{\underline{\mathbf{U}}} = \ddot{\underline{\mathbf{V}}}^H\ddot{\underline{\mathbf{V}}} = \ddot{\underline{\mathbf{I}}} \qquad \mathcal{R}\left(\ddot{\underline{\mathbf{U}}}\right) = \mathcal{R}\left(\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\right); \mathcal{R}\left(\ddot{\underline{\mathbf{V}}}\right) = \mathcal{R}\left(\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\right) \tag{5.214}$$

from which we decompose the pseudoinverse as $\ddot{\underline{\mathbf{H}}}^{\dagger}_{\dot{\underline{\Phi}}} = \ddot{\underline{\mathbf{V}}}\ddot{\underline{\mathbf{D}}}^{-1}\ddot{\underline{\mathbf{U}}}^H$ and rewrite (5.213) in the form

$$\dot{\underline{\Theta}}^H = \ddot{\underline{\mathbf{H}}}^{\dagger H}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\dot{\underline{\Phi}}^H = \ddot{\underline{\mathbf{U}}}\ddot{\underline{\mathbf{D}}}^{-H}\ddot{\underline{\mathbf{V}}}^H\ddot{\underline{\mathbf{V}}}\ddot{\underline{\mathbf{D}}}^H\ddot{\underline{\mathbf{U}}}^H\dot{\underline{\Phi}}^H = \ddot{\underline{\mathbf{U}}}\ddot{\underline{\mathbf{U}}}^H\dot{\underline{\Phi}}^H \tag{5.215}$$

Consequently, a sufficient condition for the fulfilment of (5.212) is $\ddot{\underline{\mathbf{U}}}\ddot{\underline{\mathbf{U}}}^H = \ddot{\underline{\mathbf{I}}}$, which is satisfied if $\ddot{\underline{\mathbf{U}}}$ is square and non-singular, hence $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$ necessarily a full row-rank operator. This can also be seen from the full row-rank formula

$$\ddot{\underline{\mathbf{H}}}^{\dagger}_{\dot{\underline{\Phi}}} = \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\left(\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\right)^{-1} \quad \rightarrow \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^{\dagger}_{\dot{\underline{\Phi}}} = \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\left(\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\right)^{-1} = \ddot{\underline{\mathbf{I}}} \tag{5.216}$$

that shows that the pseudoinverse is a right inverse of $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$. In other words, the hermitian adjoint of the pseudoinverse is a left inverse of $\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}$ and $\ddot{\underline{\Upsilon}} = \ddot{\underline{\mathbf{H}}}^{\dagger H}_{\dot{\underline{\Phi}}}$ whitens the noise at the TS-MMF output

$$\ddot{\underline{\mathbf{R}}}_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Theta}}}\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Theta}}}} = N_0\ddot{\underline{\mathbf{H}}}^{\dagger H}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\dot{\underline{\Phi}}^H\dot{\underline{\Phi}}\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}\ddot{\underline{\mathbf{H}}}^{\dagger HH}_{\dot{\underline{\Phi}}} = N_0\dot{\underline{\Phi}}^H\dot{\underline{\Phi}} = \mathbf{R}_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Phi}}}\dot{\underline{\mathbf{n}}}_{\dot{\underline{\Phi}}}} = N_0\ddot{\underline{\mathbf{I}}} \tag{5.217}$$

We should keep in mind that $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$ is the band-based sparse L factor resulting from a block-wise QL operator factorization, and its dimensions are $n_T N_L \times n_T L$, where n_T is the number of MIMO inputs, L is the number of transmitted vector symbols and N_L is the number of independent functions $\underline{\mathbf{H}}(t - lT_s, t)$ that carry those symbols to the output. N_L is always smaller or equal than L, so that $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$ has a lower triangular row-echelon form. For (5.212) to be valid the columns of $\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}$ must be linearly independent, i.e.

$$\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}}\dot{\mathbf{c}} = \mathbf{0} \rightarrow \dot{\mathbf{c}} = \mathbf{0} \triangleq \sum_l \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}}[l,k]}\underline{\mathbf{c}}_l = \mathbf{0}\ \forall_k \rightarrow \underline{\mathbf{c}}_l = \mathbf{0} \tag{5.218}$$

Being the channel statistically determined, so are the functions $\underline{\mathbf{H}}(t - lT_s, t)$ and so is the operator $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$. Moreover, from the statistical independence of a set of N_L domain functions $\underline{\mathbf{H}}(t - lT_s, t)$ one assesses the statistical independence of the matrix coefficients within each column of $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$, and thence the condition in (5.218) can be deemed satisfied with high probability. The whitening approach to the TS-MMF orthonormalization is depicted in the system diagram of Figure 5.15.

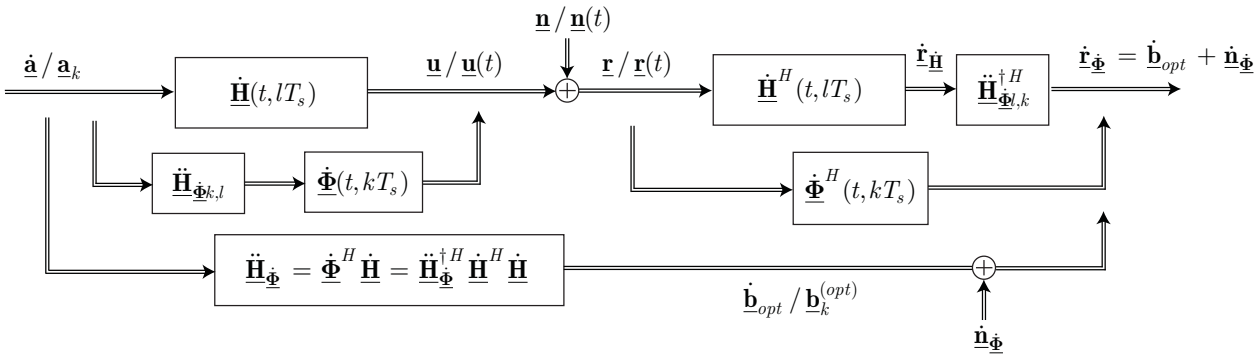

FIGURE 5.15 Orthonormalizing the TS-MMF using the Moore-Penrose pseudoinverse of the discrete channel filter.

5.5.2 FULLY DISCRETIZING THE TS-MMF AND ORTHO-TS-MMF

We have seen that the MIMO channel acts on the input symbols by means of a semi-discrete matrix operator, producing at its output a continuous vector signal. The operator works by sampling the channel TF $\underline{\mathbf{H}}(\tau, t)$ at equispaced parallel transmission domains, as seen in Figure 5.4, scaling the matrix functions thus obtained by the input vector samples and then superposing each of them in an overall weighted sum $\underline{\mathbf{u}}(t)$. The sum is continuous at any time instant and, because of the bandlimited character of the operator $\underline{\dot{\mathbf{H}}}$, it is also bandlimited, hence discretizable by sampling. The possibility of discretizing the received vector signal $\underline{\mathbf{r}}(t)$ prior to matched filtering would be rather convenient, in the sense that not continuous but discrete filters TS-MMF or ORTHO-TS-MMF would be required for optimal detection.

It was shown in Section 5.2 that the symbols $\underline{\mathbf{a}}_k$ are previously sent through a time-invariant pulse shaping filter $\underline{\mathbf{G}}(\tau)$ that outputs the signal $\underline{\mathbf{s}}(t)$ for transmission. The input-output relation has the following two equivalent forms

$$\underline{\mathbf{s}}(t) = \int_{\tau} \underline{\mathbf{G}}(\tau) \underline{\mathbf{x}}_s(t-\tau) d\tau = \sum_k \underline{\mathbf{G}}(t - kT_s) \underline{\mathbf{a}}_k \overset{\mathcal{F}}{\rightarrow} \underline{\mathbf{s}}_{\mathcal{F}}(f) = \underline{\mathbf{G}}_{\mathcal{F}}(f) \underline{\mathbf{a}}_{\mathcal{F}}\left(e^{j2\pi f T_s}\right) \tag{5.219}$$

On the other hand, the channel noiseless output is related to $\underline{\mathbf{s}}_{\mathcal{F}}(f)$ by

$$\begin{aligned} \underline{\mathbf{u}}_{\mathcal{F}}(f) &= \frac{1}{2} \int_{\nu} \underline{\mathbf{C}}_{\mathcal{F}}(f-\nu, \nu) \underline{\mathbf{s}}_{\mathcal{F}}(f-\nu) d\nu = \int_{\nu} \frac{1}{2} \underline{\mathbf{C}}_{\mathcal{F}}(f-\nu, \nu) \underline{\mathbf{G}}_{\mathcal{F}}(f-\nu) \underline{\mathbf{a}}_{\mathcal{F}}\left(e^{j2\pi(f-\nu)T_s}\right) d\nu \\ &= \int_{\nu} \underline{\mathbf{H}}_{\mathcal{F}}(f-\nu, \nu) \underline{\mathbf{a}}_{\mathcal{F}}\left(e^{j2\pi(f-\nu)T_s}\right) d\nu \end{aligned} \tag{5.220}$$

where the global frequency-domain TF is

$$\underline{\mathbf{H}}_{\mathcal{F}}(f, \nu) = \frac{1}{2} \underline{\mathbf{C}}_{\mathcal{F}}(f, \nu) \underline{\mathbf{G}}_{\mathcal{F}}(f) \tag{5.221}$$

We shall assume that the shaping filter is the squared-root raised cosine (SRRC) of less than unity roll-off factor $roll$, Nyquist bandwidth $B = f_s$, and total bandwidth $B(1 + roll)$, i.e.

$$\underline{\mathbf{G}}_{\mathcal{F}}(f) = \sqrt{RC_{f_s, roll}(f)} \mathbf{I} \tag{5.222}$$

where

$$RC_{B,roll}(f)=\begin{cases}1/B & |f|<\dfrac{B}{2}(1-roll)\\ \dfrac{1}{2B}\Big[1+\cos\Big(\pi\Big(|f|-\dfrac{B}{2}(1-roll)\Big)/Broll\Big)\Big] & \dfrac{B}{2}(1-roll)\le|f|<\dfrac{B}{2}(1+roll)\\ 0 & |f|\ge\dfrac{B}{2}(1+roll)\end{cases} \tag{5.223}$$

and the following condition is met

$$f_s\sum_k \underset{\mathcal{F}}{\mathbf{G}}^H(f-kf_s)\underset{\mathcal{F}}{\mathbf{G}}(f-kf_s)=f_s\sum_k RC_{f_s,roll}(f-kf_s)\mathbf{I}=\mathbf{I} \tag{5.224}$$

Since $\underset{\mathcal{F}}{\mathbf{G}}(f)$ is bandlimited and $\underset{\mathcal{F}}{\mathbf{C}}(f,\nu)$ has limited Doppler bandwidth, one concludes that $\underset{\mathcal{F}}{\mathbf{H}}(f,\nu)$ is bandlimited on both arguments, that is

$$\underline{\mathbf{H}}(f,\nu)=0 \qquad for \qquad \begin{cases}|f|>f_s(1+roll)/2<f_s\\ |\nu|>\nu_{\max}\end{cases} \tag{5.225}$$

and, in light of Figure 5.12, its output's bandwidth is limited to

$$B_{\underline{\mathbf{u}}}\le B_{\underline{\mathbf{H}}}+2\nu_{\max}=f_s(1+roll)+2\nu_{\max}<2f_s \tag{5.226}$$

Consequently

$$\begin{aligned}\underset{\mathcal{F}}{\underline{\mathbf{u}}}(f)&=\underset{\mathcal{F}}{\underline{\mathbf{H}}}(f-\nu,\nu)\underset{\nu}{\otimes}\underset{\mathcal{F}}{\underline{\mathbf{a}}}\big(e^{j2\pi fT_s}\big)\\ &=0 \qquad for \qquad |f|>f_s(1+roll)/2+\nu_{\max}<f_s\end{aligned} \tag{5.227}$$

and it is assumed that the signalling rate is high enough to accommodate both the roll-off and the Doppler spread. We will be sampling $\underline{\mathbf{u}}(t)$ before applying the matched filter, and to avoid aliasing we must sample at a rate higher than (5.226). Restricting the roll-off factor to $roll<1-2\nu_{\max}/f_s$, then an adequate sampling rate is $2f_s$, and so we write

$$\underset{\mathcal{F}}{\underline{\mathbf{u}}}\big(e^{j2\pi fT_s/2}\big)=\underset{\mathcal{F}}{\underline{\mathbf{u}}}_{2s}(f)=2f_s\sum_k\underset{\mathcal{F}}{\underline{\mathbf{u}}}(f-k2f_s) \tag{5.228}$$

Now we apply the TS-MMF as given in (5.195), and, knowing that it does not introduce extra bandwidth, we are free to sample its output also at twice the signalling rate, obtaining

$$\begin{aligned}\underset{\mathcal{F}}{\underline{\mathbf{u}}}^{(MF)}\big(e^{j2\pi fT_s/2}\big)&=\underset{\mathcal{F}}{\underline{\mathbf{u}}}^{(MF)}_{2s}(f)=2f_s\sum_k\underset{\mathcal{F}}{\underline{\mathbf{u}}}^{(MF)}(f-k2f_s)\\ &=2f_s\int_\nu\sum_k\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H\big(f-k2f_s,e^{-j2\pi\nu T_s/2}\big)\underset{\mathcal{F}}{\underline{\mathbf{u}}}(f-k2f_s-\nu)d\nu\end{aligned} \tag{5.229}$$

To simplify (5.229) we must look closely to the following bandwidth limits

$$\begin{aligned}\underset{\mathcal{F}}{\underline{\mathbf{H}}}^H\big(f,e^{-j2\pi\nu T_s/2}\big)\neq 0 \quad&\rightarrow\quad \begin{cases}|f|<f_s(1+roll)/2<f_s\\ |\nu-k2f_s|<\nu_{\max}\quad\forall k\end{cases}\\ \underset{\mathcal{F}}{\underline{\mathbf{u}}}(f-\nu)\neq 0 \quad&\rightarrow\quad |f-\nu|<f_s(1+roll)/2+\nu_{\max}<f_s\end{aligned} \tag{5.230}$$

from which it is possible to conclude that there is no sampling-induced overlapping on the inner convolution and, consequently, that the sum of the products in (5.229) can be ex-

changed by a product of sums, and the integration be confined to one period of length $2f_s$. The outcome is

$$\begin{aligned}\underset{\mathcal{F}}{\mathbf{u}}^{(MF)}\left(e^{j2\pi fT_s/2}\right) &= \int_{\nu/2f_s}\sum_k \underset{\mathcal{F}}{\mathbf{H}}^H\left(f-k2f_s, e^{-j2\pi\nu T_s/2}\right)2f_s\sum_l \underset{\mathcal{F}}{\mathbf{u}}(f-l2f_s-\nu)d\nu \\ &= \int_{\nu}\sum_k \underset{\mathcal{F}}{\mathbf{H}}^H\left(f-k2f_s, e^{-j2\pi\nu T_s/2}\right)\underset{\mathcal{F}}{\mathbf{u}}\left(e^{j2\pi(f-\nu)T_s/2}\right)d\nu \\ &= (2f_s)^{-1}\int_{\nu/2f_s}\underset{\mathcal{F}}{\mathbf{H}}^H\left(e^{j2\pi fT_s/2}, e^{-j2\pi\nu T_s/2}\right)\underset{\mathcal{F}}{\mathbf{u}}\left(e^{j2\pi(f-\nu)T_s/2}\right)d\nu\end{aligned} \tag{5.231}$$

which is not unexpected because it sums up to the periodic convolution resulting from filtering with a discrete time-varying filter, namely

$$\underset{\mathcal{F}}{\mathbf{u}}^{(MF)}\left(e^{j2\pi fT_s/2}\right) = \underset{\mathcal{F}}{\mathbf{H}}^H\left(e^{j2\pi fT_s/2}, e^{-j2\pi\nu T_s/2}\right)\underset{\nu/2f_s}{\otimes}\underset{\mathcal{F}}{\mathbf{u}}\left(e^{j2\pi fT_s/2}\right) \tag{5.232}$$

