1 Introduction

1.1 Motivation

We initiate the book with a classical example, which exhibits both the non-obvious behavior of large dimensional random matrices and the motivation behind their study.

Consider a random sequence $x_1, \ldots, x_n$ of n independent and identically distributed (i.i.d.) *observations* of a given random process. The classical law of large numbers states that the sequence $x_1, \frac{x_1+x_2}{2}, \ldots$, with nth term $\frac{1}{n}\sum_{k=1}^{n} x_k$ tends almost surely to the deterministic value $\mathrm{E}[x_1]$, the expectation of this process, as n tends to infinity. Denote $(\Omega, \mathcal{F}, P)$ the probability space that generates the infinite sequence $x_1, x_2, \ldots$. For a given realization $\omega \in \Omega$, we will denote $x_1(\omega), x_2(\omega), \ldots$ the realization of the random sequence $x_1, x_2, \ldots$. We recall that almost sure convergence means that there exists $A \subset \Omega$, with $P(A) = 1$, such that, for $\omega \in A$

$$\frac{1}{n}\sum_{k=1}^{n} x_k(\omega) \to \mathrm{E}[x_1] \triangleq \int_{\Omega} x_1(w)dw.$$

We also remind briefly that the notation $(\Omega, \mathcal{F}, P)$ designates the triplet composed of the space of random realizations Ω, i.e. in our case $\omega \in \Omega$ is the realization of a series $x_1(\omega), x_2(\omega), \ldots$, $\mathcal{F}$ is a σ-field on Ω, which can be seen as the space of the measurable *events* on Ω, e.g. the space $B = \{x_1(\omega) > 0\} \in \mathcal{F}$ is such an event, and P is a probability measure on $\mathcal{F}$, i.e. P is a function that assigns to every event in $\mathcal{F}$ a probability.

This law of large numbers is fundamental in the sense that it provides a deterministic feature for a process ruled by 'chance' (or more precisely, ruled by a deterministic process, the precise nature of which the observer is unaware). This allows the observer to be able to retrieve deterministic information from random variables based on any observed random sequence (within a space of probability one). If, for instance, $x_1, \ldots, x_n$ are successive samples of a stationary zero mean white noise waveform $x(t)$, i.e. $\mathrm{E}[x(t)x(t-\tau)] = \sigma^2\delta(t)$, it is the usual signal processing problem to estimate the power $\sigma^2 = \mathrm{E}[|x_1|^2]$ of the noise process; the

empirical variance σ_n^2, i.e.

$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n} |x_i|^2$$

is a classical estimate of σ^2 which, according to the law of large numbers, is such that $\sigma_n^2 \to \sigma^2$ almost surely when $n \to \infty$. It is often said that σ_n^2 is a *consistent* estimator of σ^2 as it is asymptotically and almost surely equal to σ^2. To avoid confusion with the two-dimensional case treated next, we will say instead that σ_n^2 is an *n-consistent* estimator of σ^2, as it is asymptotically accurate as n grows large. Obviously, we are never provided with an infinitely long observation time window, so that n is usually large but finite, and therefore σ_n^2 is merely an approximation of σ^2.

With the emergence of multiple antenna systems, channel spreading codes, sensor networks, etc., signal processing problems have become more and more concerned with vectorial inputs rather than scalar inputs. For a sequence of n i.i.d. random vectors, the law of large numbers still applies. For instance, for $\mathbf{x}_1, \mathbf{x}_2, \ldots \in \mathbb{C}^N$ randomly drawn from a given N-variate zero mean random process

$$\mathbf{R}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^{\mathsf{H}} \to \mathbf{R} \triangleq \mathrm{E}[\mathbf{x}_1 \mathbf{x}_1^{\mathsf{H}}] \tag{1.1}$$

