Pick a flower on Earth and you move the farthest star.

– Paul Dirac

University of Alberta

Performance Analysis and Array Design for Size Constrained Multiple Antenna Reception

by

Payam Dehghani Rahimzadeh

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Communications

Department of Electrical and Computer Engineering

Dedicated to

my family.

Abstract

Future generations of wireless cellular networks will vastly benefit from the various advantages of multiple antenna techniques. It is, however, well known that signal correlation among a closely spaced set of antennas deployed on a small wireless mobile device can dramatically degrade the performance of such multiple antenna techniques. The performances of two multiple antenna reception schemes for a size limited array of antennas are investigated. Exact closed-form expressions are derived for the bit error rate of binary phase shift keying (BPSK) in Rayleigh fading with maximal ratio combining (MRC) diversity in the presence of cochannel interference (CCI) and additive white Gaussian noise (AWGN). The desired signal and the interferer signals are all subject to correlated Rayleigh fading. In the next contribution, an analytical expression is derived for the average output signal-to-interference-plus-noise ratio (SINR) of optimum combining (OC) for a spatially correlated array of antennas in the presence of a single interferer and Rayleigh fading. Using the derived expression and based on an asymptotic analysis of the eigenvalues of dense correlation matrices, the asymptotic performance of optimum combining is evaluated as the number of the antennas increases while the total physical size of the array is fixed. The case of multiple interferers is examined by simulation and is shown to exhibit similar asymptotic behavior to the case of one interferer. Finally, a general analytical framework for the optimal design of a size constrained array of antennas is developed. It is shown that the problem of optimizing the number and the positions of the antennas within a size limited array can be formulated as a quadratic convex optimization problem which can be solved efficiently using the available numerical methods for convex optimization. Moreover, an analytical solution to this convex problem is

obtained for the special case of a linear array under exponential correlation model. Several one-dimensional (1D) and two-dimensional (2D) array design examples are presented within the proposed framework.

Acknowledgements

I am indebted to many people without whose support this thesis would not have been possible. First of all, I would like to express my deepest gratitude to my supervisor, Dr. Norman C. Beaulieu for his continued help and support during the course of my Ph.D. program. I would also like to thank the members of my examining committee for their careful reading of the thesis and for their valuable comments and suggestions.

My special thanks goes to my friends and colleagues in *i*CORE Wireless Communications Laboratory (*i*WCL) and many other friends in fifth floor, ECERF building where I spent most of my time during my Ph.D. years. Their friendship has always been a source of joy, motivation, and encouragement. In particular, I would like to thank Reza Nikjah, Kaveh Mahdaviani, and Mahmood Mazrouei from whom I have learned a lot through our fruitful and sometimes very long discussions, David young for his devotion in keeping our computers up and running, and Amirmasoud Rabiei and Moslem Noori for organizing the life saving afternoon coffee breaks.

Last, but not least, I am grateful to my beloved fiancée, Hinako Ishikawa, for her patience, support, and encouragement during all those years of hard work, frustration, and excitement. I would not have been able to complete this thesis had I not had her by my side.

Table of Contents

List of Figures

List of Symbols

$(\cdot)^*$	The complex conjugate	31
$(\cdot)^H$	The conjugate transpose (Hermitian)	10
$(\cdot)^T$	The matrix transpose	29
$\langle \cdot, \cdot \rangle$	Inner product	93
$\lfloor \cdot \rfloor$	The floor operator	21
$\lvert \cdot \rvert$	The absolute value	15
$\lvert \cdot \rvert$	The matrix determinant	76
$\star$	The convolution operator	11
$\mathbf{1}_n$	An $n \times 1$ vector whose elements are all equal to 1	57
$\mathbf{A}$	Diagonal matrix containing the eigenvalues of $\mathbf{R_a}/M$	28
$\alpha_{i,l}$	Coefficients of the partial fraction expansion of $\mathcal{A}(s)$ in (2.6b)	15
$\alpha_{i,M}$	The i'th eigenvalue of $\mathbf{R_a}/M$	28
$\tilde{\alpha}_k$	The k'th largest eigenvalue of an integral operator with kernel $R_a(x, y)$	32
β	A parameter defined in (3.23b)	34
β	Roll-off factor of a root raised-cosine pulse	21
$\mathbf{c}(s, t)$	A mapping from a subset of $\mathbb{R}^2$ to a subset of $\mathbb{R}^2$	59
$\mathbf{c}(t)$	A mapping from a subset of $\mathbb{R}$ to a subset of $\mathbb{R}^3$	31
$\mathcal{C}([0, 1])$	The set of continuously differentiable functions over $[0, 1]$	55
$c(t)$	A mapping from a subset of $\mathbb{R}$ to a subset of $\mathbb{R}$	54
$\mathrm{Cov}(\cdot, \cdot)$	Covariance operator	57
$\delta(\cdot)$	Dirac's delta function	57
δ_k	Kronecker's delta function	12
$\nabla \mathcal{J}(p)$	Variational derivative of the functional $\mathcal{J}(p)$	93

$P_{out}^{\gamma_{th}}$	The SINR outage probability of OC at threshold level γ_{th}	76
$\mathbf{Q}$	Matrix of the eigenvectors of $\mathbf{R_a}$	28
$\mathbf{Q_0}$	Matrix of the eigenvectors of $\mathbf{R_0}$	10
$\mathbf{Q_I}$	Matrix of the eigenvectors of $\mathbf{R_I}$	10
$\mathbf{q}_i$	The eigenvector corresponding to the i'th eigenvalue of $\mathbf{R_a}$	28
$q_k(\cdot)$	The k'th eigenfunction of an integral operator with kernel $R_a(x, y)$	32
$\mathbf{R}$	the short-term correlation matrix of the interference-plus-noise	28
$\mathbf{r}$	Position vector	31
$\mathbb{R}$	The set of real numbers	46
r	Radius of a circular array of antennas	21
$\mathbf{R_0}$	Covariance (correlation) matrix of the desired user's channel vector	10
$\mathbf{R_a}$	Array correlation matrix	27
$R_a(\cdot, \cdot)$	Array correlation function	31
$\mathrm{Res}(s_i)$	Complex residue of $\mathcal{A}(s)$ in (2.6b) corresponding to the pole s_i	14
$R_g(x)$	1D spatial correlation function	46
$R_g(x, y)$	2D spatial correlation function	46
$R_g(\mathbf{r_1}, \mathbf{r_2})$	Spatial correlation function	31
$R_g(\mathbf{r}_1 - \mathbf{r}_2)$	Spatial correlation function for WSS random fields	46
$\mathcal{R}_g(\omega)$	Fourier transform of $R_g(x)$	57
$\mathbf{R_I}$	Covariance (correlation) matrix of the interfering users' channel vectors	10
$s(t)$	Users' pulse shape	10
$\mathrm{sinc}(\cdot)$	The sinc function, $\mathrm{sinc}(x) = \sin(x)/x$	20
τ_i	Time delay of the i'th user's signal relative to the desired user's signal	10
T	Symbol interval	10
$\mathrm{tr}(\cdot)$	The trace operator	28
z_1	The first positive zero of $\mathrm{J}_0(2\pi x)$	78