The fully discrete TS-MMF is thus

$$\underline{\underset{\mathcal{F}}{\mathbf{MMF}}}\left(e^{j2\pi(f-\nu)T_s/2}, e^{j2\pi\nu T_s/2}\right) = \underset{\mathcal{F}}{\mathbf{H}}^H\left(e^{j2\pi fT_s/2}, e^{-j2\pi\nu T_s/2}\right) \tag{5.233}$$

and its time-domain counterpart is

$$\underline{\mathbf{MMF}}(lT_s/2, kT_s/2) = \underline{\mathbf{H}}^H(-lT_s/2, (k-l)T_s/2) \tag{5.234}$$

which is easily obtained from (5.167) as follows

$$\begin{aligned}&\underset{\mathcal{F}}{\mathbf{u}}^{(MF)}\left(e^{j2\pi fT_s/2}\right) = (2f_s)^{-1}\int_{f'/2f_s}\underset{\mathcal{F}}{\ddot{\mathbf{H}}}^H_{(2s)}(f', -f)\underset{\mathcal{F}}{\mathbf{u}}\left(e^{j2\pi f'T_s/2}\right)df' \\ &\rightarrow \dot{\underline{\mathbf{u}}}^{(MF)}_{(2s)} = \underline{\ddot{\mathbf{H}}}^H_{(2s)}\dot{\underline{\mathbf{u}}}_{(2s)} \triangleq \underline{\mathbf{u}}^{(MF)}(kT_s/2) = \sum_l \underline{\ddot{\mathbf{H}}}^H_{(2s)}(lT_s/2, kT_s/2)\underline{\mathbf{u}}(lT_s/2) \\ &\qquad = \underline{\mathbf{H}}^H(-lT_s/2, (k-l)T_s/2)\underset{l}{\otimes}\underline{\mathbf{u}}(kT_s/2)\end{aligned} \tag{5.235}$$

It becomes clear that discretizing the TS-MMF has the drawback of demanding the estimation of twice the original transmission domains. Continuity is traded by redundant transmission data. As shown in Figure 5.16, if the receiver is willing to sample the received

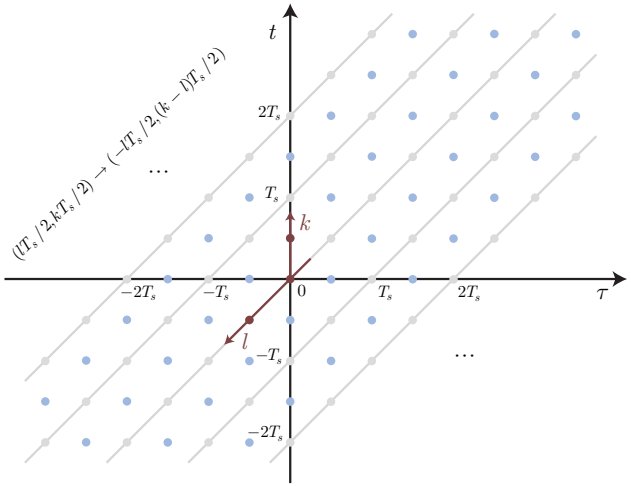

FIGURE 5.16 Effect of sampling the received signal at twice the signalling rate.

vector signal, it must at the same time gather information in the additional time domains

$(t - kT_s/2, t)$, something that may not be acceptable and/or feasible depending on the system design. If the receiver chooses the sampling path, it must then perform sample decimation at the TS-MMF output

$$\begin{aligned}\underset{\mathcal{F}}{\mathbf{u}}_s^{(MF)}(f) &= f_s \sum_k \underset{\mathcal{F}}{\mathbf{u}}^{(MF)}(f - kf_s) = M^{-1} \sum_{i=0}^{M-1} Mf_s \sum_k \underset{\mathcal{F}}{\mathbf{u}}^{(MF)}(f - if_s - kf_s) \\ &= M^{-1} \sum_{i=0}^{M-1} \underset{\mathcal{F}}{\mathbf{u}}_{2s}^{(MF)}(f - if_s) \rightarrow \underset{\mathcal{F}}{\mathbf{u}}_s^{(MF)}(f) = 2^{-1}\left[\underset{\mathcal{F}}{\mathbf{u}}_{2s}^{(MF)}(f) + \underset{\mathcal{F}}{\mathbf{u}}_{2s}^{(MF)}(f - f_s)\right]\end{aligned} \tag{5.236}$$

Aliasing will necessarily occur as part of decimation, but in this case it is a necessary ingredient of optimal detection.

Another trait of the sampling approach is related to the noise. As explained in Section 5.2, the noise introduced by the system is theoretically modelled as space-time white, and for the purpose of detection with the TS-MMF it is a valid model because the TS-MMF effectively blocks the noise for $|f| > f_s(1 + roll)/2$. However, losslessly discretizing the matched filter will require that the noise be eliminated for $|f| > f_s$ with (preferably) an ideal filter bank, otherwise aliasing and distortion will occur. Even though not explicit, this pre-filtering stage is already present in the down-conversion process of the receiver front-end (see Figure 5.1), but since the baseband noise was deemed white, this "implicit" filtering should be made explicit in the respective system diagram. Be that as it may, spectral noise filtering will always be present (e.g. as part of multiuser separation), so our only impair concern should be the ideal desideratum that such a filter never possesses in practice.

Even though the full discretization was derived for the TS-MMF, it directly extends to the ORTHO-TS-MMF

$$\begin{aligned}&\underset{\mathcal{F}}{\mathbf{u}}^{(O-MF)}\left(e^{j2\pi fT_s/2}\right) = \underset{\mathcal{F}}{\mathbf{\Phi}}^H\left(e^{j2\pi fT_s/2}, e^{-j2\pi\nu T_s/2}\right) \underset{\nu/2f_s}{\otimes} \underset{\mathcal{F}}{\mathbf{u}}\left(e^{j2\pi fT_s/2}\right) \\ &\rightarrow \underline{\mathbf{u}}^{(O-MF)}(kT_s/2) = \sum_l \underline{\ddot{\mathbf{\Phi}}}_{(2s)}^H(lT_s/2, kT_s/2)\underline{\mathbf{u}}(lT_s/2) \\ &\qquad \triangleq \underline{\dot{\mathbf{u}}}_{(2s)}^{(O-MF)} = \underline{\ddot{\mathbf{\Phi}}}_{(2s)}^H \underline{\dot{\mathbf{u}}}_{(2s)}\end{aligned} \tag{5.237}$$

and implies that the orthonormalization described in Section 5.3.1 can be discretized, yet it has to be done for twice the number of functions. The alternative is to discretize the TS-MMF instead and then commit the discrete output to the whitening approach devised in Section 5.5.1. The receiver diagram after full filter discretization is depicted in Figure 5.17.

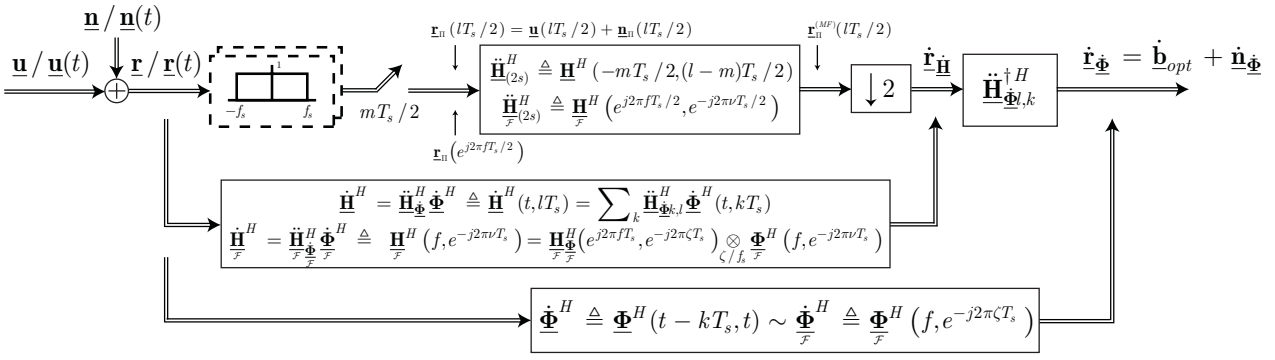

FIGURE 5.17 Optimal detection based, fully discretized receiver model for the MIMO LTV channel.

5.6 *Linear Detection for the Overspread MIMO Channel*

Optimal detection of the sequence of vector samples $\underline{\dot{\mathbf{a}}}$ is accomplished by employing the Viterbi algorithm extended to the MIMO set-up, as it yields a maximum likelihood estimate of $\underline{\dot{\mathbf{a}}}$. As usual, the penalty incurred by optimal detection is computational complexity, and in the MIMO case it becomes overwhelming. If M is the symbol constellation size and $D = \lceil \tau_{\max} / T_s \rceil$ is the channel memory span, then at each stage of detection (every signalling interval) the number of distance metrics to be computed is $M^{(D+1)n_T}$, of which M^{Dn_T} survive. Thus there is an exponential growth in the detection complexity that may become unbearable for certain channel conditions and MIMO configurations. To flank this situation we turn to *time-varying linear filtering* as an alternative method of detection, mainly because linear complexity scaling is guaranteed. The idea behind the following developments is an extension of standard methods of linear estimation (viz. [105], [106], [107], [108]) to time-varying linear operators.

5.6.1 CONSTRAINED LINEAR DETECTION

The first detector that we derive is based on the assumption that the MIMO response can be cropped (or constrained) to an extent of $D = \lceil \tau_{\max} / T_s \rceil$ vector samples. This is equivalent to say that, after applying the ORTHO-TS-MMF, $\ddot{\underline{\mathbf{H}}}_{\underline{\dot{\Phi}}}$ is structured approximately as given by (5.83) and, accordingly, the matched filtered response is expressed as

$$\begin{aligned} \underline{\dot{\mathbf{r}}}_{\underline{\dot{\Phi}}} &= \ddot{\underline{\mathbf{H}}}_{\underline{\dot{\Phi}}} \underline{\dot{\mathbf{a}}} + \underline{\dot{\mathbf{n}}}_{\underline{\dot{\Phi}}} \\ &\triangleq \underline{\mathbf{r}}_{\underline{\Phi}k} = \sum_{l=k-D}^{k} \ddot{\underline{\mathbf{H}}}_{\underline{\dot{\Phi}}k,l} \underline{\mathbf{a}}_l + \underline{\mathbf{n}}_{\underline{\Phi}k} = \sum_{l=0}^{D} \underline{\mathbf{H}}_{\underline{\Phi}l,k} \underline{\mathbf{a}}_{k-l} + \underline{\mathbf{n}}_{\underline{\Phi}k} \end{aligned} \tag{5.238}$$

From this expression one readily finds that all the information about $\underline{\mathbf{a}}_k$ is conveyed by the vector samples $\underline{\mathbf{r}}_{\underline{\Phi}k} \ldots \underline{\mathbf{r}}_{\underline{\Phi}k+D}$, which means that a linear detector may be based solely on these to determine an estimate of $\underline{\mathbf{a}}_k$, $\hat{\underline{\mathbf{a}}}_k$. Denoting the linear detector operator by $\ddot{\underline{\mathbf{D}}}$, the estimate that it provides at time instant k is obtained as follows

$$\hat{\underline{\mathbf{a}}}_k = \sum_{n=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{k,n} \underline{\mathbf{r}}_{\underline{\Phi}n} = \sum_{l=n-D}^{n} \left[\sum_{n=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{k,n} \ddot{\underline{\mathbf{H}}}_{\underline{\dot{\Phi}}n,l} \right] \underline{\mathbf{a}}_l + \sum_{n=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{k,n} \underline{\mathbf{n}}_{\underline{\Phi}n} \tag{5.239}$$

from which we understand that $\ddot{\underline{\mathbf{D}}}$ is necessarily anticausal, even though this is not an issue because a sufficient lag may be introduced.

The error resulting from (5.239) is the difference between $\underline{\mathbf{a}}_k$ and its estimate $\hat{\underline{\mathbf{a}}}_k$, i.e. $\underline{\mathbf{e}}_k = \underline{\mathbf{a}}_k - \hat{\underline{\mathbf{a}}}_k$, in terms of which $\underline{\mathbf{a}}_k$ may be written as

$$\underline{\mathbf{a}}_k = \hat{\underline{\mathbf{a}}}_k + \underline{\mathbf{e}}_k = \sum_{n=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{k,n} \underline{\mathbf{r}}_{\underline{\Phi}n} + \underline{\mathbf{e}}_k = \sum_{n=-D}^{0} \underline{\mathbf{D}}_{n,k} \underline{\mathbf{r}}_{\underline{\Phi}k-n} + \underline{\mathbf{e}}_k \tag{5.240}$$

At this point we must seek for the operator $\ddot{\underline{\mathbf{D}}}$ that provides the optimal linear estimate of $\underline{\mathbf{a}}_k$, but, since $\underline{\dot{\mathbf{r}}}_{\underline{\dot{\Phi}}}$ is statistically determined, $\ddot{\underline{\mathbf{D}}}$ must be obtained in the mean sense. To this end we shall use the *minimum mean squared error (MMSE)* criterion, which will guarantee that the average energy in the error of linear estimation is minimized. Mathematically, we de-

scribe the MMSE criterion as

$$\ddot{\underline{\mathbf{D}}}_{opt|k,k:k+D} = \underset{\ddot{\underline{\mathbf{D}}}_{k,k:k+D}}{\arg\min}\, E\left[\underline{\mathbf{e}}_k^H \underline{\mathbf{e}}_k\right] = \underset{\ddot{\underline{\mathbf{D}}}_{k,k:k+D}}{\arg\min}\, tr\left\{\ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{e}}_k \underline{\mathbf{e}}_k}\right\} \tag{5.241}$$

where the following notation is used for simplicity

$$\{k \le n \le k+D\} \rightarrow k : k+D \tag{5.242}$$

Since $\hat{\underline{\mathbf{a}}}_k \in \mathcal{R}\left(\underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}k:k+D}\right)$ it is evident that the mean error energy is minimal when $\hat{\underline{\mathbf{a}}}_k$ is the orthogonal projection of $\underline{\mathbf{a}}_k$ onto $\mathcal{R}\left(\underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}k:k+D}\right)$, that is, $\underline{\mathbf{e}}_k \in \mathcal{R}^{\perp}\left(\underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}k:k+D}\right)$, and thus

$$E\left[\underline{\mathbf{e}}_k \underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}^H\right] = \mathbf{0} \quad k \le n \le k+D \;\rightarrow \ddot{\underline{\mathbf{D}}}_{opt|k,k:k+D} \tag{5.243}$$

Expanding the error $\underline{\mathbf{e}}_k = \underline{\mathbf{a}}_k - \hat{\underline{\mathbf{a}}}_{opt|k}$ into (5.243) we are led to

$$E\left[\underline{\mathbf{a}}_k \underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}^H\right] = E\left[\hat{\underline{\mathbf{a}}}_{opt|k} \underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}^H\right] = \sum_{p=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{opt|k,p} E\left[\underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}p} \underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}^H\right] \qquad k \le n \le k+D \tag{5.244}$$

The left expectation can be simplified by expanding $\underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}$ with (5.238), yielding

$$E\left[\underline{\mathbf{a}}_k \underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}^H\right] = \sum_{l=n-D}^{n} E\left[\underline{\mathbf{a}}_k \underline{\mathbf{a}}_l^H\right] \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,l}^H = 2 \sum_{l=n-D}^{n} \ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_l} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,l}^H \tag{5.245}$$

which, assuming that the transmitted vector samples $\underline{\mathbf{a}}_k$ are independent and zero-mean, hence uncorrelated and orthogonal, simplifies to

$$\ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_l} = \ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_k} \delta_{k-l} \quad \rightarrow \quad E\left[\underline{\mathbf{a}}_k \underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}^H\right] = 2\ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_k} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,k}^H \qquad k \le n \le k+D \tag{5.246}$$