almost surely as $n \to \infty$, where the convergence is considered for any matrix norm, i.e. $\|\mathbf{R} - \mathbf{R}_n\| \to 0$ on a set of probability one. The matrix $\mathbf{R}_n$ is often referred to as the *empirical covariance matrix* or as the *sample covariance matrix*, as it is computed from observed vector samples. We will use this last phrase throughout the book. Following the same semantic field, the matrix $\mathbf{R}$ will be referred to as the *population covariance matrix*, as it characterizes the innate nature of all stochastic vectors $\mathbf{x}_i$ from the overall population of such vectors. The empirical $\mathbf{R}_n$ is again an n-consistent estimator of $\mathbf{R}$ of $\mathbf{x}_1$ and, as before, as n is taken very large *for N fixed*, $\mathbf{R}_n$ is a good approximation of $\mathbf{R}$ in the sense of the aforementioned matrix norm. However, in practical applications, it might be that the number of available snapshots $\mathbf{x}_k$ is indeed very large but not extremely large compared to the vector size N. This situation arises in diverse application fields, such as biology, finance, and, of course, wireless communications. If this is the case, as will become obvious in the following examples and against intuition, the difference $\|\mathbf{R} - \mathbf{R}_n\|$ can be far from zero even for large n.

Since the DNA of many organisms have now been entirely sequenced, biologists and evolutionary biologists are interested in the correlations between genes, e.g.: How does the presence of a given gene (or gene sequence) in an organism impact the probability of the presence of another given gene? Does the activation of a given gene come along with the activation of several other genes? To be able to study the joint correlation between a large population of the several ten thousands of human genes, call this number N, we need a large sample of genome

sequences extracted from human beings, call the number of such samples n. It is therefore typical that the $N \times n$ matrix of the n gene sequence samples does not have many more columns than rows, or, worse, may even have more rows than columns. We see already that, in this case, the sample covariance matrix $\mathbf{R}_n$ is necessarily rank-deficient (of maximum rank $N - n$), while $\mathbf{R}$ has all the chances to be full rank. Therefore, $\mathbf{R}_n$ is obviously no longer a good approximation of $\mathbf{R}$, even if n were very large in the first place, since the eigenvalues of $\mathbf{R}_n$ and $\mathbf{R}$ differ by at least $N - n$ terms.

In the field of finance, the interest of statisticians lies in the interactions between assets in the market and the joint time evolution of their stock market indices. The vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ here may be representative of n months of market index evolution of N different brands of a given product, say soda, the ith entry of the column vector $\mathbf{x}_k$ being the evolution of the market index of soda i in month k. Obviously, this case differs from the independent vector case presented up to now since the evolution at month $k + 1$ is somewhat correlated to the evolution at previous month k, but let us assume for simplicity that the month evolution is at least an uncorrelated process (which does not imply independence). Similar to the gene case for biologists, it often turns out that the $N \times n$ matrix under study contains few columns compared to the number of rows, although both dimensions are typically large compared to 1. Of specific importance to traders is the largest eigenvalue of the population covariance matrix $\mathbf{R}$ of (a centered and normalized version of) the random process $\mathbf{x}_1$, which is an indicator of the maximal risk against investment returns taken by a trader who constitutes a portfolio from these assets. From the biology example above, it has become clear that the eigenvalues of $\mathbf{R}_n$ may be a very inaccurate estimate of those of $\mathbf{R}$; thus, $\mathbf{R}_n$ cannot be relied on to estimate the largest eigenvalue and hence the trading risk. The case of wireless communications will be thoroughly detailed in Part II, and a first motivation is given in the next paragraph.

Returning to the initial sample covariance matrix model, we have already mentioned that in the scalar case the strong law of large numbers ensures that it suffices for n to be quite large compared to 1 for σ_n^2 to be a good estimator for σ^2. In the case where data samples are vectors, if n is large compared to 1, whatever N, then the (i, j) entry $R_{n,ij}$ of $\mathbf{R}_n$ is a good estimator of the (i, j) entry R_{ij} of $\mathbf{R}$. This might (mis)lead us to assume that as n is much greater than one, $\mathbf{R}_n \simeq \mathbf{R}$ in some sense. However, if both N and n are large compared to 1 but n is not large compared to N, then the peculiar thing happens: the eigenvalue distribution of $\mathbf{R}_n$ (see this as an histogram of the eigenvalues) in general converges, but does not converge to the eigenvalue distribution of $\mathbf{R}$. This has already been pointed out in the degenerated case $N > n$, for which $\mathbf{R}_n$ has $N - n$ null eigenvalues, while $\mathbf{R}$ could be of full rank. This behavior is evidenced in Figure 1.1 in which we consider $\mathbf{x}_1 \sim \mathcal{CN}(0, \mathbf{I}_N)$ and then $\mathbf{R} = \mathbf{I}_N$, for $N = 500$, $n = 2000$. In that case, notice that $\mathbf{R}_n$ converges *point-wise* to $\mathbf{I}_N$