List of Abbreviations

Chapter 1

Introduction

The focus of this thesis is on the performance analysis and array design for multiple antenna diversity receivers in the presence of multipath fading, antenna correlation, and cochannel interference (CCI). This research is motivated by the fact that wireless mobile devices, such as cell phones, laptops, etc. can vastly benefit from the various advantages of multiple antenna reception techniques. A multiple antenna receiver on a mobile device can be designed to compensate for the signal fading arising from the multipath propagation of the electromagnetic waves. A multiple antenna receiver can also be used to cancel out the undesired CCI arising from the existence of other users in the same network. However, it is well known that the fading correlation among the antennas due to the physical proximity of the antenna elements on a small receiver can severely deteriorate the predicted performance of a multiple antenna receiver. Therefore, it is important to study the effect of the antenna correlation on the performance of such receivers and to find a way to minimize its undesired effects. The contributions of this thesis can be divided into two major parts. The first part of the thesis focuses on the analytical study of two diversity reception techniques, namely, maximal ratio combining (MRC) and optimum combining (OC) in the presence of CCI and fading correlation. The second part is dedicated to the study of the optimal arrangement of the antennas within a small size device in order to minimize the unwanted effects of fading correlation. An introductory background on the subject as well as a review of the available literature is presented in the following sections before we get into the technical details of our research work.

1.1 Cellular Wireless Communication Networks

In a cellular mobile communication network, the geographic area under coverage is divided into hexagonal cells where the users inside a cell are served by the cell's base station(s). The base stations of different cells are connected to the outside world through the backhaul network. Using a multiple access (MA) technique [1] such as time division multiple access (TDMA) or code division multiple access (CDMA), the users within a cell can communicate with the cell's base station simultaneously without interfering with each other. This is done by assigning orthogonal signal signatures to different users within a cell. In TDMA, the signal orthogonality is implemented by assigning specific time slots to each user within which the other users remain silent and, in CDMA, it is done by assigning orthogonal codes to different users. The neighbouring cells use the same set of signal signatures for their users while a different carrier frequency is used by them to avoid intercell interference. However, since the radio frequency (RF) spectrum is scarce and expensive, the carrier frequency used in a cell is reused in other nonadjacent cells according to a specific pattern. This, in turn, produces a certain amount of interference known as cochannel interference (CCI). The CCI in addition to the inherent thermal noise of the receiver's electronics, modelled as an additive white Gaussian noise (AWGN), are two sources of performance degradation in cellular networks. Another source of performance degradation in a wireless channel is the so-called multipath fading. It is the result of the multiple reflections and refractions of the electromagnetic waves from physical objects such as buildings and trees in the case of an outdoor environment, or walls and furniture in the case of an indoor channel. Multipath fading refers to the fading of the received signal at certain points in space and time caused by the destructive superposition of the different copies of the signal received from multiple paths. The amount of signal fading is random due to the dynamic nature of the underlying physical channel and is commonly modelled as a Rayleigh, Ricean, or Nakagami random variable [2]. Multipath fading can severely degrade the system's performance. Therefore, finding ways of compensating its effect is crucial in the design of wireless communication systems.

1.2 Multiple Antenna Receivers and Array Processing

A well known technique to alleviate the undesirable effects of both CCI and multipath fading is the use of multiple antennas in the receiver along with a linear array processing (i.e., antenna combining) scheme. Replacing the single antenna in the receiver with an array of antennas provides extra degrees of freedom through which the interfering signals can be detected and cancelled out. It can also combat multipath fading by providing spatial diversity. Spatial diversity is based on the idea that it is unlikely for antennas located at different points in space to be in a deep fade all at the same time. An optimum linear array processing method in the presence of CCI, AWGN, and fading is the so-called optimum combining (OC) [3], also known as minimum mean square error (MMSE) combining. Essentially, OC maximizes the output signal-to-interference-plus-noise ratio (SINR) within the class of all linear combiners and is, in fact, a maximum-likelihood (ML) detector of the desired user's data when the channel state information (CSI) of the interfering users are unknown and under the assumption that the data transmitted by the interfering users are Gaussian. In practice, the data symbols of the interfering users may come from a discrete constellation, e.g., quadrature amplitude modulation (QAM), where the Gaussian assumption on the distribution of the data symbols is not valid. A truly optimum combining receiver, in this case, is obtained by applying the ML rule according to the actual distribution of the interferers' data symbols. However, the performance of this complex ML combiner is only slightly better than the conventional OC receiver which assumes Gaussian interference [1].

[1] This has been confirmed by computer simulation.

In order to implement OC, the receiver needs knowledge of the desired user's CSI as well as the short-term covariance matrix of the interference-plus-noise term. The CSI of the desired user can be estimated using pilot symbols whereas the covariance matrix of the interference-plus-noise term can be estimated using a blind estimation technique [4]. Although OC is the optimum linear processor, the suboptimum MRC scheme may be used as a simpler alternative. Maximal ratio combining is the optimum ML receiver in the presence of fading and AWGN, but not CCI. It provides the maximum output signal-to-noise ratio (SNR) among all linear combiners. To implement it, the receiver needs only knowledge of the desired user's CSI. It can be shown that OC is decomposable into an interference-plus-noise decorrelator stage followed by an MRC stage. The additional complexity of OC compared to MRC is, therefore, due to this decorrelator stage which is also responsible for the superiority of OC over MRC. There are other suboptimal antenna combining schemes that are even simpler than MRC; however, they exhibit even lower overall performance. Examples of such combining techniques are equal gain combining (EGC), selection combining (SC), and generalized selection combining (GSC) [2]. In this thesis, we will only focus on MRC and OC.

1.3 Antenna Correlation

As mentioned in Section 1.1, multipath fading can be modelled as a random variable due to the randomness of the scattering environment. In a multiple antenna receiver, the signals received by different antenna elements are faded differently since they are picked up at different points in space. If the antennas are spatially well separated, it is natural to assume that their received signals experience independent fadings. It can be shown that the independent fading model is accurate when the spacing between any pair of the antennas is at least as large as half of the wavelength of the carrier wave [5]. However, for smaller antenna spacings, the fading coefficients of different antennas become increasingly correlated and, therefore, the independence assumption is not accurate anymore. Realistic correlated fading models calculated based on the propagation properties of the electromagnetic waves are available for modelling closely spaced antennas [5]. The fading correlation between the antennas can severely degrade the performance of multiple antenna receivers. Therefore, it is important to use a correct fading model when analysing the performance of such systems. The independent fading model is appropriate for a base station tower equipped with multiple antennas where enough space is available in order to maintain the necessary spatial separation. On the other hand, for a multiple antenna receiver implemented on a small handheld device such as a cellphone where there is not enough space available, a correlated fading model is more appropriate.

1.4 Literature Review

A review of the related research works available in the literature is provided in this section. These research works are presented under three different categories which are the subjects of our studies in this thesis. The first category lists a number of the most important works on the performance analysis of MRC in the presence of CCI and under different fading scenarios. The second list includes a review of the available literature on the performance analysis of OC for different channel models. Finally, the last category lists a number of research works that study the effect of the array configuration (i.e., the number and the positions of the antennas in the array) on the performance of multiple antenna systems.