Equivalently, the right expectation in (5.244) expands as

$$\begin{aligned} E\left[\underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}p} \underline{\mathbf{r}}_{\underline{\boldsymbol{\Phi}}n}^H\right] &= \sum_{l=p-D}^{p} \sum_{m=n-D}^{n} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}p,l} E\left[\underline{\mathbf{a}}_l \underline{\mathbf{a}}_m^H\right] \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,m}^H + E\left[\underline{\mathbf{n}}_{\underline{\boldsymbol{\Phi}}p} \underline{\mathbf{n}}_{\underline{\boldsymbol{\Phi}}n}^H\right] \\ &= 2 \sum_{l=\max\{n,p\}-D}^{\min\{n,p\}} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}p,l} \ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_l \underline{\mathbf{a}}_l} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,l}^H + 2N_0 \delta_{p-n} \mathbf{I} \end{aligned} \tag{5.247}$$

where complete independence between the noise and the transmitted samples was also assumed. Inserting (5.246) and (5.247) into (5.244) we retrieve

$$\begin{aligned} \ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_k} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,k}^H &= \sum_{p=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{opt|k,p} \left[\sum_{l=\max\{n,p\}-D}^{\min\{n,p\}} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}p,l} \ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_l \underline{\mathbf{a}}_l} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,l}^H + N_0 \delta_{p-n} \mathbf{I} \right] \quad k \le n \le k+D \\ &= \sum_{p=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{opt|k,p} \ddot{\underline{\mathbf{T}}}_{p,n} \qquad k \le n \le k+D \end{aligned} \tag{5.248}$$

where

$$\ddot{\underline{\mathbf{T}}}_{p,n} = \sum_{l=\max\{n,p\}-D}^{\min\{n,p\}} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}p,l} \ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_l \underline{\mathbf{a}}_l} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\boldsymbol{\Phi}}}n,l}^H + N_0 \delta_{p-n} \mathbf{I} \tag{5.249}$$

In matrix notation, (5.248) is equivalent to

$$\ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_k} \begin{pmatrix} \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}} k,k} \\ \vdots \\ \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}} k+D,k} \end{pmatrix}^T = \begin{pmatrix} \ddot{\underline{\mathbf{D}}}_{opt|k,k} \\ \vdots \\ \ddot{\underline{\mathbf{D}}}_{opt|k,k+D} \end{pmatrix}^T \begin{pmatrix} \ddot{\underline{\mathbf{T}}}_{k,k} & \cdots & \ddot{\underline{\mathbf{T}}}_{k,k+D} \\ \vdots & \ddots & \vdots \\ \ddot{\underline{\mathbf{T}}}_{k+D,k} & \cdots & \ddot{\underline{\mathbf{T}}}_{k+D,k+D} \end{pmatrix} \tag{5.250}$$

or, more compactly, to

$$\underline{\mathbf{R}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_k} \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}} k,k:k+D} = \ddot{\underline{\mathbf{D}}}_{opt|k,k:D} \ddot{\underline{\mathbf{T}}}_{k:k+D,k:k+D} \tag{5.251}$$

Determining $\ddot{\underline{\mathbf{D}}}_{opt|k,k:D}$ can now be done by standard Gaussian elimination or computing the inverse of $\ddot{\underline{\mathbf{T}}}_{k:k+D,k:k+D}$, which, being singular, still has a generalized inverse (e.g. Drazin inverse or Moore-Penrose pseudoinverse), i.e.

$$\ddot{\underline{\mathbf{D}}}_{opt|k,k:k+D} = \underline{\mathbf{R}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_k} \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}} k,k:k+D} \ddot{\underline{\mathbf{T}}}^{\dagger}_{k:k+D,k:k+D} \tag{5.252}$$

The optimal estimate so determined is

$$\begin{aligned} \hat{\underline{\mathbf{a}}}_{opt|k} &= \sum_{n=k}^{k+D} \ddot{\underline{\mathbf{D}}}_{opt|k,n} \underline{\mathbf{r}}_{\underline{\Phi} n} = \begin{pmatrix} \ddot{\underline{\mathbf{D}}}_{opt|k,k} & \cdots & \ddot{\underline{\mathbf{D}}}_{opt|k,k+D} \end{pmatrix} \begin{pmatrix} \underline{\mathbf{r}}_{\underline{\Phi} k} \\ \vdots \\ \underline{\mathbf{r}}_{\underline{\Phi} k+D} \end{pmatrix} \\ &= \ddot{\underline{\mathbf{D}}}_{opt|k,k:k+D} \underline{\mathbf{r}}_{\underline{\Phi} k:k+D} \end{aligned} \tag{5.253}$$

which is finally applied to a symbol-by-symbol slicer bank that compares each distance to each constellation symbol and chooses the ones with the lowest distance.

The optimal linear detector can be further simplified when the symbols are mutually orthogonal, $\ddot{\underline{\mathbf{R}}}_{\underline{\mathbf{a}}_k \underline{\mathbf{a}}_k} = (P_{\underline{\mathbf{a}}}/2)\mathbf{I}$, where P_a is the average power of the symbol constellation, and the result is

$$\ddot{\underline{\mathbf{D}}}_{opt|k,k:k+D} = \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}} k,k:k+D} \ddot{\underline{\mathbf{X}}}^{\dagger}_{k:k+D,k:k+D} \tag{5.254}$$

where

$$\ddot{\underline{\mathbf{X}}}_{p,n} = \sum_{l=\max\{n,p\}-D}^{\min\{n,p\}} \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}} p,l} \ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\Phi}} n,l} + (1/\gamma_t)\delta_{p-n}\mathbf{I} \tag{5.255}$$

and $\gamma_t = P_{\underline{\mathbf{a}}}/\sigma^2_{\underline{\mathbf{n}}_{\dot{\underline{\Phi}}}}$ is the average branch SNR with reference to the MIMO input (where $\sigma^2_{\underline{\mathbf{n}}_{\dot{\underline{\Phi}}}} = 2N_0$ is the branch noise power at the output of the ORTHO-TS-MMF).

5.6.2 UNCONSTRAINED LINEAR DETECTION

The constrained linear detector was built upon a sample by sample approach, and each estimate was based on a truncated sequence of received vector samples $\underline{\mathbf{r}}_{\underline{\Phi} k} \cdots \underline{\mathbf{r}}_{\underline{\Phi} k+D}$. Yet more expensive, a better detector can be derived by including in the process the entire sequence of vector samples $\underline{\mathbf{r}}_{\underline{\Phi} 0} \cdots \underline{\mathbf{r}}_{\underline{\Phi} L}$ or, equivalently, $\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Phi}}}$, which means that all the information available in $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}}$ is used to build the detector.

The transmission model is unchanged

$$\dot{\underline{\mathbf{r}}}_{\dot{\underline{\Phi}}} = \dot{\underline{\mathbf{\Phi}}}^H \underline{\mathbf{r}} = \dot{\underline{\mathbf{\Phi}}}^H \left(\underline{\dot{\mathbf{H}}} \underline{\dot{\mathbf{a}}} + \underline{\mathbf{n}} \right) = \ddot{\underline{\mathbf{H}}}_{\dot{\underline{\Phi}}} \underline{\dot{\mathbf{a}}} + \underline{\dot{\mathbf{n}}}_{\dot{\underline{\Phi}}} \tag{5.256}$$

and accounts for the channel and orthogonal matched filtering as well. The linear estimator of $\underline{\dot{\mathbf{a}}}$ is

$$\hat{\underline{\dot{\mathbf{a}}}} = \underline{\ddot{\mathbf{D}}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}} = \underline{\ddot{\mathbf{D}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\dot{\mathbf{a}}} + \underline{\ddot{\mathbf{D}}}\underline{\dot{\mathbf{n}}}_{\dot{\Phi}} = \underline{\dot{\mathbf{a}}} + \left(\underline{\ddot{\mathbf{D}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}} - \ddot{\mathbf{I}}\right)\underline{\dot{\mathbf{a}}} + \underline{\ddot{\mathbf{D}}}\underline{\dot{\mathbf{n}}}_{\dot{\Phi}} \tag{5.257}$$

and $\underline{\ddot{\mathbf{D}}}$ is still the linear detector operator. The error induced by the estimator is the sum of two independent entities: the symbol distortion and the residual noise, i.e.

$$\underline{\dot{\mathbf{a}}} = \hat{\underline{\dot{\mathbf{a}}}} + \underline{\dot{\mathbf{e}}} = \underline{\ddot{\mathbf{D}}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}} + \underline{\dot{\mathbf{e}}} \quad \rightarrow \quad \underline{\dot{\mathbf{e}}} = \underline{\dot{\mathbf{a}}} - \hat{\underline{\dot{\mathbf{a}}}} = \underbrace{\underline{\dot{\mathbf{a}}} - \underline{\ddot{\mathbf{D}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\dot{\mathbf{a}}}}_{\underline{\dot{\mathbf{a}}}_{distortion}} - \underbrace{\underline{\ddot{\mathbf{D}}}\underline{\dot{\mathbf{n}}}_{\dot{\Phi}}}_{\underline{\dot{\mathbf{n}}}_{residual}} \tag{5.258}$$

which amounts to the divergence of the estimator from the original sequence of vector samples. The error is once more minimized when it is orthogonal to the matched filter's output, that is

$$E\left[\underline{\dot{\mathbf{e}}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = \mathbf{0} \rightarrow \underline{\ddot{\mathbf{D}}}_{opt} = \arg\min_{\underline{\ddot{\mathbf{D}}}} E\left[\underline{\dot{\mathbf{e}}}^{H}\underline{\dot{\mathbf{e}}}\right] = \arg\min_{\underline{\ddot{\mathbf{D}}}} tr\left\{\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{e}}\dot{\mathbf{e}}}\right\} \tag{5.259}$$

Expanding the error in (5.259) we have

$$E\left[\underline{\dot{\mathbf{e}}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = E\left[\left(\underline{\dot{\mathbf{a}}} - \hat{\underline{\dot{\mathbf{a}}}}_{opt}\right)\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = E\left[\left(\underline{\dot{\mathbf{a}}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}\right)\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = \mathbf{0} \rightarrow E\left[\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = \underline{\ddot{\mathbf{D}}}_{opt}E\left[\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] \tag{5.260}$$

which means that two statistical expectations need to be taken in order to proceed, $E\left[\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right]$ and $\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{r}}_{\dot{\Phi}}\dot{\mathbf{r}}_{\dot{\Phi}}} = E\left[\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right]$. Because the noise is independent from the transmitted symbols, both are trivially obtained. The first one is

$$E\left[\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = E\left[\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}^{H}\right]\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} = 2\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} \tag{5.261}$$

and the second one

$$E\left[\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = \underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}E\left[\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}^{H}\right]\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} + E\left[\underline{\dot{\mathbf{n}}}_{\dot{\Phi}}\underline{\dot{\mathbf{n}}}_{\dot{\Phi}}^{H}\right] = 2\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} + 2\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{n}}_{\dot{\Phi}}\dot{\mathbf{n}}_{\dot{\Phi}}} \tag{5.262}$$

which, since

$$\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{n}}_{\dot{\Phi}}\dot{\mathbf{n}}_{\dot{\Phi}}} = \frac{1}{2}E\left[\underline{\dot{\mathbf{n}}}_{\dot{\Phi}}\underline{\dot{\mathbf{n}}}_{\dot{\Phi}}^{H}\right] = \underline{\dot{\mathbf{\Phi}}}^{H}\frac{1}{2}E\left[\underline{\mathbf{n}}\underline{\mathbf{n}}^{H}\right]\underline{\dot{\mathbf{\Phi}}} = N_0\underline{\ddot{\mathbf{I}}} \tag{5.263}$$

simplifies to

$$\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{r}}_{\dot{\Phi}}\dot{\mathbf{r}}_{\dot{\Phi}}} = E\left[\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}^{H}\right] = \underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} + N_0\underline{\ddot{\mathbf{I}}} \tag{5.264}$$

Equating (5.261) and (5.264) we get

$$\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} = \underline{\ddot{\mathbf{D}}}_{opt}\left(\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} + N_0\underline{\ddot{\mathbf{I}}}\right) \tag{5.265}$$

and after applying the generalized inverse (or Gaussian elimination), yields the linear detector operator sought for

$$\underline{\ddot{\mathbf{D}}}_{opt} = \underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H}\left(\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} + N_0\underline{\ddot{\mathbf{I}}}\right)^{\dagger} \tag{5.266}$$

A further simplification is achieved for independently transmitted zero-mean symbols

$$\underline{\ddot{\mathbf{R}}}_{\dot{\mathbf{a}}\dot{\mathbf{a}}} = \frac{1}{2}E\left[|\underline{a}|^2\right]\underline{\ddot{\mathbf{I}}} = \left(P_{\underline{\dot{\mathbf{a}}}}/2\right)\underline{\ddot{\mathbf{I}}} \quad \rightarrow \underline{\ddot{\mathbf{D}}}_{opt} = \underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H}\left(\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}^{H} + \frac{1}{\gamma_t}\underline{\ddot{\mathbf{I}}}\right)^{\dagger} \tag{5.267}$$

where again $\gamma_t = P_{\underline{\dot{\mathbf{a}}}} / \sigma^2_{\underline{\dot{\mathbf{u}}}_{\dot{\Phi}}}$, and is somewhat resemblant of (5.254) and (5.255).

The average energy in the error produced by the optimal linear detector is minimal, and provides a symbol-by-symbol low complexity alternative to the TV-MLSE detector. The minimum average error energy can be determined as follows. Since the estimator is orthogonal to the error, the latter's autocorrelation is given by

$$\begin{aligned}\ddot{\mathbf{R}}^{(opt)}_{\underline{\dot{\mathbf{e}}}\underline{\dot{\mathbf{e}}}} &= \frac{1}{2}E\left[\underline{\dot{\mathbf{e}}}\underline{\dot{\mathbf{a}}}^H\right] = \frac{1}{2}E\left[\left(\underline{\dot{\mathbf{a}}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}\right)\underline{\dot{\mathbf{a}}}^H\right] = \ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}} - \frac{1}{2}\underline{\ddot{\mathbf{D}}}_{opt}E\left[\underline{\dot{\mathbf{r}}}_{\dot{\Phi}}\underline{\dot{\mathbf{a}}}^H\right] \\ &= \ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}} = \left(\ddot{\mathbf{I}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\right)\ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}}\end{aligned} \tag{5.268}$$

and consequently

$$E\left[\underline{\dot{\mathbf{e}}}^H\underline{\dot{\mathbf{e}}}\right]_{\min} = 2tr\left\{\ddot{\mathbf{R}}^{(opt)}_{\underline{\dot{\mathbf{e}}}\underline{\dot{\mathbf{e}}}}\right\} = 2tr\left\{\left(\ddot{\mathbf{I}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\right)\ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}}\right\} \tag{5.269}$$

However, from (5.265) we deduce that

$$\left(\ddot{\mathbf{I}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\right)\ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}}\underline{\ddot{\mathbf{H}}}^H_{\dot{\Phi}} = N_0\underline{\ddot{\mathbf{D}}}_{opt} \tag{5.270}$$

which can be related to (5.269) by introducing the generalized inverse of $\underline{\ddot{\mathbf{H}}}^H_{\dot{\Phi}}$, so that we get the following solution

$$\left(\ddot{\mathbf{I}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\right)\ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}} = N_0\underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}^{\dagger H}_{\dot{\Phi}} \quad \rightarrow \ddot{\mathbf{R}}^{(opt)}_{\underline{\dot{\mathbf{e}}}\underline{\dot{\mathbf{e}}}} = \left(\ddot{\mathbf{I}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\right)\ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}} = N_0\underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}^{\dagger H}_{\dot{\Phi}} \tag{5.271}$$

and, ultimately

$$E\left[\underline{\dot{\mathbf{e}}}^H\underline{\dot{\mathbf{e}}}\right]_{\min} = 2tr\left\{\left(\ddot{\mathbf{I}} - \underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\right)\ddot{\mathbf{R}}_{\underline{\dot{\mathbf{a}}}\underline{\dot{\mathbf{a}}}}\right\} = 2tr\left\{N_0\underline{\ddot{\mathbf{D}}}_{opt}\underline{\ddot{\mathbf{H}}}^{\dagger H}_{\dot{\Phi}}\right\} \tag{5.272}$$

5.6.3 ZERO-FORCING LINEAR DETECTION

The zero-forcing detector is derived by the same steps as the two previous detectors, except that the presence of channel noise is entirely ignored. The constrained zero-forcing linear detector is found to be given by

$$\underline{\ddot{\mathbf{D}}}_{opt|k,k:k+D} = \underline{\ddot{\mathbf{H}}}^H_{\dot{\Phi}k,k:k+D}\left(\sum_{l=\max\{n,p\}-D}^{\min\{n,p\}}\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}p,l}\underline{\ddot{\mathbf{H}}}^H_{\dot{\Phi}n,l}\right)^{\dagger} \tag{5.273}$$

and, by the same reasoning, the unconstrained zero-forcing linear detector by

$$\underline{\ddot{\mathbf{D}}}_{opt} = \underline{\ddot{\mathbf{H}}}^H_{\dot{\Phi}}\left(\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}\underline{\ddot{\mathbf{H}}}^H_{\dot{\Phi}}\right)^{\dagger} \tag{5.274}$$

which, whenever $\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}$ is invertible, reduces to $\underline{\ddot{\mathbf{D}}}_{opt} = \underline{\ddot{\mathbf{H}}}^{-1}_{\dot{\Phi}}$. The zero-forcing detector performance is inferior to the optimal linear detector, in particular because if $\underline{\ddot{\mathbf{H}}}^{-1}_{\dot{\Phi}}$ contains large entries the noise will be greatly enhanced while the input symbols will not. Nevertheless, if the channel is indeed noiseless, the zero-forcing detector utterly removes the channel distortion provided that $\underline{\ddot{\mathbf{H}}}_{\dot{\Phi}}$ is not badly conditioned and the channel state information is sufficiently accurate.