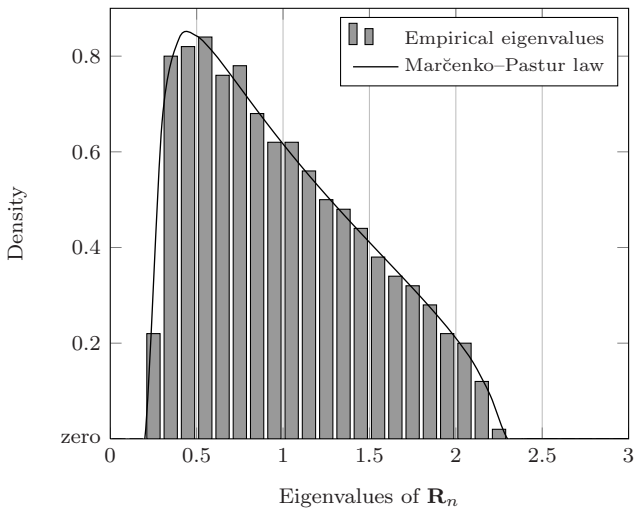

Figure 1.1 Histogram of the eigenvalues of $\mathbf{R}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_k \mathbf{x}_k^{\mathsf{H}}$, $\mathbf{x}_k \in \mathbb{C}^N$, for $n = 2000$, $N = 500$.

when n is large, as $R_{n,ij}$, the entry (i, j) of $\mathbf{R}_n$, is given by:

$$R_{n,ij} = \frac{1}{n}\sum_{k=1}^{n} x_{ik} x_{jk}^*$$

which is close to one if $i = j$ and close to zero if $i \neq j$. This is obviously irrespective of N, which is not involved in the calculus here. However, the eigenvalues of $\mathbf{R}_n$ do not converge to a single mass in 1 but are spread around 1. This apparent contradiction is due to the fact that N grows along with n but n/N is never large. We say in that case that, while $\mathbf{R}_n$ is an n-consistent estimator of $\mathbf{R}$, it is not an (n, N)-*consistent* estimator of $\mathbf{R}$. The seemingly paradoxical behavior of the eigenvalues of $\mathbf{R}_n$, while $\mathbf{R}_n$ converges *point-wise* to $\mathbf{I}_N$, lies in fact in the rate convergence of the entries of $\mathbf{R}_n$ towards the entries of $\mathbf{I}_N$. Due to central limit arguments for the sample mean of scalar i.i.d. random variables, $R_{n,ij} - \mathrm{E}[R_{n,ij}]$ is of order $O(1/\sqrt{n})$. When determining the eigenvalues of $\mathbf{R}_n$, the deviations around the means are negligible when n is large and N fixed. However, for N and n both large, these residual deviations of the entries of $\mathbf{R}_n$ (their number is N^2) are no longer negligible and the eigenvalue distribution of $\mathbf{R}_n$ is not a single mass in 1. In some sense, we can see $\mathbf{R}_n$ as a matrix close to the identity but whose entries all contain some small residual "energy," which becomes relevant as much of such small energy is cumulated.