1.4.1 Performance Analysis of MRC in the Presence of CCI

The performance of MRC in the presence of AWGN and CCI has been extensively analyzed by many researchers. The outage performance of MRC in this scenario is investigated in [6]–[9]. Closed-form solutions for the bit error rate (BER) of coherent binary phase shift keying (BPSK) and binary differential phase shift keying (DPSK) using MRC in correlated Rician fading, CCI and AWGN are derived in [10] where the probability density function (PDF) of the decision variable, assuming +1 transmitted, is computed by taking the inverse Laplace transform of its moment generating function (MGF) using Cauchy's residue theorem [11],[12]. Then the average BER (i.e., the probability of the decision variable being less than zero) is calculated by integrating this PDF from minus infinity to zero. A different MGF based method is used in [13] and [14] to derive a closed-form solution for the BER of BPSK using MRC in correlated Rayleigh fading, CCI, and AWGN where the correlation matrices of the desired user's channel and that of the interfering users' channels are assumed to have the same eigenvector matrices. In contrast with [10], the PDF of the decision variable in [13] is derived by partial fraction expansion of its MGF and then taking the inverse Laplace transform of each term separately.

1.4.2 Performance Analysis of OC

There is also a large body of literature on the performance analysis of OC both under the uncorrelated array assumption, e.g., [15]–[19] and the correlated array assumption, e.g., [20]–[23]. However, due to the complex form of the expression for the output SINR of OC, its analysis particularly in the correlated case is, in general, very difficult. Therefore, various simplifying assumptions and approximations have been adopted in the literature. For instance, in [20] and [22], it is assumed that only the array signals received from the interfering users are correlated while those received from the desired user are independent. In [22] and [19], it is assumed that the received signal is corrupted with only one interfering user and with AWGN. Shah *et al.* [23] and Zhang *et al.* [21] analyze the outage performance of OC while assuming an interference dominant environment where the effect of AWGN is neglected. Under this assumption, it is shown in [23] that the PDF of the output signal-to-interference ratio (SIR) of OC is independent of the correlation between the antennas and that the average output SIR grows linearly with the number of antennas as long as the number of interferers is large enough. However, by neglecting the AWGN, the OC receiver exists only if the number of the interfering users is no less than the number of the antennas. As a result, the analysis provided in [23] and [21] is only valid for this special case.

1.4.3 The Effect of Array Configuration on Multiple Antenna Systems

Motivated by the application of multiple receive antennas in size constrained mobile units, it is important to understand how the correlations between the antennas affect the asymptotic performance of the receiver as the number of antennas in the array increases while the total physical size of the array is limited to the available space on the mobile device. An asymptotic analysis of this type has been done for MRC by Beaulieu and Zhang [24], [25] where the average SINR performance of MRC is analyzed as the number of antennas inside a fixed length linear array or a fixed radius circular array increases. A similar asymptotic analysis for the capacity of a multiple input multiple output (MIMO) system with a receiver equipped with a fixed size circular array of antennas is considered in [26]. The research works in [24], [25], and [26] assume a uniform array structure where the antenna spacings are the same for different pairs of adjacent antennas in the array. The problem of optimizing the locations of the antennas has also been investigated by a few researchers. In [27], the authors consider the optimization of the base station antenna spacing for a uniform linear and/or circular array in a space division multiple access (SDMA) system. They do so by plotting the average of the squared correlation coefficient between all the antenna pairs as a function of the antenna spacing and by searching for a point on the plot that corresponds to the minimum average correlation. A similar problem is considered in [28] for a SDMA scenario with four fixed (i.e., non-mobile) users and four antennas on the base station whose locations are to be optimized. The work in [28] assumes line of sight (LOS) channels for all the users and, using a genetic algorithm, searches for a 2D antenna arrangement that minimizes the average MMSE of the users.

1.5 Thesis Outline and Major Contributions

The BER performance of MRC is investigated in Chapter 2 where two novel exact and closed-form expressions for the average BER of binary phase shift keying using MRC in the presence of correlated Rayleigh fading, CCI and AWGN are derived. The derived BER expressions are much simpler than those available in the literature. Moreover, using the derived BER expressions, it is shown through some numerical examples that for a fixed number of antennas, there is a best radius for a uniform circular array at which the array performance is almost as if all antennas in the array were uncorrelated. The average SINR performance analysis of OC in the presence of CCI, AWGN, and a general form of correlated Rayleigh fading is discussed in Chapter 3. In the case of OC, our main focus is on determining the change in the average output SINR with the number of antennas when the total physical size of the array is fixed. In particular, we show that if the average received power per antenna is fixed, the average output SINR of OC increases linearly with the number of antennas despite the fact that the array becomes more correlated as more antennas are introduced to the array. This is not true for MRC according to the results of [24] and [25]. The correlation between a set of antennas in a fixed size array is a function of their number as well as their relative positions. Therefore, one should be able to maximize the array's performance by optimizing these parameters. This is the subject of our research in Chapter 4 where we develop a mathematical framework for the optimization of the positions of the antennas in a multiple antenna receiver. We adopt the average output SINR of OC derived in Chapter 3 as a measure of the array's performance and attempt to maximize it by properly arranging the antennas while assuming that the total physical dimensions of the array are fixed. We start with the case of a linear 1D array and later generalize it to the case of a 2D array with an arbitrary shape. The results on the performance analysis of OC presented in Chapter 3 as well as those on optimum antenna placement in Chapter 4 are valid for any modulation scheme. Finally, Chapter 5 concludes the thesis.

Chapter 2

BER of MRC in Correlated Rayleigh Fading and CCI

In this chapter[1], we derive two new exact and closed-form expressions for the average BER of BPSK using MRC in the presence of correlated Rayleigh fading, CCI and AWGN where the assumption of [13] on the eigenvector matrices of the correlation matrices is adopted. Our derviations are based on the MGF method [2] and the application of the residue theorem [12]. However, in contrast with [10], [13], and [14] (see Section 1.4), the average BER is directly calculated from the MGF of the decision variable through a single integral resulting in new simpler expressions. Using the derived expressions, it is shown through some numerical examples that for a fixed number of antennas, there is a best radius for a uniform circular array at which the array performance is almost as if all antennas in the array were uncorrelated. Counter-intuitively, at this radius, the adjacent antennas in the array are not uncorrelated. Moreover, as an additional contribution, it is shown that the assumption on the equality of the eigenvector matrices of the desired user's and the interferers' correlation matrices adopted in [13] and in this work corresponds to a number of different physical realizations and is not only limited to the case of equal correlation matrices.

---

[1] *A version of this chapter has been published in part in IEEE Transactions on Communications, vol. 57, no. 3, pp. 630-634, 2009 and in Proceedings of IEEE International Conference on Communications (ICC), 2008, pp. 1370-1376.*

2.1 System Model

The equivalent baseband signal vector received by an $M-$element array of antennas in the presence of cochannel interference and AWGN is given by

$$\mathbf{r}(t) = \sum_{k=-\infty}^{\infty} \left\{ \sqrt{P_0}\mathbf{g_0}d_k^{(0)}s(t-kT) + \sum_{i=1}^{N} \sqrt{P_i}\mathbf{g_i}d_k^{(i)}s(t-kT-\tau_i) \right\} + \mathbf{n}(t) \quad (2.1)$$

where $s(t)$ is a unit energy signal waveform used by all users, T is the symbol interval, N is the number of interfering users , $d_k^{(0)}$ is the transmitted data symbol of the desired user at the k'th time slot, and $d_k^{(i)}$, $(i = 1, \ldots, N)$ is the transmitted data symbol of the i'th interfering user at the k'th time slot. The $d_k^{(i)}$'s are independent and identically distributed (i.i.d.) and chosen from a BPSK constellation each assuming a value of $+1$ or -1 with a probability of 0.5. The $\mathbf{g_i}$'s $(i = 0, \ldots, N)$ are independent M by 1 zero-mean complex symmetric Gaussian random vectors representing the Rayleigh channel between each user and the array elements in the receiver. The desired user's channel vector $\mathbf{g_0}$ has covariance matrix $\mathbf{R_0}$ and the interfering users' channel vectors $\mathbf{g_i}$'s $(i = 1, \ldots, N)$ all have the same covariance matrix $\mathbf{R_I}$. τ_i is the time delay between the reception of the i'th interfering user's signal and the desired user's signal. The τ_i's are all equal to zero in a synchronous CCI scenario. However, In the asynchronous CCI case, they are modeled as i.i.d. random variables uniformly distributed over $[0, T)$. The parameter P_i is the received power associated with the i'th user's signal, and $\mathbf{n}(t)$ denotes the complex AWGN vector whose elements are independent and each has a power spectral density of $N_0/2$ per dimension.