5.7 Error Performance of the ORTHO-TS-MMF with Linear Detection

Having derived a detection oriented, TV-MLSE based, optimal discrete model for the overspread, ORTHO-TS-MMF filtered MIMO system, and the corresponding optimal linear TV-MMSE detector, the aim of this section is to study its performance in terms of the average error rate of the transmitted sequence of symbols. To accomplish this, we shall divert from the pure theoretical standpoint and engage in Monte-Carlo based numerical simulation of the MIMO system.

5.7.1 Channel Model for Simulation

We begin with the continuous passband channel model from the j-th input to the i-th output, as given by (5.3) and repeated here for convenience

$$c_{ij}(\tau,t)=\sum_{n=1}^{N_{ij}(t)}\alpha_n^{(ij)}(t)\delta(\tau-\tau_n^{(ij)}(t)) \tag{5.275}$$

Its baseband version is

$$\underline{c}_{ij}(\tau,t)=2c_{ij}(\tau,t)e^{-j2\pi f_c\tau}=2\sum_{n=1}^{N_{ij}(t)}\alpha_n^{(ij)}(t)e^{j\varphi_n^{(ij)}(t)}\delta(\tau-\tau_n^{(ij)}(t)) \tag{5.276}$$

where the multipath delays and associated phases are

$$\begin{aligned}\tau_n^{(ij)}(t)&=\tau_n^{(ij)}-(\omega_n^{(ij)}/\omega_c)t\\ \varphi_n^{(ij)}(t)&=-2\pi f_c\tau_n^{(ij)}(t)=-2\pi f_c(\tau_n^{(ij)}-(\omega_n^{(ij)}/\omega_c)t)=\varphi_n^{(ij)}+\omega_n^{(ij)}t\end{aligned} \tag{5.277}$$

The carrier frequency is ω_c and the angular Doppler shift is given by

$$\omega_n^{(ij)}=2\pi\nu_n^{(ij)}=2\pi\frac{\upsilon}{\lambda}\cos\phi_n^{(ij)}=2\pi f_D\cos\phi_n^{(ij)} \tag{5.278}$$

where υ is the relative velocity between transmitter and receiver, λ is the wavelength, $\phi_n^{(ij)}$ are the different angles of arrival relative to the direction of motion, and f_D is the maximum Doppler shift.

To simplify the model some assumptions need to be made. The sole purpose is to streamline the simulation by invoking some typical channel characteristics, meaning they do not induce any loss of generality whatsoever. The imposed assumptions are:

1. the carrier frequency is much larger than the Doppler shift, so that the multipath delays may be considered time-invariant, i.e. $\tau_n^{(ij)}(t)=\tau_n^{(ij)}$;
2. the relative motion is small in scale, which guarantees that the statistical properties of the channel remain unchanged during the simulation; also, the number of multipath replicas and respective attenuations do not vary with time;
3. the distance between transmitter and receiver is sufficiently large so that each multipath component arrives at different receiver antennas by parallel paths; this means that $\phi_n^{(ij)}=\phi_n^{(1j)}$; moreover, the multipath attenuations are equal, i.e. $\alpha_n^{(ij)}=\alpha_n^{(1j)}$;
4. each transmitter produces the same number of multipath components, and each one impinges on each receiver antenna, i.e. $N_{ij}=N$; additionally, it is assumed that the

multipath components coming from different transmitters are completely independent;

5. the receiver moves in parallel to the axis of the receiver array, as shown in Figure 5.18;

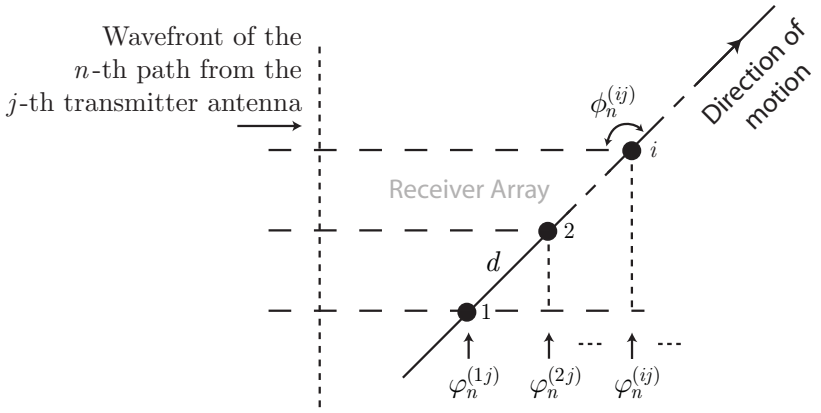

FIGURE 5.18 Illustrating the different multipath phases of a moving array with colinear axis.

this means that the different phases are related by

$$\begin{aligned}\varphi_n^{(ij)} &= \varphi_n^{(i-1,j)} + \beta d\cos(\phi_n^{(1j)}) = \varphi_n^{(1j)} + \beta d(i-1)\cos(\phi_n^{(1j)}) \\ &= -2\pi f_c\left(\tau_n^{(1j)} - (d/c)(i-1)\cos(\phi_n^{(1j)})\right)\end{aligned} \tag{5.279}$$

where $\beta = 2\pi/\lambda$, d is the antenna separation, and c is the speed of light. There will be two situation for the antenna separation, one in which all the antennas will be equispacedly fit within $d_{\max} = 0.5\lambda$, and other in which the array will scale freely as more antennas are added, i.e. $d_{\max} = dn_R = 0.5\lambda n_R$.

The chosen statistical model for the multipath delays is the exponential model, which is completely defined by the RMS delay spread σ_τ as

$$\tau_n^{(1j)} = -\sigma_\tau \ln(1-u_n^{(1j)}) \quad \rightarrow f_\tau(\tau_n^{(1j)}) = \frac{1}{\sigma_\tau} e^{-\tau_n^{(1j)}/\sigma_\tau} \tag{5.280}$$

where $u_n^{(1j)}$ is uniformly distributed in the interval $(0,1)$. The delays to the other antennas are computed from

$$\tau_n^{(ij)} = \tau_n^{(1j)} - (d/c)(i-1)\cos(\phi_n^{(1j)}) \tag{5.281}$$

Equivalently, the attenuations are exponentially decaying $\alpha_n^{(1j)} = e^{-\tau_n^{(1j)}/2\sigma_\tau}$, simulating the characteristic of the multipath power delay profile. The angles of arrival $\phi_n^{(1j)}$ will be considered uniformly distributed in the interval $(0,2\pi)$, simulating the isotropic scattering model. The RMS delay spread will be the rule-of-thumb of a typical mobile propagation environment, i.e. $\sigma_\tau = 3\,\mu s$, and unless otherwise noted, the number of multipath components will be $N = 8$. Since it is a natural candidate for 4G LTE licensing, the frequency of operation $f_c = 2.6\,GHz$ will be chosen for the simulation. At this carrier frequency and mobile speeds of about $v = 85\,km/h$, the maximum Doppler shift is $f_D \approx 200\,Hz$, which will also be the simulation default.

The pulse shaping filter is the 1/3 roll-off SRRC with impulse response

$$srrc(\tau) = \frac{\cos(\pi(1+roll)f_s\tau) + (4rollf_s\tau)^{-1}\sin(\pi(1-roll)f_s\tau)}{\left(4roll\sqrt{f_s}\right)^{-1}\pi\left(1-(4rollf_s\tau)^2\right)} \tag{5.282}$$

delayed an integer number of symbols $srrc_{delay}$ and truncated on both sides

$$g(\tau) = \begin{cases} srrc(\tau - srrc_{delay}/f_s) & 0 \le \tau \le 2srrc_{delay}/f_s \\ 0 & otherwise \end{cases} \tag{5.283}$$

which, cascaded with the wireless channel yields the global input-output response

$$\underline{h}_{ij}(\tau,t) = \frac{1}{2}\int_{\alpha} \underline{c}_{ij}(\alpha,t)g(\tau-\alpha)d\alpha = \sum_{n=1}^{N} \alpha_n^{(1j)} e^{j\left(\varphi_n^{(ij)} + \omega_n^{(1j)}t\right)} g(\tau - \tau_n^{(ij)}) \tag{5.284}$$

5.7.2 Discrete Channel Model for Simulation

We know that the noiseless continuous output of the MIMO channel is given by

$$\underline{\mathbf{u}} = \underline{\dot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} \triangleq \underline{\mathbf{u}}(t) = \sum_{l} \underline{\dot{\mathbf{H}}}(t, lT_s)\underline{\mathbf{a}}_l \tag{5.285}$$

and it was shown in section Section 5.5.2 that a sampling rate of $2f_s$ was required to avoid aliasing with a comfortable margin. Consequently, we will be sampling (5.285) at a time rate that is twice the signalling rate, i.e.

$$\underline{\mathbf{u}}(kT_s/2) = \sum_{l} \underline{\ddot{\mathbf{H}}}(kT_s/2, lT_s)\underline{\mathbf{a}}_l \triangleq {}_2\underline{\dot{\mathbf{u}}} = {}_2\underline{\ddot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} \tag{5.286}$$

where we have to pay attention to the fact that the operator has two different discrete rates on each of its sides. If the noise is confined to a bandwidth $B = f_s$, then the global noisy discrete channel model is

$$\begin{aligned} \underline{\mathbf{r}}(kT_s/2) &= \sum_{l} \underline{\ddot{\mathbf{H}}}(kT_s/2, lT_s)\underline{\mathbf{a}}_l + \underline{\mathbf{n}}(kT_s/2) \\ &\triangleq {}_2\underline{\dot{\mathbf{r}}} = {}_2\underline{\ddot{\mathbf{H}}}\underline{\dot{\mathbf{a}}} + {}_2\underline{\dot{\mathbf{n}}} \end{aligned} \tag{5.287}$$

The entries of $\underline{\ddot{\mathbf{H}}}(kT_s/2, lT_s)$, $\underline{\ddot{h}}_{ij}(kT_s/2, lT_s)$, are one-dimensional operators that define the response between the j-th input to the i-th output of the MIMO channel. They can be related to (5.284) as follows

$$\begin{aligned} \underline{\ddot{h}}_{ij}(kT_s/2, lT_s) &= \underline{h}_{ij}(kT_s/2 - lT_s, kT_s/2) \\ &= \sum_{n=1}^{N} \alpha_n^{(1j)} e^{j\left(\varphi_n^{(ij)} + \omega_n^{(1j)} kT_s/2\right)} g\left(kT_s/2 - lT_s - \tau_n^{(ij)}\right) \end{aligned} \tag{5.288}$$

and provide all the information required to build the fully discrete channel operator for simulation.

5.7.3 Building the ORTHO-TS-MMF with Block-wise Householder Reflections

The ORTHO-TS-MMF may be computed using a discretized version of the orthogonalization procedure described in section Section 5.3.1. Unfortunately, experience has shown

that it is rather slow to complete and even worst, has poor numerical stability. This renders the orthonormalization very time consuming if one is to address stability with variable precision arithmetic. An alternative version of the orthogonalization based on the modified Gram-Schmidt orthogonalization was implemented, but despite being more stable it was still slower than desired.

To address both issues we now explain a novel method based on backward block-wise Householder reflections that proved to be incredibly fast and stable. Our goal is to perform the following decomposition

$$ {}_2\ddot{\underline{\mathbf{H}}} = {}_2\ddot{\underline{\mathbf{\Phi}}}\ddot{\underline{\mathbf{H}}}_{\ddot{\mathbf{\Phi}}} \triangleq \ddot{\underline{\mathbf{H}}}(kT_s/2, lT_s) = \sum_m \ddot{\underline{\mathbf{\Phi}}}(kT_s/2, mT_s)\ddot{\underline{\mathbf{H}}}_{\ddot{\mathbf{\Phi}}m,l} \tag{5.289} $$

where $\ddot{\underline{\mathbf{H}}}_{\ddot{\mathbf{\Phi}}}$ is block lower triangular with block size $n_T \times n_T$ and ${}_2\ddot{\underline{\mathbf{\Phi}}}$ is block orthogonal, each block consisting of n_T columns. We remember the reader that the indexes in (5.289) are indexing the blocks and not the entries themselves. The idea is to sequentially eliminate the entries above each block in the lower right diagonal of ${}_2\ddot{\underline{\mathbf{H}}}$, where the lower right diagonal is interpreted as the diagonal that starts with the utmost lower right block of ${}_2\ddot{\underline{\mathbf{H}}}$. To eliminate the entries above the utmost lower right block we begin by finding the QR factorization of the utmost right n_T columns of ${}_2\ddot{\underline{\mathbf{H}}}$, i.e.

$$ \ddot{\underline{\mathbf{H}}}(:, LT_s) = \mathbf{QR} \tag{5.290} $$

where $\mathbf{R}$ is $n_T \times n_T$ lower triangular. Next we retrieve the utmost lower right block of $\mathbf{Q}$ and compute

$$ \mathbf{B} = \mathbf{Q}(last, :) \rightarrow \mathbf{P} = \mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1/2} \tag{5.291} $$

which lets as define

$$ \mathbf{v} = \mathbf{Q} - \mathbf{eP}\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1/2} = \mathbf{Q} - \mathbf{eP} \tag{5.292} $$

where $\mathbf{e}$ is the same size of $\mathbf{Q}$ and is all zeros except on the lower right diagonal, where it is all ones. The block $\mathbf{v}$ can be used to build the first Householder reflector

$$ \mathbf{H}_r = \mathbf{I} - 2\mathbf{v}(\mathbf{v}^H\mathbf{v})^{-1}\mathbf{v}^H \tag{5.293} $$

which multiplied by $\mathbf{Q}$, $\mathbf{H}_r\mathbf{Q} = \mathbf{Q} - 2\mathbf{v}(\mathbf{v}^H\mathbf{v})^{-1}\mathbf{v}^H\mathbf{Q}$, will have a very interesting property that we now demonstrate. The products $\mathbf{v}^H\mathbf{v}$ and $\mathbf{v}^H\mathbf{Q}$ can be expanded as

$$ \begin{aligned} \mathbf{v}^H\mathbf{v} &= \mathbf{I} + \mathbf{P}^H\mathbf{P} - \mathbf{B}^H\mathbf{P} - \mathbf{P}^H\mathbf{B} \\ \mathbf{v}^H\mathbf{Q} &= \mathbf{I} - \mathbf{P}^H\mathbf{B} \end{aligned} \tag{5.294} $$