This observation has very important consequences, which motivate the need for singling out the study of large empirical covariance matrices and more generally of large random Hermitian matrices as a unique field of mathematics. Wireless communications may be the one research field in which large matrices have started to play a fundamental role. Indeed, current and more importantly

future wireless communication systems are multi-dimensional in several respects (spatial with antennas, temporal with random codes, cellular-wise with large number of users, multiple cooperative network nodes, etc.) and random in other respects (time-varying fading channels, noisy communications, etc.). The study of the behavior of large wireless communication systems therefore calls for advanced mathematical tools that can easily deal with large dimensional random matrices. Consider for instance a multiple input multiple output (MIMO) complex channel matrix $\mathbf{H} \in \mathbb{C}^{N \times n}$ between an n-antenna transmitter and an N-antenna receiver, the entries of which are independent and complex Gaussian with zero mean and variance $1/n$. If uniform power allocation across the antennas is used at the transmit antenna array and the additive channel noise is white and Gaussian, the achievable transmission rates over this channel are all rates less than the channel mutual information

$$\mathcal{I}(\sigma^2) = \mathrm{E}\left[\log_2 \det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)\right] \tag{1.2}$$

where σ^{-2} denotes now the signal-to-noise ratio (SNR) at the receiver and the expectation is taken over the realizations of the random channel $\mathbf{H}$, varying according to the Gaussian distribution. Now note that $\mathbf{H}\mathbf{H}^{\mathsf{H}} = \sum_{i=1}^{n} \mathbf{h}_i \mathbf{h}_i^{\mathsf{H}}$ with $\mathbf{h}_i \in \mathbb{C}^N$ the ith column of $\mathbf{H}$, $\mathbf{h}_1, \ldots, \mathbf{h}_n$ being i.i.d. random vectors. The matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ can then be seen as the sample covariance matrix of some hypothetical random N-variate variable $\sqrt{n}\mathbf{h}_1$. From our previous discussion, denoting $\mathbf{H}\mathbf{H}^{\mathsf{H}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$ the spectral decomposition of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, we have:

$$\mathcal{I}(\sigma^2) = \mathrm{E}\left[\log_2 \det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{\Lambda}\right)\right] = \mathrm{E}\left[\sum_{i=1}^{N} \log_2\left(1 + \frac{\lambda_i}{\sigma^2}\right)\right] \tag{1.3}$$

with $\lambda_1, \ldots, \lambda_N$ the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, which again are not all close to one, even for n and N large. The achievable transmission rates are then explicitly dependent on the eigenvalue distribution of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$. More generally, it will be shown in Chapters 12–15 that random matrix theory provides a powerful framework, with multiple methods, to analyze the achievable transmission rates and rate regions of a large range of multi-dimensional setups (MIMO, CDMA, multi-user transmissions, MAC/BC channels, etc.) and to derive the capacity-achieving signal covariance matrices for some of these systems, i.e. determine the non-negative definite matrix $\mathbf{P} \in \mathbb{C}^{N \times N}$, which, under some trace constraint $\mathrm{tr}\,\mathbf{P} \leq P$, maximizes the expression

$$\mathcal{I}(\sigma^2; \mathbf{P}) = \mathrm{E}\left[\log \det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{H}\mathbf{P}\mathbf{H}^{\mathsf{H}}\right)\right]$$

for numerous fading channel models for $\mathbf{H}$.

1.2 History and book outline

The present book is divided into two parts: a first part on the theoretical fundamentals of random matrix theory, and a second part on the applications of random matrix theory to the field of wireless communications. The first part will give a rather broad, although not exhaustive, overview of fundamental and recent results concerning random matrices. However, the main purpose of this part goes beyond a listing of important theorems. Instead, it aims on the one hand at providing the reader with a large, yet incomplete, range of techniques to handle problems dealing with random matrices, and on the other hand at developing sketches of proofs of the most important results in order to provide further intuition to the reader. Part II will be more practical as it will apply most of the results derived in Part I to problems in wireless communications, such as system performance analysis, signal sensing, parameter estimation, receiver design, channel modeling, etc. Every application will be commented on with regard to the theoretical results developed in Part I, for the reader to have a clear understanding of the reasons why the practical results hold, of their main limitations, and of the questions left open. Before moving on to Part I, in the following we introduce in detail the objectives of both parts through a brief historical account of eighty years of random matrix theory.