$\mathbf{R_0}$ and $\mathbf{R_I}$ are non-negative Hermitian matrices. Therefore, their eigenvalues are all real and non-negative. Let $\mathbf{R_0} = \mathbf{Q_0}\mathbf{\Lambda_0}\mathbf{Q_0}^H$ ($\mathbf{R_I} = \mathbf{Q_I}\mathbf{\Lambda_I}\mathbf{Q_I}^H$) denote the eigenvalue decomposition (EVD) of $\mathbf{R_0}$ ($\mathbf{R_I}$) where $\mathbf{Q_0}$ ($\mathbf{Q_I}$) is a unitary matrix containing the eigenvectors of $\mathbf{R_0}$ ($\mathbf{R_I}$) in its columns, $\mathbf{\Lambda_0}$ ($\mathbf{\Lambda_I}$) is a diagonal matrix containing the eigenvalues of $\mathbf{R_0}$ ($\mathbf{R_I}$), and $(\cdot)^H$ is the Hermitian (conjugate transpose) operator. We adopt the same assumption as in [13] that there exist eigenvalue decompositions of $\mathbf{R_0}$ and $\mathbf{R_I}$ in which $\mathbf{Q_0}$ is equal to $\mathbf{Q_I}$, however, $\mathbf{\Lambda_0}$ and $\mathbf{\Lambda_I}$ can in general be different. As an example of a physical realization in which this assumption holds, one can consider a circular array of antennas where the antenna elements are uniformly distributed on the circumference of a circle located in the horizontal plane and where the azimuth angle of arrival is uniformly distributed over $[0, 2\pi)$ (i.e., omnidirectional reception in the horizontal plane) for both the desired user's signal and the interfering signals. The symmetry arising from the omnidirectional reception in the horizontal plane guarantees that the correlation coefficient between each pair of the antennas in the array is only a function of their distance [29], [30]. This, along with the circular structure of the array, implies that the correlation matrices $\mathbf{R_0}$ and $\mathbf{R_I}$ are both circulant [31] and, therefore, both have the same eigenvector matrix given by the discrete Fourier transform (DFT) matrix (i.e., $\mathbf{Q_0} = \mathbf{Q_I} = \mathbf{DFT}$) [31]. However, the set of eigenvalues of $\mathbf{R_0}$ and that of $\mathbf{R_I}$ which are given by the DFT transform of the first row of $\mathbf{R_0}$ and $\mathbf{R_I}$, respectively, may be different due to a different distribution of the elevation angle of arrival for the desired signal and the interference. Other examples that satisfy the above-mentioned assumption are the cases where either $\mathbf{R_0}$ or $\mathbf{R_I}$ is the identity matrix $\mathbf{I}_M$, or where $\mathbf{R_0} = \mathbf{R_I}$. The last case corresponds, for example, to an urban environment, where the multiple antenna receiver is surrounded by a large number of local scatterers with no line of sight to the desired or the interfering base stations.

Let $x(t) = s(t) \star s(-t)$ where $\star$ denotes convolution. The signal waveform $s(t)$ is designed such that $x(t)$ is a Nyquist pulse [1]. After match-filtering, sampling and maximal ratio combining of the received signal vector in (2.1), the decision variable for the detection of the desired user's data at the m'th time interval is given by [1]

$$\begin{aligned} y_m =& \mathrm{Re}\left\{\mathbf{g_0}^H \int_{-\infty}^{\infty} \mathbf{r}(t) s(t - mT) dt\right\} \\ =& \sqrt{P_0}\mathbf{g_0}^H \mathbf{g_0} d_m^{(0)} + \sum_{i=1}^{N} \sqrt{P_i} \mathrm{Re}\left\{\mathbf{g_0}^H \mathbf{g_i}\right\} a_i \\ &+ \mathrm{Re}\left\{\mathbf{g_0}^H \mathbf{n}\right\} \end{aligned} \tag{2.2a}$$

where

$$a_i = \sum_{k=-\infty}^{\infty} d_{k+m}^{(i)} x(kT + \tau_i) \tag{2.2b}$$

and $\mathbf{n}$ is a complex AWGN vector whose elements are independent each with variance N_0. An estimate of $d_m^{(0)}$ is then obtained by taking the sign of y_m.

2.2 BER Calculation

Due to the symmetry, the total probability of error for detection of $d_m^{(0)}$ is equal to the conditional error probability when $d_m^{(0)} = +1$ is transmitted and is given by

$$P_e = \Pr\left(y_m < 0 | d_m^{(0)} = 1\right) = F_{Y_m}(0|d_m^{(0)} = 1) \tag{2.3}$$

where y_m is defined in (2.2) and $F_{Y_m}(y|d_m^{(0)} = 1)$ is the conditional cumulative distribution function (CDF) of y_m. From the definition of the characteristic function and the integration property of the Fourier transform, one can express the conditional CDF of y_m as [32]

$$F_{Y_m}(y|d_m^{(0)} = 1) = \frac{1}{2} - \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\Phi(\omega)}{j\omega}\, e^{-j\omega y}\, d\omega \tag{2.4}$$

where $\Phi(\omega)$ is the characteristic function of y_m given $d_m^{(0)} = 1$. Because of the circular symmetry of the PDF of the channel vectors $\mathbf{g_i}$'s, the phases of a_i's in (2.2) do not have any effect on the distribution of y_m. Therefore, each a_i in (2.2a) can be replaced by its magnitude $|a_i|$ without changing the probability of error in (2.3).