Moreover, from (5.291) we know that

$$ \begin{aligned} \mathbf{P}^H\mathbf{P} &= \left(\mathbf{B}^H\mathbf{B}\right)^{-1/2}\mathbf{B}^H\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1/2} = \mathbf{I} \\ \mathbf{P}^H\mathbf{B} &= \mathbf{B}^H\mathbf{P} = \left(\mathbf{B}^H\mathbf{B}\right)^{1/2} \end{aligned} \tag{5.295} $$

which inserted into (5.294) reveals that

$$ \mathbf{v}^H\mathbf{v} = 2\mathbf{I} - 2\left(\mathbf{B}^H\mathbf{B}\right)^{1/2} = 2\mathbf{v}^H\mathbf{Q} \tag{5.296} $$

This condition means that the product of the Householder reflector by $\mathbf{Q}$ is in fact

$$\mathbf{H}_r\mathbf{Q} = \mathbf{Q} - 2\mathbf{v}(\mathbf{v}^H\mathbf{v})^{-1}\mathbf{v}^H\mathbf{Q} = \mathbf{Q} - \mathbf{v} = \mathbf{eP} \tag{5.297}$$

and we immediately conclude that the entries above the utmost lower right block of $\mathbf{Q}$ have been eliminated. Since the purpose of all of this is to perform the elimination in ${}_2\ddot{\underline{\mathbf{H}}}$, we multiply (5.297) by $\mathbf{R}$, which yields the same elimination

$$\mathbf{H}_r\mathbf{QR} = \mathbf{H}_r\ddot{\underline{\mathbf{H}}}(:, LT_s) = \mathbf{ePR} \tag{5.298}$$

We thus apply the Householder reflector to the entire ${}_2\ddot{\underline{\mathbf{H}}}$,

$${}_2\ddot{\underline{\mathbf{H}}} \overset{set}{=} \mathbf{H}_r\,{}_2\ddot{\underline{\mathbf{H}}} = {}_2\ddot{\underline{\mathbf{H}}} - 2\mathbf{v}(\mathbf{v}^H\mathbf{v})^{-1}\mathbf{v}^H\,{}_2\ddot{\underline{\mathbf{H}}} \tag{5.299}$$

and then normalize the bottom block row with

$$\mathbf{E} = \left(\mathbf{R}^H\mathbf{R}\right)^{-1/2}\mathbf{R}^H\mathbf{P}^H \quad \rightarrow \ddot{\underline{\mathbf{H}}}(last,:) \overset{set}{=} \mathbf{E}\ddot{\underline{\mathbf{H}}}(last,:) \tag{5.300}$$

The first step of the decomposition ends here, yet further steps are required to eliminate the rest of the entries above the block diagonal. The procedure described so far is still valid, except that it is applied to ${}_2\ddot{\underline{\mathbf{H}}}$ excluding the block columns that have already been eliminated and corresponding block rows. The process ends when the first block column of ${}_2\ddot{\underline{\mathbf{H}}}$ undergoes elimination, and at this point the lower $Ln_T \times Ln_T$ entries of ${}_2\ddot{\underline{\mathbf{H}}}$ are in fact $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}$ in its block lower triangular form.

There is still one subtlety with the method just described, and it concerns the product $\mathbf{B}^H\mathbf{B}$ in (5.291). To avoid the inherent loss of numerical information in computing this particular product, we use the following trick: we perform the singular value decomposition of $\mathbf{B}$ and then compute $\mathbf{P}$ as

$$\mathbf{B} = \mathbf{USV}^H \rightarrow \mathbf{P} = \mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1/2} = \mathbf{UV}^H \tag{5.301}$$

which circumvents the problem and replaces the formula in (5.291).

To this point, we have described how to obtain the operator $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}$ for simulation, but the ORTHO-TS-MMF ${}_2\ddot{\underline{\mathbf{\Phi}}}^H$ is yet to be determined. One obvious option is to apply the procedure just described to the identity matrix instead of ${}_2\ddot{\underline{\mathbf{H}}}$, which after normalization with (5.300) yields ${}_2\ddot{\underline{\mathbf{\Phi}}}^H$. What this approach has in obviousness, it also has in cost. It is slow. Another approach was tested in the forward direction, but was also slow, so will not be described. Experimentation revealed that the swiftest approach was to perform Gaussian elimination on the system $\ddot{\underline{\mathbf{H}}}^H_{\dot{\underline{\mathbf{\Phi}}}}\,{}_2\ddot{\underline{\mathbf{\Phi}}}^H = {}_2\ddot{\underline{\mathbf{H}}}^H$, since ${}_2\ddot{\underline{\mathbf{H}}}$ and $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}$ are already known. ${}_2\ddot{\underline{\mathbf{\Phi}}}$ will not be strictly required for the simulation since $\ddot{\underline{\mathbf{H}}}_{\dot{\underline{\mathbf{\Phi}}}}$ completely defines the noiseless channel after matched filtering, and also because the average noise power at the input of ${}_2\ddot{\underline{\mathbf{\Phi}}}$ remains unchanged by it

$$\begin{aligned} size\left(\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Phi}}}}\right)P_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Phi}}}}} &= E\left[\dot{\underline{\mathbf{n}}}^H_{\dot{\underline{\mathbf{\Phi}}}}\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Phi}}}}\right] = E\left[tr\left(\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Phi}}}}\dot{\underline{\mathbf{n}}}^H_{\dot{\underline{\mathbf{\Phi}}}}\right)\right] \\ &= tr\left({}_2\ddot{\underline{\mathbf{\Phi}}}\,{}_2\ddot{\underline{\mathbf{\Phi}}}^H E\left[\,{}_2\dot{\underline{\mathbf{n}}}\,{}_2\dot{\underline{\mathbf{n}}}^H\right]\right) = size\left(\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Phi}}}}\right)P_{{}_2\dot{\underline{\mathbf{n}}}} \quad \rightarrow P_{\dot{\underline{\mathbf{n}}}_{\dot{\underline{\mathbf{\Phi}}}}} = P_{{}_2\dot{\underline{\mathbf{n}}}} \end{aligned} \tag{5.302}$$

5.7.4 Monte Carlo Simulation, Numerical Results and Discussion

Monte Carlo simulation was performed using the Matlab® programming language, and

the implemented code is made available in Appendix C. Simulation starts by randomly generating and modulating a message of 64 complex symbols, either using BPSK, QPSK, 8-QAM or 16-QAM, yet the default modulation for simulation will be 16-QAM. After that, the channel operator ${}_2\ddot{\underline{\mathbf{H}}}$ is built using (5.288) and the ensemble of random parameters described previously, namely delays, phases, Doppler shifts and attenuations. The random message is applied to ${}_2\ddot{\underline{\mathbf{H}}}$ and noise is added, simulating the MIMO transmission. Then the ORTHO-TS-MMF ${}_2\ddot{\underline{\mathbf{\Phi}}}^H$ is computed and applied to the noisy output of the channel operator. The output is finally filtered by the unconstrained linear detector derived in Section 5.6.2, and applied to a slicer that makes the decisions on which were the transmitted constellation symbols. To yield reliable error rates this process is repeated so that at least 100000 symbol transmissions are simulated. Several plots were obtained for typical, randomly chosen channel realizations, and each plot expresses the error rate (or estimate of the average probability of symbol error) as a function of the SNR (stepped at 2 dB) at the output of the MIMO channel, before matched filtering. The plots will now be described.

Figure 5.19 depicts the error rate for the different types of modulation and for two

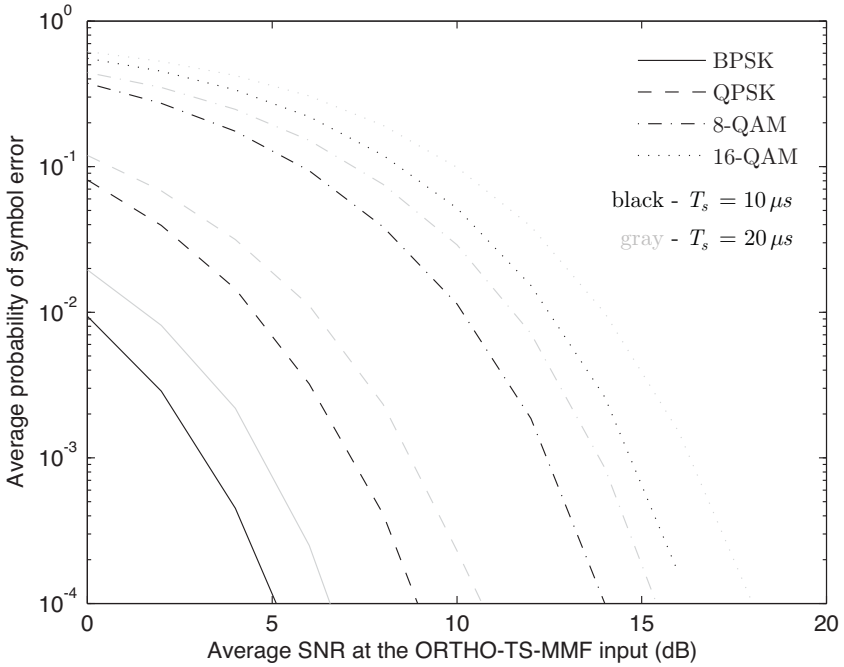

FIGURE 5.19 Average error rates for several symbol constellation sizes and two signalling intervals, for $n_T = 2$ and $n_R = 4$ with equispaced fit within $d_{\max} = 0.5\lambda$.

signalling intervals for a MIMO arrangement of $n_T = 2$ and $n_R = 4$ with equispaced fit within $d_{\max} = 0.5\lambda$. We conclude that enlarging the constellation has the expected consequence of increasing the average error probability, as for a given SNR the symbols will be closer to one another as the constellation enlarges. In addition, slightly increasing the signalling interval from $T_s = 10\,\mu s$ to $T_s = 20\,\mu s$ also increases the error rate. This happens because for $\sigma_\tau = 3\,\mu s$, such an increase reduces the ability of the detector to resolve the multipath components, which means that taking advantage of delay diversity will be less possible.

Plots for the average error rate as the number of transmitter antennas increases are shown in Figure 5.20. We see that increasing the number of transmitter antennas has the

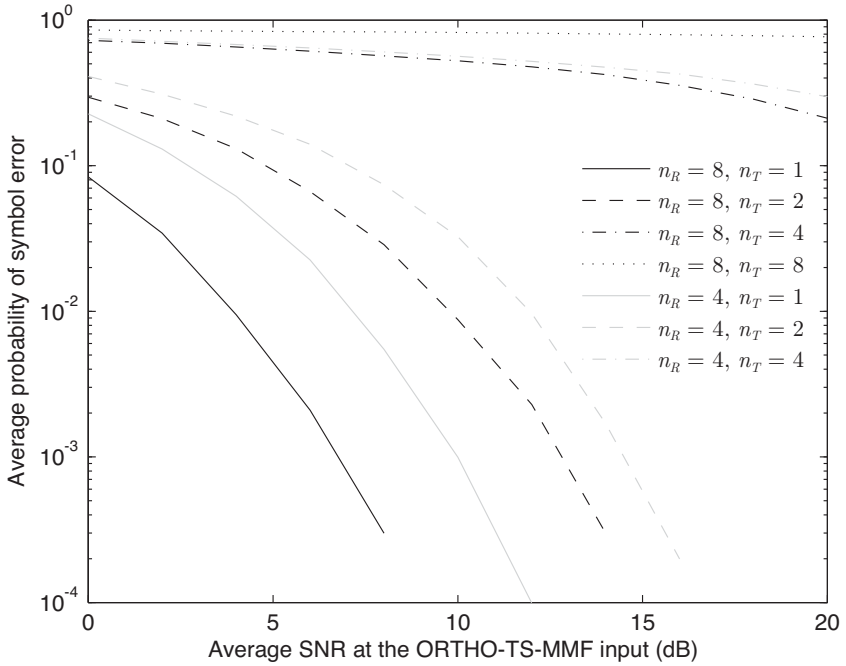

FIGURE 5.20 Average error rates for two receiver sizes with equispaced fit within $d_{\max} = 0.5\lambda$ and increasing number of transmitter antennas, for $T_s = 100\,\mu s$.

effect of increasing the error probability, which is legitimate because the symbols transmitted from different transmitter antennas will interfere with one another, causing confusion at the receiver. The penalty for higher data rates (using the same bandwidth and thus improving the spectral efficiency) is an increase in the probability of symbol error. Nevertheless, for $n_T = 2$ the error rates are indeed acceptable, becoming even lower as the number of receiver antennas increases.

A different situation appears in Figure 5.21. One transmitter antenna is assumed, and the number of receiver antennas increases. We readily gather that increasing the size of the receiver array always has the consequence of decreasing the error probability, which means that the receiver is able to use reception diversity to improve the error performance. This improvement is even larger when free scaling of the receiver array is permitted, which is natural because the receiver antennas will be positioned further apart, therefore decreasing the cross-correlation between the received signals and increasing diversity. Equivalent plots but for the case $n_T = 2$ and $T_s = 1\,\mu s$ are shown in Figure 5.22, and we infer the same behaviour. The best performance takes place for larger n_R, but when the receiver antennas' placement is restricted to a maximum length it becomes harder to attain lower error rates. Finally, plots for $n_T = 4$ are displayed in Figure 5.23, but this time for 10 times the signalling rate, $T_s = 0.1\,\mu s$, which allows for much lower error rates gained by multipath resolution. Again, we observe the expected behaviour of lower error rates for larger receiver arrays, confirming the previous results and showing that it is the same regardless of the number of transmitter antennas.

To assess the variation of the error performance with the signalling rate, plots for sev-

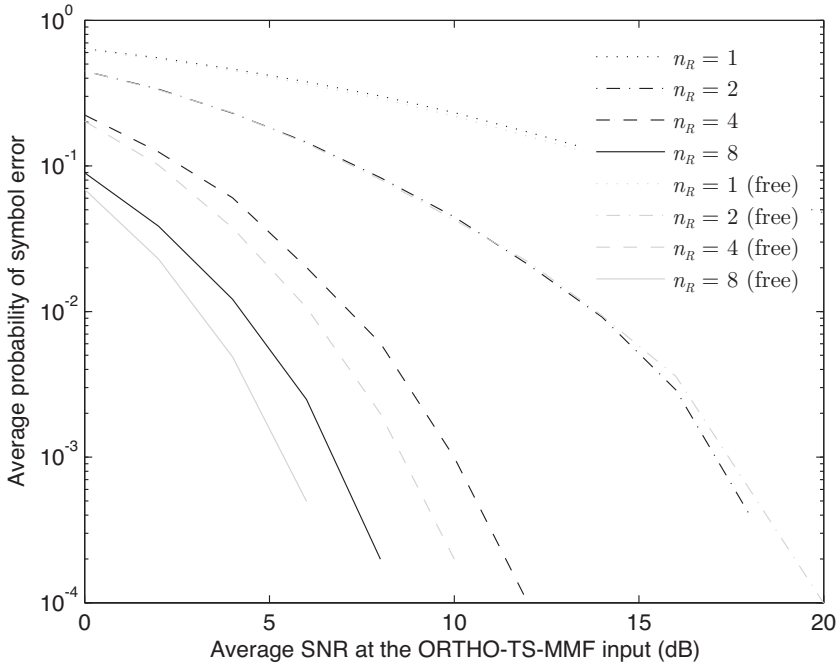

FIGURE 5.21 Average error rates for one transmitter antenna and several receiver sizes (equispaced fit within $d_{\max} = 0.5\lambda$ and free scaling considered), for $T_s = 100\,\mu s$.