The origin of the study of random matrices is usually said to date back to 1928 with the pioneering work of the statistician John Wishart [Wishart, 1928]. Wishart was interested in the behavior of sample covariance matrices of i.i.d. random vector processes $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{C}^N$, in the form of the matrix $\mathbf{R}_n$ previously introduced

$$\mathbf{R}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^{\mathsf{H}}. \tag{1.4}$$

Wishart provided an expression of the joint probability distribution of the entries of such a matrix when its column vector entries are themselves independent and have an identical standard complex Gaussian distribution, i.e. $x_{ij} \sim \mathcal{CN}(0, 1)$. These normalized matrices with i.i.d. standard Gaussian entries are now called *Wishart matrices*. Wishart matrices were thereafter generalized and extensively studied. Today there exists in fact a large pool of properties on the joint distribution of the eigenvalues, the distribution of the extremes eigenvalues, the distribution of the ratios between extreme eigenvalues, etc.

The first asymptotic considerations, i.e. the first results on matrices of asymptotically large dimensions, appeared with the work of the physician Eugene Wigner [Wigner, 1955] on nuclear physics, who considered (properly scaled) symmetric matrices with independent entries uniformly distributed in $\{1, -1\}$ and proved the convergence of the marginal probability distribution of its eigenvalues towards the deterministic *semi-circle law*, as the dimension of the matrix grows to infinity. Hermitian $n \times n$ matrices with independent upper-

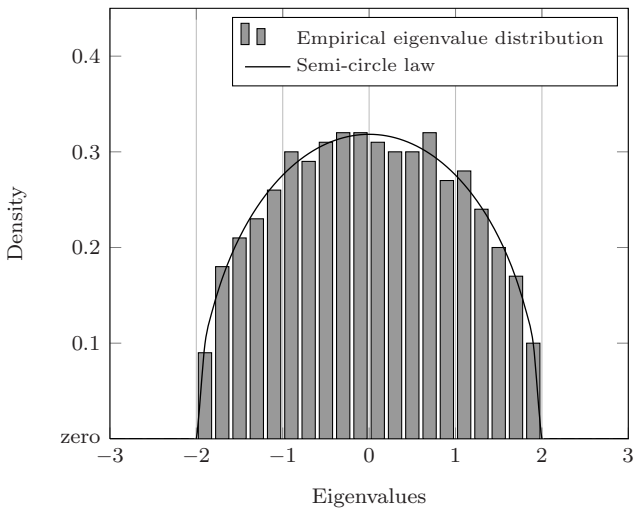

Figure 1.2 Histogram of the eigenvalues of a Wigner matrix and the semi-circle law, for $n = 500$.

triangular entries of zero mean and variance $1/n$ are now referred to as *Wigner matrices*. The empirical eigenvalues of a large Wigner matrix and the semi-circle law are illustrated in Figure 1.2, for a matrix of size $n = 500$. From this time on, infinite size random matrices have drawn increasing attention in many domains of physics [Mehta, 2004] (nuclear physics [Dyson, 1962a], statistical mechanics, etc.), finance [Laloux *et al.*, 2000], evolutionary biology [Arnold *et al.*, 1994], etc. The first accounts of work on large dimensional random matrices for wireless communications are attributed to Tse and Hanly [Tse and Hanly, 1999] on the performance of large multi-user linear receivers, Verdú and Shamai [Verdú and Shamai, 1999] on the capacity of code division multiple access (CDMA) systems, among others. The pioneering work of Telatar [Telatar, 1995] on the transmission rates achievable with multiple antennas, paralleled by Foschini [Foschini and Gans, 1998], is on the contrary a particular example of the use of small dimensional random matrices for capacity considerations. In its final version of 1999, the article also mentions asymptotic laws for capacity [Telatar, 1999]. We will see in Chapter 13 that, while Telatar's original proof of the capacity growth rate for increasing number of antennas in a multiple antenna setup is somewhat painstaking, large random matrix theory provides a straightforward result. In Chapter 2, we will explore some of the aforementioned results on random matrices of small dimension, which will be shown to be difficult to manipulate for simply structured matrices and rather intractable to extend to more structured matrices.