2.2.1 Synchronous CCI

In the case of synchronous CCI, the τ_i's in (2.1) are all equal to zero. Since $x(t)$ is a Nyquist pulse, $x(kT) = \delta_k$ where δ_k is the Kronecker's delta function. Therefore a_i in (2.2b) will be reduced to $d_m^{(i)}$ (i.e., $|a_i| = 1$). In this case, under the assumption of equal eigenvector matrices for $\mathbf{R_0}$ and $\mathbf{R_I}$, it is shown in [13] that $\Phi(\omega)$ can be expressed as

$$\Phi(\omega) = \prod_{i=1}^{M} \frac{1}{1 - j\omega\lambda_0^i + \omega^2\left(\lambda_0^i\lambda_I^i/\Gamma_I + \lambda_0^i/\Gamma_0\right)/4} \tag{2.5}$$

where the λ_0^i's and λ_I^i's are the eigenvalues of $\mathbf{R_0}$ and $\mathbf{R_I}$, respectively, $\Gamma_0 = P_0/N_0$ is the input average SNR, and $\Gamma_I = P_0/\left(\sum_{i=1}^{N} P_i\right)$ is the input average

SIR. If the correlation matrices of the interfering users are not the same, as long as they all have the same eigenvector matrices, one only needs to replace λ_I^i in (2.5) by a weighted average of the i'th eigenvalues of the interfering users where the weights are given by $w_j = P_j / \left(\sum_{j=1}^{N} P_j\right)$. Therefore, the calculations in the sequel are all applicable to this general case as well. From (2.3), (2.4), and (2.5) one can see that as long as the total interference power $\sum_{i=1}^{N} P_i$ is kept fixed, $\Phi(\omega)$ and consequently the probability of error are not a function of either the number of interfering users N or the distribution of the interfering users' powers $P_1, \ldots, P_N$. This is not the case for asynchronous CCI which will be discussed in Section 2.2.2. By substituting (2.5) in (2.4), setting $y = 0$, and changing the integration variable to $s = j\omega$, the probability of error in (2.3) can be expressed as

$$P_e = \frac{1}{2} - \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \mathcal{A}(s) ds \tag{2.6a}$$

where

$$\mathcal{A}(s) = \frac{1}{s} \prod_{i=1}^{M} \frac{1}{1 - \lambda_0^i s - \frac{1}{4}\left(\lambda_0^i \lambda_I^i / \Gamma_I + \lambda_0^i / \Gamma_0\right) s^2}. \tag{2.6b}$$

It is easy to show that the poles of $\mathcal{A}(s)$ are all real. There is a pole at $s_0 = 0$, M negative poles located at

$$s_N^i = -2 \left(\frac{1 + \sqrt{1 + \frac{\lambda_I^i}{\Gamma_I \lambda_0^i} + \frac{1}{\Gamma_0 \lambda_0^i}}}{\frac{\lambda_I^i}{\Gamma_I} + \frac{1}{\Gamma_0}} \right), \quad i = 1, \ldots, M \tag{2.7a}$$

and M positive poles located at

$$s_P^i = -2 \left(\frac{1 - \sqrt{1 + \frac{\lambda_I^i}{\Gamma_I \lambda_0^i} + \frac{1}{\Gamma_0 \lambda_0^i}}}{\frac{\lambda_I^i}{\Gamma_I} + \frac{1}{\Gamma_0}} \right), \quad i = 1, \ldots, M. \tag{2.7b}$$

In general, some of the negative poles in (2.7a) can have the same values (i.e., repeated poles). This can happen for the positive poles in (2.7b) as well. Since the derivation of the BER expression depends on the multiplicity of the poles, the following definitions will be required in the sequel. Suppose that there are N_n different values of s_N^i in (2.7a), namely, s_i, $i = 1, \ldots, N_n$, each having a multiplicity of m_i, $i = 1, \ldots, N_n$, such that $\sum_{i=1}^{N_n} m_i = M$. Similarly,

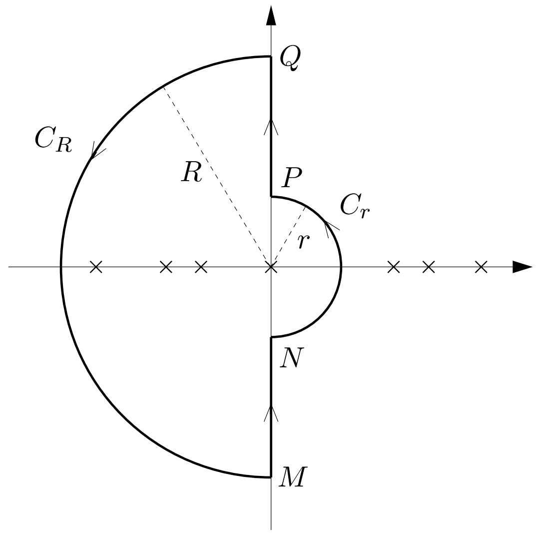


Figure 2.1. An illustration of the contour integration of $\mathcal{A}(s)$ in the complex s-plane. The $\times$ symbols represent the poles of $\mathcal{A}(s)$.

suppose that there are N_p different values of s_P^i in (2.7b), namely, s_i, $i = N_n+1, \ldots, N_n+N_p$, each having a multiplicity of m_i, $i = N_n+1, \ldots, N_n+N_p$, such that $\sum_{i=N_n+1}^{N_n+N_p} m_i = M$. The pole at $s_0 = 0$ is always a simple pole, therefore, its multiplicity is $m_0 = 1$.

The integral in (2.6a) can be calculated using the method of residue integration [12]. As shown in Fig.2.1, the integral of $\mathcal{A}(s)$ over the $j\omega$-axis can be expressed as

$$\int_{-j\infty}^{j\infty} \mathcal{A}(s)ds = \lim_{R\to\infty} \lim_{r\to 0} \left(\oint_C \mathcal{A}(s)ds - \int_{C_R} \mathcal{A}(s)ds - \int_{C_r} \mathcal{A}(s)ds \right) \tag{2.8}$$

where, in Fig. 2.1, C_R is the large half-circle with radius R, C_r is the small half-circle with radius r, and $C = \overline{MN} \cup C_r \cup \overline{PQ} \cup C_R$ is the entire loop. As soon as R becomes large enough such that all non-positive poles of $\mathcal{A}(s)$ are located inside C and r becomes small enough such that all positive poles of $\mathcal{A}(s)$ are positioned outside C, the first term in the right of (2.8) becomes independent of R and r, according to the residue theorem [12], and is given by

$$\lim_{R\to\infty} \lim_{r\to 0} \oint_C \mathcal{A}(s)ds = 2\pi j \sum_{i=0}^{N_n} \text{Res}(s_i) \tag{2.9}$$

where $\text{Res}(s_i)$ is the complex residue of $\mathcal{A}(s)$ at s_i. To calculate the second term in the right of (2.8) one should note that the magnitude of $\mathcal{A}(s)$ decreases at least as fast as $1/R^3$. Therefore

$$\left| \int_{C_R} \mathcal{A}(s) ds \right| \leq \int_{C_R} |\mathcal{A}(s)| ds \leq \frac{K}{R^3} \cdot \pi R = \frac{K\pi}{R^2} \tag{2.10}$$

for some positive number K. This results in

$$\lim_{R \to \infty} \int_{C_R} \mathcal{A}(z) dz = 0. \tag{2.11}$$

The third term has, however, a non-zero value. The integral of $\mathcal{A}(s)$ over C_r can be calculated by setting $s = re^{j\omega}$ where ω varies from 0 to π. Let $\mathcal{B}(s) \triangleq s\mathcal{A}(s)$. Then,

$$\lim_{r \to 0} \int_{C_r} \mathcal{A}(s) ds = \lim_{r \to 0} \int_{C_r} \frac{\mathcal{B}(s)}{s} ds = \lim_{r \to 0} \int_0^{\pi} j\mathcal{B}(re^{j\omega}) d\omega = j\pi \lim_{s \to 0} \mathcal{B}(s) = j\pi \mathrm{Res}(s_0) \tag{2.12}$$

where the last equality is simply the definition of the residue of $\mathcal{A}(s)$ at $s = s_0 = 0$. Moreover, from the definition of $\mathcal{A}(s)$ in (2.6b) one can easily show that $\mathrm{Res}(s_0) = 1$. Combining this with (2.6a) and (2.8) – (2.12), the probability of error for the synchronous CCI case can be written as

$$P_e^{SC}(\Gamma_0, \Gamma_I, \mathbf{R_0}, \mathbf{R_I}) = -\sum_{i=1}^{N_n} \mathrm{Res}(s_i). \tag{2.13}$$