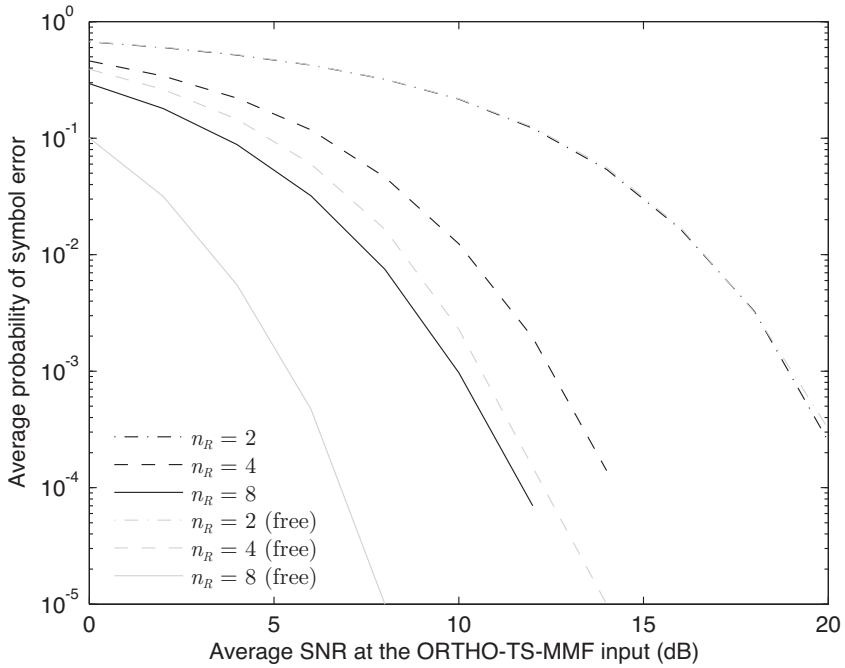

FIGURE 5.22 Average error rates for two transmitter antennas and several receiver sizes (equispaced fit within $d_{\max} = 0.5\lambda$ and free scaling considered), for $T_s = 1\,\mu s$.

eral T_s were obtained as shown in Figure 5.24. We see that the lowest signalling interval is responsible for the lowest error rates, which was already explained to be a consequence of multipath diversity resolution. As the signalling interval increases below $\sigma_\tau = 3\,\mu s$, the error rate also steadily increases, which is explained by an increasing difficulty of resolving the multipath components since these are now more and more concentrated and approach a situation of flat fading. When the RMS delay spread becomes much lower than the sig-

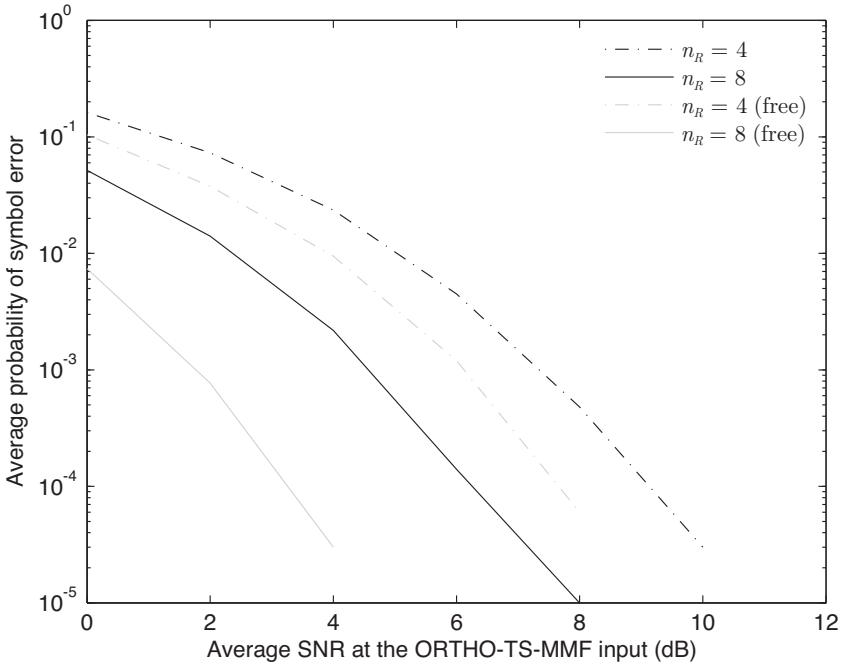

FIGURE 5.23 Average error rates for four transmitter antennas and several receiver sizes (equispaced fit within $d_{\max} = 0.5\lambda$ and free scaling considered), for $T_s = 0.1\,\mu s$.

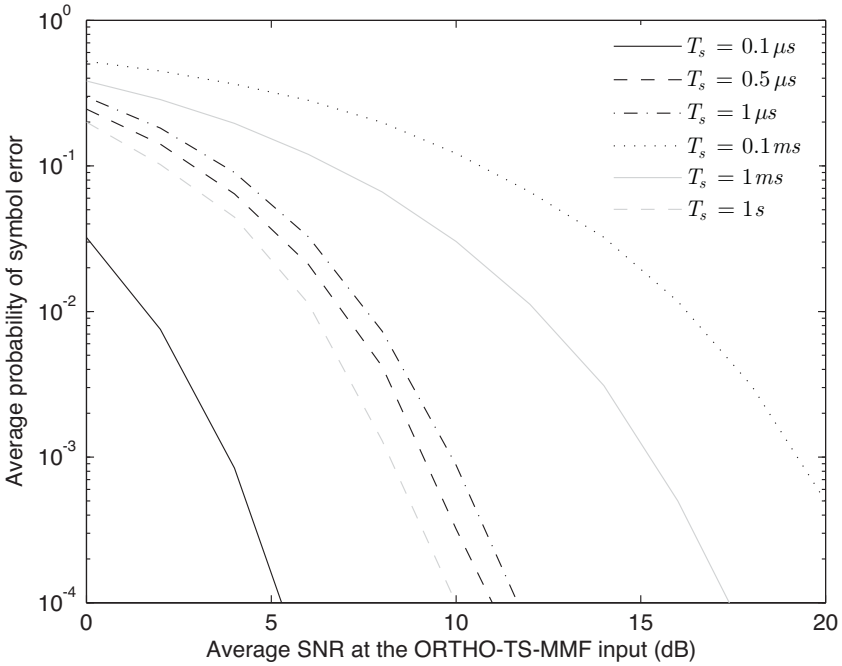

FIGURE 5.24 Average error rates for several signalling rates, and a MIMO configuration of $n_T = 2$ and $n_R = 4$ with equispaced receiver antenna fit within $d_{\max} = 0.5\lambda$.

nalling rate the error performance severely degrades, as can be seen in the error rate curve of $T_s = 0.1\,ms$. In this situation there is no multipath diversity available for resolution and the receiver cannot compensate for the ISI introduced. However, as the signalling rates get lower and lower, the ISI will become negligible and, despite the fact that the channel may enter fast-fading, the error rates will be significantly reduced. The curves for $T_s = 1\,ms$ and $T_s = 1\,s$ demonstrate this fact.

To study the influence of the Doppler shift on the error performance we have included the plots in Figure 5.25 and Figure 5.26. We see that when the channel is fast-fading

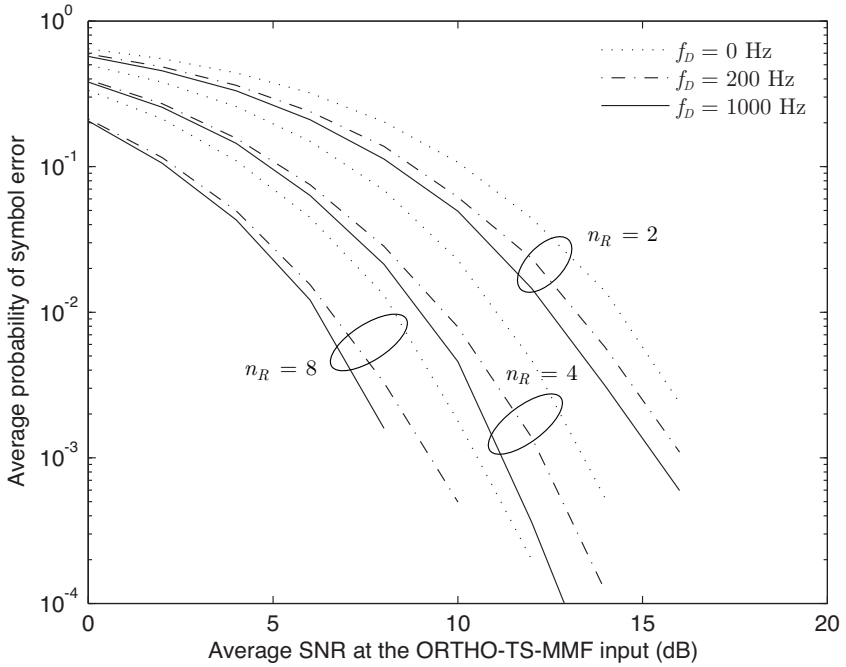

FIGURE 5.25 Average error rates for several maximum Doppler shifts, and a MIMO configuration of $n_T = 2$ and several n_R with equispaced receiver antenna fit within $d_{\max} = 0.5\lambda$, for $T_s = 1s$.

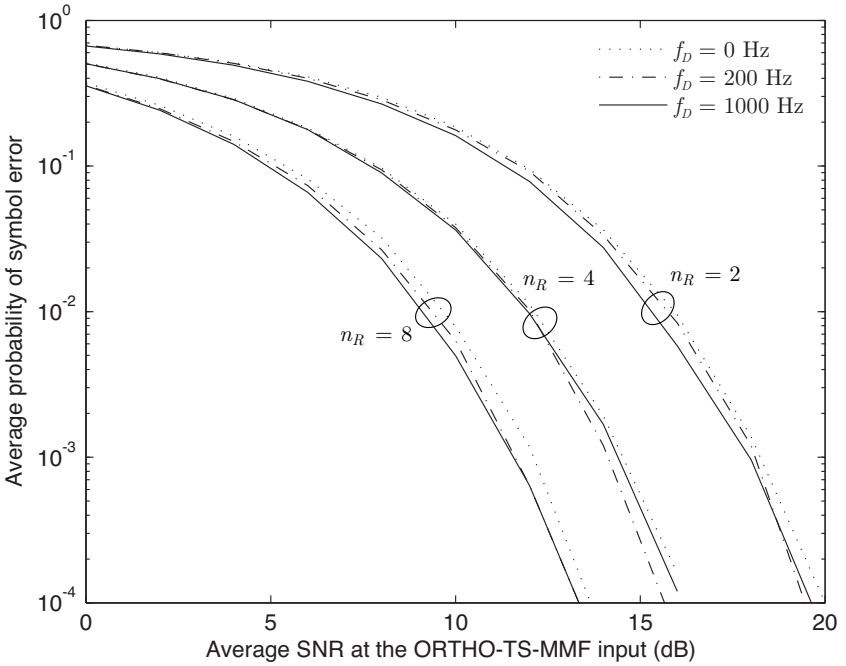

FIGURE 5.26 Average error rates for several maximum Doppler shifts, and a MIMO configuration of $n_T = 2$ and several n_R with equispaced receiver antenna fit within $d_{\max} = 0.5\lambda$, for $T_s = 1\mu s$.

(the channel varies in a signalling interval basis), which happens for $T_s = 1s$, the receiver can use the extra variability introduced by the Doppler shift to improve the detection, irrespective of the number of receiver antennas. If this is not the case, i.e. if the signalling

rate is much larger than the Doppler shift, which happens for $T_s = 1\mu s$, then the MIMO channel is practically invariant during a large number of signalling intervals, making it difficult for the receiver to accomplish a significant reduction in the error rate. This is a manifestation of Doppler diversity absence, or lack of time-selectivity. One realizes that one of the main advantages of the ORTHO-TS-MMF is its channel variability awareness, which makes it feasible to exploit the inherent Doppler diversity to provide better estimates of the transmitted symbols.

To complete the simulation results, plots for error rate variation with the maximum antenna separation are provided in Figure 5.27. It becomes clear that, as the maximum

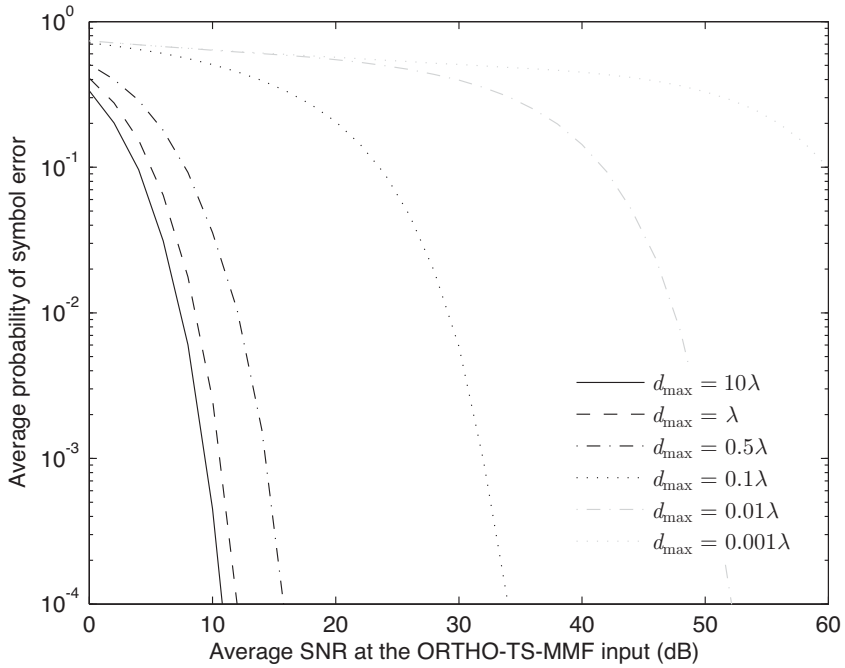

FIGURE 5.27 Average error rates for several maximum receiver antenna separations, and a MIMO configuration of $n_T = 2$ and $n_R = 4$ with equispaced receiver antenna fit within $d_{\max}$, for $T_s = 1\mu s$.

separation decreases the error performance deteriorates. As the antennas become closer and closer to one another they become increasingly correlated to the extent of not permitting the compensation of both intra-sample ISI and time selectivity.

5.8 *Conclusion*

This chapter successfully solved a number of difficulties that persisted in the treatment of the overspread MIMO LTV channel. Using a baseband set-up it started by defining a continuous input/output model of digital transmission that did not ignore the time-varying nature of the wireless channel. Then it aimed for an optimal model discretization using a channel-extracted orthogonal basis. This was accomplished by orchestrating a theory of continuous, discrete, and hybrid matrix operators that impressively simplified all MIMO calculations. It was possible to build a semi-orthogonal version of the channel that, in turn, was sufficient to devise a simple discrete version of the input/output model that was in all respects similar to a flat-fading MIMO channel. A semi-orthogonal matched filter was also derived which, above all: 1) maintained the white-nature of the input noise and 2)

maintained the optimality from a maximum likelihood detection perspective.

In sequence, a full frequency description of the time-varying MIMO channel was given in terms of frequency operators, a description that was extended to the discrete input/output model just derived. Using this approach, novel formulas for the spectral factorization of time-varying MIMO filters were obtained. A section on noise whitening and full discretization of the non-orthogonal matched filter was also presented, which made new incursions into the consequences of conventional sampling within the MIMO system.