The methods used for random matrix-based calculus are mainly segmented into: (i) the analytical methods, which treat asymptotic eigenvalue distributions of large matrices in a comprehensive framework of analytical tools, among

which the important *Stieltjes transform*, and (ii) the moment-based methods, which establish results on the successive moments of the asymptotic eigenvalues probability distribution.[1] The analytical framework allows us to solve a large range of problems in wireless communications such as those related to capacity evaluation in both random and orthogonal CDMA networks and in large MIMO systems, but also to address questions such as signal sensing in large networks or statistical inference, i.e. estimation of network parameters. These analytic methods are mostly used when the random matrices under consideration are sample covariance matrices, doubly correlated i.i.d. matrices, information plus noise matrices (to be defined later), isometric matrices or the sum and products of such matrices. They are generally preferred over the alternative moment-based methods since they consider the eigenvalue distribution of large dimensional random matrices as the central object of study, while the moment approach is dedicated to the specific study of the successive moments of the distribution. Note in particular that not all distributions have moments of all orders, and for those that do have moments of all orders, not all are uniquely defined by the series of their moments. However, in some cases of very structured matrices whose entries are non-trivially correlated, as in the example of Vandermonde matrices [Ryan and Debbah, 2009], the moment-based methods convey a more accessible treatment. Both analytical and moment-based methods are not completely disconnected from one another as they share a common denominator when it comes to dealing with unitarily invariant random matrices, such as standard Gaussian or Haar matrices, i.e. unitarily invariant unitary matrices. This common denominator, namely the field of *free probability theory*, bridges the analytical tools to the moment-based methods via derivatives of the Stieltjes transform, the R-transform, and the S-transform. The latter can be expressed in power series with coefficients intimately linked to moments and cumulants of the underlying random matrix eigenvalue distributions. The free probability tool, due to Voiculescu [Voiculescu *et al.*, 1992], was not initially meant to deal specifically with random matrices but with more abstract non-commutative algebras, large dimensional random matrices being a particular case of such algebras. The extension of classical probability theory to free probability provides interesting and often surprising results, such as a strong equivalence between some classical probability distributions, e.g. Poisson, Gaussian, and the asymptotic probability distribution of the eigenvalues of some random matrix models, e.g. Wishart matrices and Wigner matrices. Some classical probability tools, such as the characteristic function, are also extensible through analytic tools of random matrix theory.

[1] Since the terminology *method of moments* is already dedicated to the specific technique which aims at constructing a distribution function from its moments (under the condition that the moments uniquely determine the distribution), see, e.g. Section 30 of [Billingsley, 1995], we will carefully avoid referring to any random matrix technique based on moments as the method of moments.

The division between analytical and moment-based methods, with free probability theory lying in between, can be seen from another point of view. It will turn out that the analytical methods, and most particularly the Stieltjes transform approach, take full advantage of the independence between the entries of large dimensional random matrices. As for moment-based methods, from a free probability point of view, they take full advantage of the invariance properties of large dimensional matrices, such as the invariance by the left or right product with unitary matrices. The theory of orthogonal polynomials follows the same pattern, as it benefits from the fact that the eigenvalue distribution of unitarily invariant random matrices can be studied regardless of the (uniform) eigenvector distribution. In this book, we will see that the Stieltjes transform approach can solve most problems involving random matrices with invariance properties as well. This makes this distinction between random matrices with independent entries and random matrices with invariance properties not so obvious to us. For this reason, we will keep distinguishing between the analytical approaches that deal with the eigenvalue distribution as the central object of concern and the moment-based approaches that are only concerned with successive moments. We will also briefly introduce the rather old theory of orthogonal polynomials which has received much interest lately regarding the study of limiting laws of largest eigenvalues of random matrices but which requires significant additional mathematical effort for proper usage, while applications to wireless communications are to this day rather limited, although in constant expansion. We will therefore mostly state the important results from this field, particularly in terms of limit theorems of extreme eigenvalues, see Chapter 9, without development of the corresponding proofs.