Similarly, one can calculate the integral in (2.6a) by integrating $\mathcal{A}(s)$ over the mirror reflection of the contour in Fig. 2.1 about the $j\omega$-axis. This results in an alternate expression for the probability of error given by

$$P_e^{SC}(\Gamma_0, \Gamma_I, \mathbf{R_0}, \mathbf{R_I}) = 1 + \sum_{i=N_n+1}^{N_n+N_p} \mathrm{Res}(s_i). \tag{2.14}$$

The partial fraction expansion of $\mathcal{A}(s)$ in (2.6b) can be written as

$$\mathcal{A}(s) = \sum_{i=0}^{N_n+N_p} \sum_{l=1}^{m_i} \frac{\alpha_{i,l}}{(s - s_i)^l} \tag{2.15}$$

where the expansion coefficients $\alpha_{i,l}$'s are given in [8], [13], and [14]. In [13] and [14], the calculation of all of the expansion coefficients is required for the derivation of the BER. In contrast, based on (2.13) or (2.14), we only need to compute either the residues at the negative poles, i.e., the $\alpha_{i,l}$'s, for $l = 1$ and $i = 1, \ldots, N_n$ or the residues at the positive poles, i.e., the $\alpha_{i,l}$'s, for $l = 1$ and $i = N_n + 1, \ldots, N_n + N_p$. This simplifies the calculations significantly. The values of these residues are given by [8], [13], [14]

$$\operatorname{Res}(s_i) = \alpha_{i,1} = \frac{(-1)^{M+m_i-1}}{\prod_{n=1}^{M} b_n} \sum_{Q(i)} \prod_{\substack{k=0 \\ k \neq i}}^{N_n+N_p} \frac{\binom{m_k+q_k-1}{q_k}}{(s_i - s_k)^{m_k+q_k}}\ , \quad i = 1, \ldots, N_n + N_p \tag{2.16}$$

where $b_n = \frac{1}{4}\left(\lambda_0^n \lambda_I^n / \Gamma_I + \lambda_0^n / \Gamma_0\right)$ and $Q(i) = \{(q_0, \ldots, q_{N_n+N_p}) \mid q_k \in \mathbb{N}_0,\ q_i = 0,\ \sum_{k=0}^{N_n+N_p} q_k = m_i - 1\}$ with $\mathbb{N}_0$ being the set of all nonnegative integers. Using (2.7) and (2.16), one can use either (2.13) or (2.14) to calculate the probability of error. However, the application of one may be computationally more efficient than the other. In fact, based on (2.16), the complexity of the BER calculation in (2.13) or (2.14) is dominated by the complexity of the computation of the residue at the pole with the highest order. Therefore, if $\max\{m_i \mid i = 1, \ldots, N_n\} < \max\{m_i \mid i = N_n + 1, \ldots, N_n + N_p\}$, it is computationally more efficient to use (2.13), otherwise, using (2.14) is preferred. When the negative poles of $\mathcal{A}(s)$ are all simple (i.e., $N_n = M$ and $m_i = 1, \text{for } i = 1, \ldots, M$), the final expression for the probability of error using (2.13) and (2.16) is simplified to

$$P_e^{SC}\left(\Gamma_0, \Gamma_I, \mathbf{R_0}, \mathbf{R_I}\right) = \frac{1}{2} \sum_{i=1}^{M} \pi_i \left(1 - \sqrt{\frac{\bar{\gamma}_i}{1 + \bar{\gamma}_i}}\right) \tag{2.17a}$$

where

$$\bar{\gamma}_i = \frac{\lambda_0^i}{\frac{\lambda_I^i}{\Gamma_I} + \frac{1}{\Gamma_0}} \tag{2.17b}$$

$$\pi_i = \prod_{\substack{k=1 \\ k \neq i}}^{M} \frac{1}{1 - \lambda_0^k s_N^i - \frac{1}{4}\left(\lambda_0^k \lambda_I^k / \Gamma_I + \lambda_0^k / \Gamma_0\right)(s_N^i)^2} \tag{2.17c}$$

and where the s_N^i's are defined in (2.7a). A similar expression can be derived using (2.14) and (2.16) when the positive poles of $\mathcal{A}(s)$ are all simple. In the uncorrelated case where both $\mathbf{R_0}$ and $\mathbf{R_I}$ are identity matrices, all the eigenvalues are equal to one and, therefore, $\mathcal{A}(s)$ has exactly one negative pole and one positive pole each with a multiplicity of M (given by (2.7a) and (2.7b)).

In this case, the average BER expression in (2.13) reduces to the results given in [1, eq. (14.4-15)] and [33]. Eqs. (2.13), (2.14), and (2.17) provide new closed-form expressions for the BER of binary phase shift keying with MRC diversity in correlated synchronous CCI and Rayleigh fading. The derived expressions are much more efficient than Monte Carlo methods for computing the average BER of MRC, specially at high SNR values. This is due to the fact that the complexity of these expressions is not a function of the SNR. In contrast, the complexity of a Monte Carlo method grows (exponentially) with SNR as it is increasingly less likely to generate an error event at high SNR values. For instance, using the derived expressions, the BER values on the order of 10^{-8} can be calculated in a fraction of a second using MATLAB on a desktop computer while it takes on the order of an hour to reliably calculate it using Monte Carlo simulation.

2.2.2 Asynchronous CCI

In the presence of asynchronous CCI, the τ_i's in (2.2b) are i.i.d. random variables uniformly distributed over $[0, T)$. Unlike the synchronous case where $|a_i| = 1$, $(i = 1, \ldots, N)$, the $|a_i|$'s in this case are i.i.d. random variables whose common PDF is, in general, a function of the pulse shape used by the users. Although the distribution of y_m is different in this case, its conditional distribution given the $|a_i|$'s remains the same as the synchronous case provided that the interference powers P_i's in (2.2a) are replaced by $P_i|a_i|^2$. Therefore, one can still use (2.13), (2.14), and (2.17) to find the conditional probability of error given the $|a_i|$'s after replacing P_i with $P_i|a_i|^2$, $(i = 1, \ldots, N)$. The average probability of error is then obtained by integrating the conditional probability of error over the joint distribution of the $|a_i|$'s. However, since (2.13), (2.14), and (2.17) are related to the interference powers only through $\Gamma_I = P_0/\left(\sum_{i=1}^{N} P_i\right)$, the probability of error in asynchronous cochannel interference, P_e^{AC}, is obtained by a single integral and is given by

$$P_e^{AC}(\Gamma_0, \Gamma_1, \ldots, \Gamma_N, \mathbf{R_0}, \mathbf{R_I}) = \int_0^\infty P_e^{SC}(\Gamma_0, \frac{P_0}{P_I}, \mathbf{R_0}, \mathbf{R_I}) f(P_I) dP_I \quad (2.18a)$$