Finally, to test the performance of the new model a new section on linear detection was introduced. Three linear detector operators were designed, one of these with the intent of optimal detection in the MMSE sense. The latter was tested with success using Monte Carlo simulation, allowing for considerable insight into the error performance of the Doppler-aware MIMO matched filter.

CHAPTER 6

CONCLUSIONS AND FUTURE RESEARCH

6.1 *Conclusions*

This work has produced a number of significant results. The information-theoretic study of MIMO systems under receiver-sided correlated flat-fading has revealed exact, closed-form formulas for the ergodic capacity of the MIMO channel. The more involved case of more receiver antennas than transmitter antennas was also successfully addressed. The results were extended to situations of independent envelope fading, yielding new closed-form formulas. Abundant plots of the derived formulas show a substantial capacity reduction when correlation and/or non-isotropic scattering is present, yet confirm the enormous advantages of using MIMO systems. For the case of a single transmitter antenna it was confirmed that the maximal-ratio combiner attains capacity. Transmitter-sided correlation was studied by assuming no CSI at the transmitter, showing that a pronounced capacity stagnation occurs for more than three transmitter antennas. In addition, it was concluded that the greatest advantage comes from equipping a mobile unit with more antennas than the base station, irrespective of the correlation. Fixed-length antenna arrays withhold the capacity potential of independent fading, and it was shown that equispaced linear arrays do not attain the maximum capacity. A method to improve capacity was devised and validated.

The study of overspreading in linear time-varying MIMO transmission was also a success. It is a general study that contemplates every single type of MIMO channel that one can imagine. It studies delay overspreading and Doppler overspreading, the phenomenons responsible for frequency selectivity and time-selectivity, respectively, but also encompasses flat-fading. An extremely valuable introduction of continuous, discrete and hybrid operators was responsible for attaining the results that were pursued. The time-varying MIMO channel was factorized (both in time and frequency) into semi-orthogonal and causal terms, creating the necessary momentum for the full discretization of the input/

output model of the MIMO system and the derivation of the optimal semi-orthogonal time-shear matrix matched filter, the designated ORTHO-TS-MMF. With the purpose of studying the error performance of the model, time-varying linear detectors were deduced in the MMSE sense, and several Monte Carlo simulations were performed. Essentially, they showed that the error performance is boosted when the number of receiver antennas increases, and that the penalty of achieving higher data rates (within the same bandwidth) with more transmitter antennas is an increase in error probability. In some cases the penalty can be overcome by adding more antennas to receiver. Additionally, they confirmed that the error performance deteriorates as correlation increases, i.e. when the antenna separations decrease. More importantly, they unveiled that the ORTHO-TS-MMF combined with the unconstrained linear detector can take advantage of both delay-diversity and Doppler diversity, significantly reducing the error probability. This means that frequency selectivity and time-selective can be put to one's favour because they are the effects of delay-Doppler diversity. Accordingly, the flat-fading channel has proven to dictate the worst error performance overall. In effect, a form of time-frequency equalization in the context of MIMO systems has been accomplished, all this without the need of an adaptive equalization algorithm: provided that channel tracking is accurate, the ORTHO-TS-MMF and the detector are themselves already adapted to the channel variations.

6.2 *Future Research*

Despite the self-contained, all-inclusive nature of this thesis, it is clear that open issues remain that point to new directions of research. With time they may lead to conclusions that set new prospects on the performance of MIMO systems. The author has identified the following lines of investigation:

1. Find closed-form formulas for the information-theoretic capacity of MIMO channels under transmitter-sided correlated fading, when channel state information is available at the transmitter; or, more generically, under two-sided correlated fading;

2. Extend the derivations of the analytic expressions to the ergodic capacity of MIMO channels under Nakagami-m distributed correlated fading and, eventually, doubly selective wireless channels;

3. Test the performance of the ORTHO-TS-MMF when the Viterbi algorithm is used for maximum likelihood sequence estimation; study reduced-complexity ML detectors;

4. Compare the error performance of the ORTHO-TS-MMF to the error performance of MIMO adaptive equalization techniques available in the literature;

5. Derive the minimum mean-square error decision-feedback (MMSE-DF) detector for the discretized input/output model and test its performance; one should expect a superior performance as compared to that of the unconstrained linear detector of this thesis;

6. Study the performance of the discretized input/output model when applied to MIMO-OFDM, MIMO MC-CDMA and MIMO WCDMA systems;

7. Derive analytic formulas for the probability of symbol error of the discretized input/output MIMO model for different types of modulations;

8. Derive the optimal time-varying pulse shaping MIMO filter when channel state information is available at the transmitter;

9. Study the impact of space-time coding (e.g. space-time turbo trellis codes) on the results of the thesis and develop strategies of improvement.

APPENDIX

APPENDIX A

ALGORITHM IN THE MATHEMATICA® PROGRAMMING LANGUAGE TO COMPUTE THE COEFFICIENTS $a_{k,n}$ IN (4.291):

```
coef[k_,n_]:=
Block[{l,i,coefsums=Range[k]},
    For[l=1,l<n-k,l++,
        For[i=k,i>0,i--,coefsums[[i]]=
            i*Sum[coefsums[[j]],{j,1,i}]]];
            If[k==n,Return[1],
            Return[Sum[coefsums[[j]],{j,1,k}]
        ]
    ]
]
```

APPENDIX B

ALGORITHM IN THE MATHEMATICA® PROGRAMMING LANGUAGE TO COMPUTE THE ROW (COLUMN) INDICES (ANTENNA POSITIONS WITHIN $d_{\max}$) OF THE $p \times p$ PRINCIPAL SUBMATRIX OF Σ WITH MINIMUM FROBENIUS NORM:

```
LeastCorrPos[covarmatrix_,p_]:=
Block[{i,currsum,minsum,minset,subsets=
Subsets[Range[Length[covarmatrix]],{p}]},
    For[i=1,i<=Length[subsets],i++,currsum=
        Sum[Abs[covarmatrix[[subsets[[i]],
        subsets[[i]]]][[k,l]],{k,1,p},{l,1,p}];
        If[i==1,minsum=currsum;minset=subsets[[1]],
            If[currsum<minsum,minsum=currsum;
                minset=subsets[[i]]
            ]
        ]
    ];
    Return[minset]
]
```

APPENDIX C

CODE OF THE MONTE CARLO SIMULATION OF THE ORTHO-TS-MMF WITH UNCONSTRAINED LINEAR DETECTION, IN THE MATLAB® PROGRAMMING LANGUAGE:

mimo_channel_operator_generate.m

```
function H=mimo_channel_operator_generate(L,nT,nR,Ts,Nmulti,dshift,rho,u,uaoa)
%--------declare needed variables
fs=1/Ts; %signalling rate
rmsdspread=3e-6; %rms delay spread
fc=2.6e9; %center frequency
c=3e8; %speed of light;
lambda=c/fc;
d=rho*lambda; %antenna separation
srrcdelay=20; %srrc filter delay
roll=1/3; %filter roll-off

Dopplerw=zeros(nR,nT,Nmulti);
tau=zeros(nR,nT,Nmulti);
phi=zeros(nR,nT,Nmulti);
alpha=zeros(nR,nT,Nmulti);

for j=1:nT
    for n=1:Nmulti
        Dopplerw(1,j,n)=2*pi*dshift*cos(2*pi*uaoa(1,j,n));
        tau(1,j,n)=-rmsdspread*log(1-u(1,j,n));
        phi(1,j,n)=-mod(2*pi*fc*tau(1,j,n),2*pi);
        alpha(1,j,n)=exp(-tau(1,j,n)/(2*rmsdspread));
    end
end

for i=2:nR
    for j=1:nT
        for n=1:Nmulti
            Dopplerw(i,j,n)=Dopplerw(1,j,n);
            tau(i,j,n)=tau(1,j,n)+d/c*(i-1)*cos(2*pi*uaoa(1,j,n));
            %phi(i,j,n)=-2*pi*fc*tau(i,j,n);
            phi(i,j,n)=phi(1,j,n)-2*pi*(rho)*(i-1)*cos(2*pi*uaoa(1,j,n));
            alpha(i,j,n)=alpha(1,j,n);
        end
    end
end

maxdelay=2*srrcdelay+ceil(fs*max(max(max(tau))));
H=zeros(2*nR*(L+maxdelay),nT*L);
for l=0:L-1
    for k=2*l:2*(L+maxdelay)-1
        H(k*nR+1:k*nR+nR,l*nT+1:l*nT+nT)=sum(alpha(1:nR,1:nT,1:Nmulti).*...
            exp(complex(0,phi(1:nR,1:nT,1:Nmulti)+Dopplerw(1:nR,1:nT,1:Nmulti)*k*Ts/2)).*...
            (4*roll*(cos((1+roll)*pi*1/2*(k-2*l-2*srrcdelay-2*fs*tau(1:nR,1:nT,1:Nmulti)))+...
            (sin((1-roll)*pi*1/2*(k-2*l-2*srrcdelay-2*fs*tau(1:nR,1:nT,1:Nmulti))))./...
            (4*roll*1/2*(k-2*l-2*srrcdelay-2*fs*tau(1:nR,1:nT,1:Nmulti)))).*...
            (pi*sqrt(Ts)*(1-(4*roll*1/2*(k-2*l-2*srrcdelay-2*fs*tau(1:nR,1:nT,1:Nmulti))).^2))),3);
    end
end

end
```

mimo_channel_operator_orthogonalize.m

```
function [O,Ho]=mimo_channel_operator_orthogonalize(H,L,nT)
Horiginal=H;
H_height=size(H,1); H_width=size(H,2);

%------------------- O - solution 1 ---------------
%EI=eye(H_width);
%uMat=zeros(H_height,H_width);
%O=[zeros(H_width,H_height-H_width) eye(H_width)];
%-------------------------------------------
%------------------- O - solution 2 ---------------
%EI=eye(H_width);
%O=eye(H_height);
%-------------------------------------------

for j=L:-1:1
    Hcol=H(1:H_height-(L-j)*nT,(j-1)*nT+1:j*nT);
    [Q,R]=qr(Hcol,0); %Q=Q(:,1:nT); R=R(1:nT,:);% Q=flipud(Q);
    B=Q(H_height-(L-j+1)*nT+1:H_height-(L-j)*nT,:);
    [U,S,V]=svd(B,0); P=U*V';

    u=Q; u(H_height-(L-j+1)*nT+1:H_height-(L-j)*nT,:)=...
    u(H_height-(L-j+1)*nT+1:H_height-(L-j)*nT,:)-P;%*sqrtm(Q'*Q);

    E=sqrtm(R'*R)\R'*P';

    H(1:H_height-(L-j)*nT,1:j*nT)=(H(1:H_height-(L-j)*nT,1:j*nT)...
    -2*(u/(u'*u)*u'*H(1:H_height-(L-j)*nT,1:j*nT)));
    H(H_height-(L-j+1)*nT+1:H_height-(L-j)*nT,1:j*nT)=...
    E*H(H_height-(L-j+1)*nT+1:H_height-(L-j)*nT,1:j*nT);


    %------------------- O - solution 1 ---------------
    %uo=[u;zeros((2*L-j)*nT,nT)]; uMat(:,H_width-(2*L-j+1)*nT+1:H_width-(2*L-j)*nT)=uo;
    %EI(H_width-(2*L-j+1)*nT+1:H_width-(2*L-j)*nT,H_width-(2*L-j+1)*nT+1:H_width-(2*L-j)*nT)=E;
    %-------------------------------------------

    %------------------- O - solution 2 ---------------
    %EI(H_width-(2*L-j+1)*nT+1:H_width-(2*L-j)*nT,H_width-(2*L-j+1)*nT+1:H_width-(2*L-j)*nT)=E;
    %O(1:H_height-(2*L-j)*nT,1:H_height-(2*L-j)*nT)=(eye(H_height-(2*L-j)*nT)-...
    %2*(u/(u'*u)*u'))*O(1:H_height-(2*L-j)*nT,1:H_height-(2*L-j)*nT);
    %-------------------------------------------

end
Ho=H(H_height-L*nT+1:H_height,:);
%------------------- O - solution 1 ---------------
%for j=1:2*L
%    uo=uMat(:,(j-1)*nT+1:j*nT);
%    O=O-2*O*(uo/(uo'*uo)*uo');
%end
%O=O'; O=O/EI;
%-------------------------------------------
%------------------- O - solution 2 ---------------
%O=O'; O=O(:,H_height-H_width+1:end); O=O/EI;
%-------------------------------------------
%------------------- O - solution 3 ---------------
O=Horiginal/Ho;
%-------------------------------------------
end
```

mimo_linear_detection_simulation.m

```
%------------------------ detection simulation ----------

msg_length=64; Nmulti=8; snrdB_min=0; snrdB_max=60; snrdB_step=2;
Pevec=zeros(1,floor((snrdB_max-snrdB_min)/snrdB_step)+1);

count=0; test_count=0;
```

```
for M=16 %constellation size

h = modem.qammod(M); modsym=modulate(h,0:M-1); %modulator object

u=rand(1,32,Nmulti); uaoa=rand(1,32,Nmulti); %random uniform samples

%--------------test routine--------------

niter=ceil(100000/msg_length); %number of iterations

msg_test=randint(msg_length,niter,M); %build message
a_test=modulate(h,msg_test); Pa=mean(mean(abs(a_test).^2)); PadB=10*log10(Pa);

dscount=0;
for dshift=200 %maximum Doppler shift
    dscount=dscount+1;

    nTcount=0;
    for nT=2
        nR=4;
        L=msg_length/nT;
        if nR>1, rho=0.5/(nR-1); else rho=0; end %antenna separation
        %rho=0.5; %free length separation
        nTcount=nTcount+1;

        Tscount=0;
        for Ts=100e-6
            Tscount=Tscount+1;test_count=test_count+1;

            H=mimo_channel_operator_generate(L,nT,nR,Ts,Nmulti,dshift,rho,u,uaoa); %generate channel
            [O,Ho]=mimo_channel_operator_orthogonalize(H,L,nT); %orthogonalize channel

            rup_test=H*a_test; %simulate transmission
            nup_test_unit=(randn(size(rup_test,1),niter)+...
                  i*randn(size(rup_test,1),niter)); %generate noise
            Prup=mean(mean(abs(rup_test).^2)); PrupdB=10*log10(Prup);

            %--------- compute error rate for several snr -------------
            snr_count=0;
            for snr_rupdB=snrdB_min:snrdB_step:snrdB_max
                count=count+1; snr_count=snr_count+1; Pe=0; nerrors=0;

                snr_rup=10^(snr_rupdB/10); PnupdB=PrupdB-snr_rupdB; Pnup=Prup/snr_rup;
                sqrthalfPnup=sqrt(Pnup/2); %noise standard deviation in each quadrature component
                nup_test=sqrthalfPnup*nup_test_unit; %scale noise power
                no_test=O'*nup_test; Pno=mean(mean(abs(no_test).^2)); %compute noise power
                rupn_test=rup_test+nup_test; %add noise

                ron_test=O'*rupn_test; %simulate ortho-ts-mmf

                %---------- build D ------------
                Id=(Pno/Pa)*eye(length(Ho),length(Ho));
                D=Ho'/(Ho*Ho'+Id); %detector operator

                aest_test=D*ron_test; %simulate detector

                %-------------- slicer ---------
                for iter=1:niter
                    for n=1:msg_length
                        mindist=abs(aest_test(n,iter)-modsym(1));
                        adetect=modsym(1);
                        for m=2:M
                            dist=abs(aest_test(n,iter)-modsym(m));
                            if dist<mindist
                                mindist=dist;
                                adetect=modsym(m);
                            end
                        end
                        if adetect~=a_test(n,iter)
                            nerrors=nerrors+1;
                        end
```

```
                    end

                end
                %waitbar(count/(4*(floor((snrdB_max-snrdB_min)/snrdB_step)+1)));
                Pevec(test_count,snr_count)=nerrors/(msg_length*niter); %error rates

            end
        end
    end
end
end
x=snrdB_min:snrdB_step:snrdB_max;
semilogy(x,Pevec); axis([0 20 1e-4 1]) %plot
```

REFERENCES

[1] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[2] T. S. Rappaport, *Wireless Communications: Principles and Practice,* 2nd ed. NJ: Prentice-Hall, 2002.