In Chapter 3, Chapter 4, and Chapter 5, we will introduce the analytical and moment-based methods, as well as notions of free probability theory, which are fundamental to understand the important concept of asymptotic freeness for random matrices. We will also provide in these chapters a sketch of the proof of the convergence of the eigenvalue distribution of the Wishart and Wigner matrices to the Marčenko–Pastur law, depicted in Figure 1.1, and the semi-circle law, depicted in Figure 1.2, using the Stieltjes transform and the method of moments, respectively. Generic methods to determine (almost sure) limiting distributions of the eigenvalues of large dimensional random matrices, as well as other functionals of such large matrices (e.g. log determinant), will be reviewed in detail in these chapters. Chapter 6 will discuss the alternative methods used when the empirical eigenvalue distribution of large random matrices do not necessarily converge when the dimensions increase: in that case, in place of limit distributions, we will introduce the so-called *deterministic equivalents*, which provide deterministic approximations of functionals of random matrices of finite size. These approximations are (almost surely) asymptotically accurate as the matrix dimensions grow to infinity, making them consistent with the methods developed in Chapter 3. In addition to limiting eigenvalue distributions and deterministic equivalents, in Chapter 3 and Chapter 6, central

limit theorems that extend the convergence theorems to a higher precision order will be introduced. These central limit theorems constitute a first step into a more thorough analysis of the asymptotic deviations of the spectrum around its almost sure limit or around its deterministic equivalent. Chapter 7 will discuss advanced results on the spectrum of both the sample covariance matrix model and the information plus noise model, which have been extensively studied and for which many results have been provided in the literature, such as the proof of the asymptotic absence of eigenvalues outside the support of the limiting distribution. Beyond the purely mathematical convenience of such a result, being able to characterize where the eigenvalues, and especially the extreme eigenvalues, are expected to lie is of fundamental importance to perform hypothesis testing decisions and in statistical inference. In particular, the characterization of the spectrum of sample covariance matrices will be used to retrieve information on functionals of the population covariance matrix from the observed sample covariance matrix, or functionals of the signal space matrix from the observed information plus noise matrix. Such methods will be referred to as *eigen-inference techniques* and are developed in Chapter 8. The first part will then conclude with Chapter 9, which extends the analysis of Section 7.1 to the expression of the limiting distributions of the extreme eigenvalues. We will also introduce in this chapter the *spiked models*, which have recently received a lot of attention for their many practical implications. These objects are necessary tools for signal sensing in large dimensional networks, which are currently of major interest with regard to the recent incentive for cognitive radios. In Chapter 10, the essential results of Part I will finally be summarized and rediscussed with respect to their applications to the field of wireless communications.

The second part of this book is dedicated to the application of the different methods described in the first chapter to different problems in wireless communications. As already mentioned, the first applications of random matrix theory to wireless communications are exclusively related to asymptotic system performance analysis, and especially channel capacity considerations. The idea of considering asymptotically large matrix approximations was initially linked to studies in CDMA communications, where both the number of users and the length of the spreading codes are potentially very large [Li *et al.*, 2004; Tse and Hanly, 1999; Tse and Verdú, 2000; Tse and Zeitouni, 2000; Zaidel *et al.*, 2001]. It then occurred to researchers that large matrix approximations work rather well when the size of the effective matrix under study is not so large, e.g. for matrices of size 8×8 or even 4×4 (in the case of random unitary matrices, simulations suggest that approximations for matrices of size 2×2 are even acceptable). This motivated further studies in systems where the number of relevant parameters is moderately large. In particular, studies of MIMO communications [Chuah *et al.*, 2002; Hachem *et al.*, 2008b; Mestre *et al.*, 2003; Moustakas and Simon, 2005; Müller, 2002], designs of multi-user receivers [Honig and Xiao, 2001; Müller and Verdú, 2001], multi-cell communications [Abdallah and Debbah, 2004; Couillet *et al.*, 2011a; Peacock *et al.*, 2008], multiple access channels and