where

$$P_I = \sum_{i=1}^{N} P_i |a_i|^2 \tag{2.18b}$$

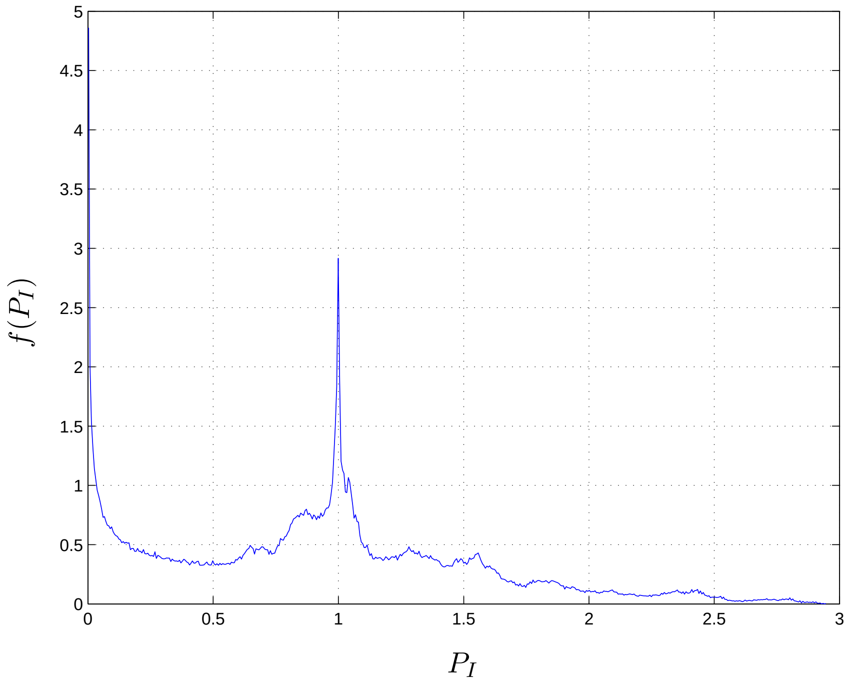
and

$$\begin{aligned}\Gamma_i &= \frac{P_0}{P_i}\left[\mathrm{E}\{|a_i|^2\}\right]^{-1} \\ &= \frac{P_0}{P_i}\left(\frac{1}{T}\int_{-\infty}^{\infty} x^2(t)dt\right)^{-1}, \quad i = 1, \ldots, N\end{aligned} \tag{2.18c}$$

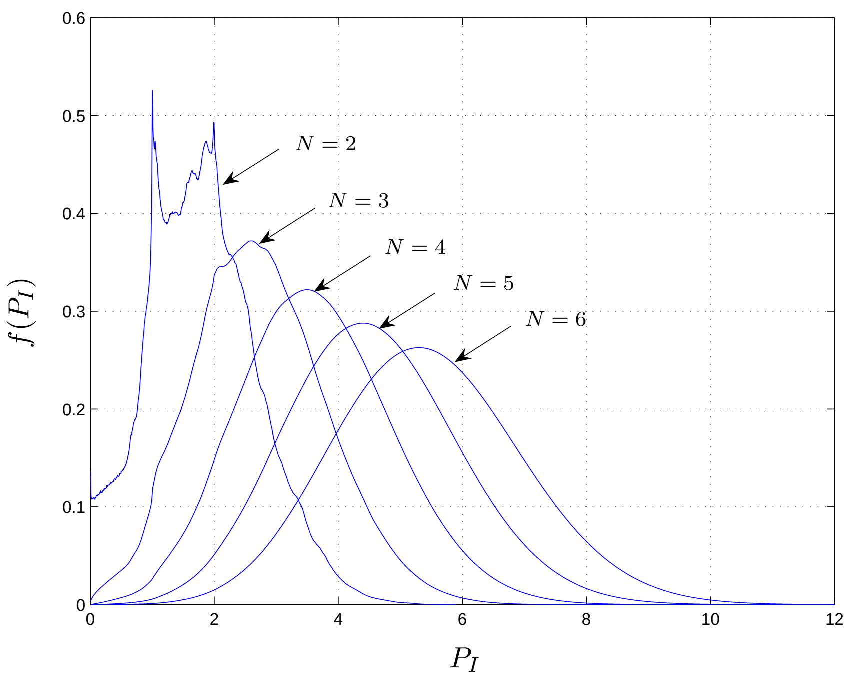
is the average input SIR corresponding to the i'th interfering user, $f(P_I)$ is the probability density function of P_I, and $\mathrm{E}\{\cdot\}$ is the expectation operator. In general, $f(P_I)$ is a function of the pulse shape $s(t)$ used by the users as well as the number of interfering users and their power distribution. Therefore, unlike the synchronous CCI case, the probability of error in the presence of asynchronous CCI given in (2.18a) varies with the number of interfering users and their power distribution, even when the total average interference power is fixed. Although it seems difficult to calculate $f(P_I)$ analytically, one can always find it numerically. For example, Fig. 2.2 shows $f(P_I)$ for different values of N when the P_i's are all normalized to unity and $x(t)$ is a raised-cosine pulse with roll-off factor $\beta = 0.35$. From Fig. 2.2 one can observe that $f(P_I)$ rapidly approaches a Gaussian density as N increases. This is an immediate result of the central limit theorem since P_I defined in (2.18b) is a sum of N i.i.d. random variables. Once $f(P_I)$ is found, the integral in (2.18a) can be calculated numerically. At high SNR values, this semi-analytical approach in calculating the average BER of MRC in the presence of asynchronous CCI is still much more efficient than Monte Carlo simulation as its complexity does not grow with SNR.

2.3 Numerical Results and Discussion

This section includes some numerical examples where the expressions derived in Section 2.2 are used to calculate the average BER of the desired user in an MRC diversity receiver in the presence of AWGN, correlated CCI and correlated Rayleigh fading. In the first example, a circular array of $M = 5$ antennas



(a)



(b)

Figure 2.2. The probability density function of the total interference power P_I for different values of the number of interfering users. The pulse shape $s(t)$ used by all users is a root raised-cosine pulse with a roll-off factor $\beta = 0.35$. (a) $N = 1$ (b) $N = 2, \ldots, 6$.

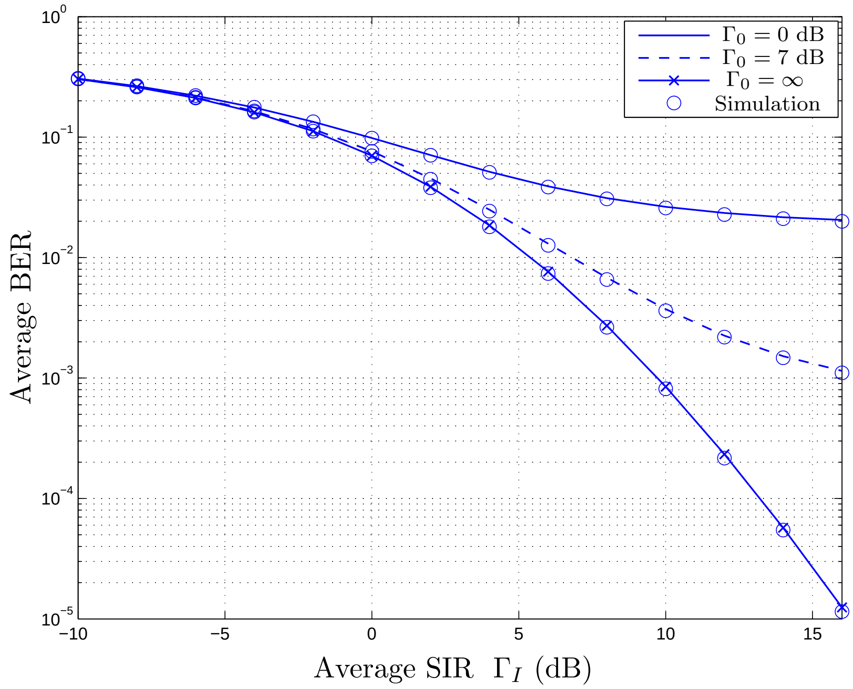


Figure 2.3. The average BER versus SIR for MRC in the presence of synchronous CCI and AWGN using a circular array of radius 0.1λ containing $M = 5$ uniformly distributed antennas with 2D-omnidirectional reception of the desired signal and 3D-omnidirectional reception of the interference signals.