[3] J. D. Parsons, *The Mobile Radio Propagation Channel,* 2nd ed. Chichester: John Wiley & Sons, 2000.

[4] M. Patzold, *Mobile Fading Channels*. Chichester: John Wiley & Sons, 2002.

[5] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. on Inf. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.

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

[7] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. on Commun.*, vol.42, no. 234, pp. 1740-1751, Feb./Mar./ Apr. 1994.

[8] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. on Inf. Theory*, vol. 50, no. 9, pp. 1859-1874, Sept. 2004.

[9] S. N. Diggavi et al., "Great expectations: the value of spatial diversity in wireless networks," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 219-270, Feb. 2004.

[10] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[11] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," in *Wireless Personal Communications*. Norwell, MA: Kluwer Academic, pp. 311–335, 1998.

[12] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when Using Multi-Element Antennas", *Bell Labs Technical Journal*, Vol. 1, No. 2, Autumn 1996, pp 41-59.

[13] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications - a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198-218, Feb. 2004.

[14] H. Bölcskei and A. J. Paulraj, "Multiple-input multiple-output (MIMO) wireless systems" in *The Communications Handbook*, 2nd ed., J. Gibson, ed., CRC Press, pp. 90.1 - 90.14, 2002.

[15] M. A. Jensen and J. W. Wallace, "A review of antennas and propagation for MIMO wireless communications," *IEEE Trans. Antennas Propagat.*, vol. 52, no. 11, pp. 2810-2824, Nov. 2004.

[16] D. Gesbert, M. Shafi, D. S. Shiu, P. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[17] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[18] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[19] D. Barry, E. Lee, and D. Messerschmitt. *Digital Communication*, 3rd ed. New York: Springer, 2003.

[20] C. A. Balanis, *Antenna Theory: Analysis and Design*, 2nd ed. New York: John Wiley, 1997.

[21] D.K. Cheng, *Fundamentals of Engineering Electromagnetics*, 2nd ed. Addison Wesley, 1993.

[22] W. C. Y. Lee. *Mobile Communications Engineering*, 2nd ed. New York: McGraw-Hill, 1998.

[23] A. Ishimaru, "Wave propagation and scattering in random media and rough surfaces," *Proceedings of the IEEE*, vol. 79, no. 10, pp. 1359-1366, Oct. 1991.

[24] G. L. Turin et al., "A statistical model of urban multipath propagation," *IEEE Trans. Veh. Technol.*, vol. VT-21, pp. 1–9, Fev. 1972.

[25] H. Suzuki, "A Statistical Model for Urban Radio Propagation", *IEEE Trans. on Commun.,* vol. COM-25, no. 7, pp. 673-680, 1977.

[26] L. Boithias, *Radio Wave Propagation*. New York: McGraw-Hill, 1987.

[27] D. C. Cox and R. P. Leck, "Distributions of multipath delay spread and average excess delay for 910 MHz urban mobile radio paths." *IEEE Trans. Ant. Prop.*, vol. AP-23, no. 2, pp. 206-213, Mar. 1975.

[28] D. C. Cox, "Delay Doppler characteristics of multipath propagation at 910 MHz in a suburban mobile radio environment," *IEEE Trans. Ant. Propag.*, vol. AP-20, no. 5, pp. 625-635, Sept. 1972.

[29] T. Aulin, "A Modified Model for the Fading Signal at a Mobile Radio Channel", *IEEE Trans. Veh. Techn.*, vol. VT-28, pp. 182-203, Aug. 1979.

[30] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: IEEE Press, 1996.

[31] M. Nakagami, "The m-distribution—A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. C. Hoffman, Ed. New York: Pergamon, pp. 3–36, 1960.

[32] G.L. Turin, "Introduction to spread-spectrum antimultipath techniques and their application to urban digital radio," *Proc. IEEE.*, vol. 68, no. 3, pp. 328-353, Mar. 1980.

[33] T. Aulin, "Characteristics of a digital mobile radio channel," I*EEE Trans. Veh. Technol.*, vol. VT-30, no. 2, pp. 45–53, May 1981.

[34] L. Rubio et al., "The use of semi-deterministic propagation models for the prediction of the short-term fading statistics in mobile channels," *Proc. IEEE VTC*, vol. 3, pp. 1460–1464, 1999.

[35] W. R. Braun and U. Dersch, "A physical mobile radio channel model," *IEEE Trans. Veh. Technol.*, vol. VT-40, no. 2, pp. 472-482, May 1991.

[36] R. J. C. Bultitude, R. F. Hahn, and R. J. Davies, "Propagation considerations for the design of an indoor broad-band communications system at EHF," *IEEE Trans. Veh. Technol.*, vol. 47, no. 1, pp. 235-245, Feb. 1998.

[37] A. Domazetovic et al., "Estimating the Doppler Spectrum of a Short-Range Fixed Wireless Channel", *IEEE Communications Letters*, vol. 7, no. 5, pp. 227-229, May 2003.

[38] P. Bello, "Characterization of Randomly Time-Variant Linear Channels," *IEEE Trans. Commun. Systems*, vol. 11, no. 4, pp. 360-393, Dec. 1963.

[39] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th Edition. New York: McGraw-Hill, 2002.

[40] J. H. Winters, "Smart antennas for wireless systems," *IEEE Personal Commun. Mag.*, vol. 5, no. 1, pp. 23-27, Feb. 1998.

[41] F. Patenaude, J. H. Lodge, and J.-Y Chouinard, "Noncoherent diversity reception over Nakagami-fading channels," *IEEE Trans. Commun.*, vol. 46, no. 8, pp. 985-991, Aug. 1998.

[42] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population", *Biometrika 20A*, pp. 32–43, 1928.

[43] A. T. James, "The distribution of the latent roots of the covariance matrix," *Ann. Math. Statist.*, vol. 31, pp. 151–158, 1960.

[44] A. Edelman, "Eigenvalues and Condition Numbers of Random Matrices," Ph.D. dissertation, MIT, Cambridge, MA, 1989.

[45] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[46] H. Shin, M. Z. Win, J. H. Lee, and M. Chiani, "On the capacity of doubly correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 2253–2265, May 2006.

[47] M. Kiessling and J. Speidel, "Exact ergodic capacity of MIMO channels in correlated Rayleigh fading environments," in *Proc. Int. Zurich Seminar on Communications*, Zurich, Switzerland, Feb. 2004, pp. 128–131.

[48] M. Kiessling and J. Speidel, "Mutual information of MIMO channels in correlated Rayleigh fading environments—A general solution," in *Proc. IEEE Int. Conf. Communications*, Paris, France, Jun. 2004, vol. 2, pp. 814–818.

[49] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlatedMIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[50] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, 2nd rev. ed. New York: Wiley-Interscience, 2005.

[51] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Hoboken, NJ: John Wiley & Sons, 2003.

[52] M. L. Mehta, "Random matrices," 3rd ed. *Pure and Applied Mathematics (Amsterdam), 142*. Amsterdam: Elsevier/Academic Press, 2004.

[53] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia: SIAM, 2000.

[54] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Comm.*, vol. 16, no. 8, pp. 1451-1458, 1998.

[55] J. Fuhl, J. P. Rossi, and E. Bonek, "High-resolution 3-D direction-of-arrival determination for urban mobile radio," *IEEE Trans. Antennas Propagat.*, vol. 45, no. 4, pp. 672–682, Apr. 1997.

[56] P. C. Fannin and A. Molina, "Analysis of mobile radio channel sounding measurements in inner city dublin at 1.808 GHz," *Proc. Inst. Elect. Eng. Commun.*, vol. 143, pp. 311–316, 1996.

[57] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[58] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: Moment-generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.

[59] A. M. Sengupta and P. P. Mitra, "Capacity of multivariate channels with multiplicative noise: Random matrix techniques and large-N expansions (2)," *J. Statist. Phys.*, vol. 125, no. 5/6, Dec. 2006.

[60] M. Kang and M. S. Alouini, "Capacity of correlated MIMO Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[61] C. Itzykson and J.-B. Zuber, "Planar approximation 2," *J. Math. Phys.*, vol. 21, pp. 411–421, 1980.

[62] R. W. Butler and A. T. A. Wood, "Laplace approximations for hypergeometric functions with matrix argument," *Ann. Statist. 30*, no. 4, pp. 1155–1177, 2002.

[63] R. Gutiérrez, J. Rodríguez, and A. J. Sáez. "Approximation of Hypergeometric Functions with Matricial Argument trough their Development in Series of Zonal Polynomials." *Electronic Transactions in Numerical Analysis*, Vol.11, pp. 121-130, 2000.

[64] P. Koev and A. Edelman, "The efficient evaluation of the hypergeometric function of a matrix argument", *Mathematics of Computation*, vol. 75, no. 254, pp. 833-846, 2006.

[65] E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*. New York: Academic, 1959.

[66] H. Weyl, *The Theory of Groups and QuantumMechanics*. New York: Dover, 1950.

[67] G. James and M. Liebeck, *Representations and Characters of Groups*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[68] W. Fulton and J. Harris, *Representation Theory: A First Course*. New York: Springer, 1999.

[69] H. Weyl, *The Classical Groups: Their Invariants and Representations*. Princeton, NJ: Princeton Univ. Press, 1997.

[70] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed. Oxford University Press, 1999.

[71] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm.*, vol. 10, no. 6, pp. 585-596, Nov. 1999.

[72] P. M. Marques and S. A. Abrantes, "On the Derivation of the Exact, Closed-Form Capacity Formulas for Receiver-Sided Correlated MIMO Channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1139–1161, Mar. 2008.

[73] A. Abdi and M. Kaveh, "A space-time correlation model for multi-element antenna systems in mobile fading channels," *IEEE J. Select. Areas Commun.*, vol. 20, no. 3, pp. 550-560, Apr. 2002.

[74] V. A. Aalo, "Performance of maximal-ratio diversity systems in a correlated Nakagami-fading environment," *IEEE Trans. Commun.*, vol. 43, no. 8, pp. 2360–2369, Aug. 1995.

[75] A. Abdi and M. Kaveh, "A versatile spatio-temporal correlation function for mobile fading channels with nonisotropic scattering," in *Proc. IEEE Workshop Statistical Signal Array Processing, Pocono Manor*, PA, pp. 58–62, 2000.

[76] O. Popescu, C. Rose, and D. C. Popescu, "Maximizing the determinant for a special class of block-partitioned matrices," *Math. Probl. in Engineering*, vol. 2004, no. 1, pp. 49–61, May 2004.

[77] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. SP-50, no. 5, pp. 1065-1076, May 2002.

[78] K.-J. Lin and Y. Lee, "Adaptive Minimum Variance Detection for Spatial Multiplexing over Time-Varying MIMO Channels," *IEEE Vehicular Technology Conference*, vol., no., pp. 1360-1364, May 11-14, 2008.

[79] S. Kun and Z. Xudong, "Equalization of time-varying MIMO systems using modified Kalman filter algorithm," *Proc. of 7th International Conference on Signal Processing*, vol. 2, no., pp. 1670-1673 vol. 2, 31 Aug.-4 Sept. 2004.

[80] G. T. Zhou, Y. Kim, and G. B. Giannakis, "Estimation and Equalization of Time-Selective Channels using Precoding," *Proc. of 33rd Asilomar Conf. on Signals, Systems, and Computers*, vol. 1, pp. 248-252, Pacific Grove, CA, Oct. 24-27, 1999.

[81] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Channel estimation and equalization in fading," *Proc. of 33rd Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1159–1163, Pacific Grove, CA, Oct. 24-27, 1999.

[82] S. He, X. Shan, and Y. Ren, "A new adaptive equalization scheme for MIMO channels," *Proc. of the 3rd International Symp. on Image and Signal Processing and Analysis*, vol. 2, no., pp. 1028-1033, Sept. 18-20, 2003.

[83] P. W. Wolniansky, G. Foschini, G. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel" in *Proc. URSI International Symposium on Signals, Systems, and Electronics.*, vol., no., pp. 295-300, 29 Sep.-2 Oct., 1998.

[84] J. Choi and H. Y. Lee, "Adaptive MIMO decision feedback equalization for receivers in time-varying channels," *The 57th IEEE Semiannual Vehicular Technology Conference,* vol.3, no., pp. 1851-1856, April 22-25, 2003.

[85] D. K. C. So and R. S. Cheng, "Layered maximum likelihood detection for MIMO systems in frequency selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 752-762, April 2006.

[86] A. M. Sayeed and B. Aazhang, "Joint multipath-doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 123--132, Jan. 1999.

[87] S. Bhashyam, A. M. Sayeed, and B. Aazhang, "Time-Selective Signaling and Reception for Communication over Multipath Fading Channels," *IEEE Trans. Commun.*, vol. 48, no. 1, pp. 83-94, Jan. 2000.

[88] X. Ma and G. B. Giannakis, "Maximum-Diversity Transmissions over Doubly-Selective Wireless Channels." *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1832-1840, July 2003.

[89] R. Boudreau, J.-Y. Chouinard and A. Yongacoglu, "Exploiting Doppler-diversity in flat, fast fading channels," *Canadian Conference on Electrical and Computer Engineering*, vol. 1, no., pp. 270-274 vol. 1, 2000.

[90] B. D. Hart and D. P. Taylor, "Extended MLSE diversity receiver for the time- and frequency-selective channel," *IEEE Trans. on Commun.*, vol. 45, no. 3, pp. 322-333, Mar. 1997.

[91] H. Zhu, B. Farhang-Boroujeny, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication systems," *IEEE Commun. Lett.*, vol. 7, no. 1, pp. 30–32, Jan. 2003.

[92] J. Gao and H. Liu, "Decision-directed estimation of MIMO time-varying Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1412-1417, July 2005.

[93] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12-25, 2004.

[94] D. Samardzija and N. Mandayam, "Pilot assisted estimation of MIMO fading channel response and achievable data rates," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2882–2890, Nov. 2003.

[95] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science.* New York: Springer, 2000.

[96] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space.* New York: Dover, 1993.

[97] K. Jörgens, "Linear integral operators," *Surveys and Reference Works in Mathematics, 7.* Boston, MA: Pitman, 1982.

[98] H.-N. Lee and G. J. Pottie, "Fast adaptive equalization/diversity combining for time-varying dispersive channels," *IEEE Trans. Commun.*, vol. 46, no. 9, pp. 1146-1162, Sep. 1998.

[99] W. H. Chin, D. B. Ward, and A. G. Constantinides, "Algorithm for exploiting channel time selectivity in pilot-aided MIMO systems," *IET Communications*, vol. 1, no. 6, pp. 1267-1273, Dec. 2007.

[100] S. Barbarossa and A. Scaglione, "On the capacity of linear time-varying channels," *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Proc.*, vol. 5, pp. 2627-2630, March 15-19, 1999.

[101] G. Matz and F. Hlawatsch, "Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1067-1086, Mar. 2006.

[102] F. Hlawatsch, G. Matz, H. Kirchauer, and W. Kozek, "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation," *IEEE Trans. on Signal Process.*, vol. 48, no. 5, pp.1417-1432, May 2000.

[103] N. Wiener and P. Masani, "The prediction theory of multivariate stochastic processes, I - The regulariy condition," *Acta Math.*, 98, pp. 111–150, 1957.

[104] N. Wiener and P. Masani, "The prediction theory of multivariate stochastic processes, II - The linear predictor," *Acta Math.*, 99, pp. 93–137, 1958.

[105] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[106] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: John Wiley & Sons, 2001.

[107] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[108] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice-Hall, 1998.