located in the horizontal plane is considered. It is assumed that the desired user's signal arrive only from the broadside direction and its azimuth angle of arrival is uniformly distributed. This is known as two-dimensional omnidirectional (2D-omnidirectional) scattering [25]. The interference signals are assumed to arrive from any direction in the space with a uniform distribution.This is known as three-dimensional omnidirectional (3D-omnidirectional) scattering [34], [30]. Under these assumptions, the correlation coefficient between the signals received by any pair of the antenna elements is given by $\mathrm{J}_0(2\pi d)$ for the desired user's signal and by $\mathrm{sinc}(2\pi d)$ for the interferer signals [30], [5], [34] where d is the ratio of the distance between the two antennas to the carrier wavelength λ, $\mathrm{J}_0(\cdot)$ is the zeroth-order Bessel function of the first kind, and $\mathrm{sinc}(x) = \sin(x)/x$. The CCI is assumed to be synchronous. Fig. 2.3 shows the average BER of this system as a function of the average SIR, Γ_I for different values of the average SNR, Γ_0. The normalized array radius (i.e.

the ratio of the array radius to the wavelength) is set to $r = 0.1$. Except for the infinite SNR curve which corresponds to a CCI limited scenario, the other curves exhibit a BER floor since the performance of the system in the large SIR regime is mostly limited by the AWGN or equivalently the average SNR. The open circles show the simulation results. It can be seen that the analytical results are in excellent agreement with the simulation results. It is worth noting that because of the circular structure of the array in this example, the desired signal's correlation matrix $\mathbf{R_0}$ has the form

$$\mathbf{R_0} = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_2 & \rho_1 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_1 & \rho_2 & \rho_2 & \rho_1 & 1 \end{pmatrix} \tag{2.19}$$

with ρ_1 and ρ_2 being some real numbers between -1 and $+1$. The interference correlation matrix $\mathbf{R_I}$ has the same form but with a different set of values for ρ_1 and ρ_2. This is a circulant matrix [31]. Therefore, its eigenvalues are the DFT of its first row [31]. But, because of the symmetric structure of the first row one can show that its DFT is a real sequence with the same symmetric structure. This means that there are at most $\lfloor M/2 \rfloor + 1$ different eigenvalues where $\lfloor x \rfloor$ is the largest integer smaller than or equal to x. Therefore, some of the poles of $\mathcal{A}(s)$ in (2.7) have a multiplicity of greater than or equal to two. This is true for any uniform circular array with an arbitrary number of antennas under the assumption of omnidirectional reception.

In the second example, a similar scenario to the first example is considered only with asynchronous CCI this time. The pulse shape $s(t)$ used by all the users is chosen to be a root raised-cosine pulse with a roll-off factor $\beta = 0.35$. Moreover, it is assumed that the SIRs corresponding to the interfering users (defined in (2.18c)) are all the same and equal to $\Gamma_i = 17$ dB. Fig. 2.4 shows the average BER of this system versus the average input SNR for different values of the number of interfering users N. These curves are obtained by numerically calculating the integral in (2.18a) using $f(P_I)$ shown in Fig. 2.2. Again, there exists an error floor in all of the curves since the system's performance for high SNRs is limited by the interference. As expected, this error floor moves

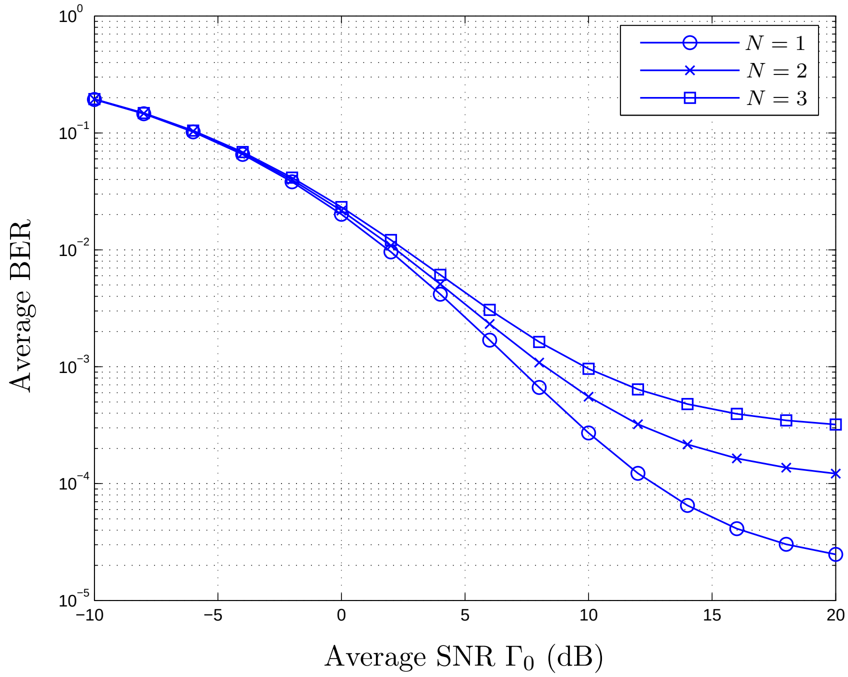


Figure 2.4. The average BER versus SNR for MRC in the presence of asynchronous CCI and AWGN with the number of interfering users N as a parameter. A root raised-cosine pulse with $\beta = 0.35$ is used by all users. The receiver consists of a circular array of radius 0.1λ with $M = 5$ uniformly distributed antennas where 2D-omnidirectional reception of the desired signal and 3D-omnidirectional reception of the interference signals are assumed. The SIRs corresponding to the interfering users are all set to $\Gamma_i = 17$ dB.

upward as the number of interfering users increases.

The next example considers the effect of antenna spacing on the BER performance of the system. Fig. 2.5 shows the BER of a circular array versus the normalized radius of the array for different values of M and average SIR Γ_I where the CCI is synchronous. It is assumed that the desired signal and the interferer signals are both arriving from the broadside direction (i.e., 2D-omnidirectional reception for all signals). The average SNR is set to infinity corresponding to a CCI-limited scenario. A number of observations can be made from this figure. First, one can see that for small values of the normalized radius r, the BER is almost independent of M for any given value of the average SIR. The reason is that when r is small, the array is extremely packed and, therefore, both the desired signal vector and the interference vector are highly correlated. If the optimal detector (i.e., decorrelator plus MRC) was

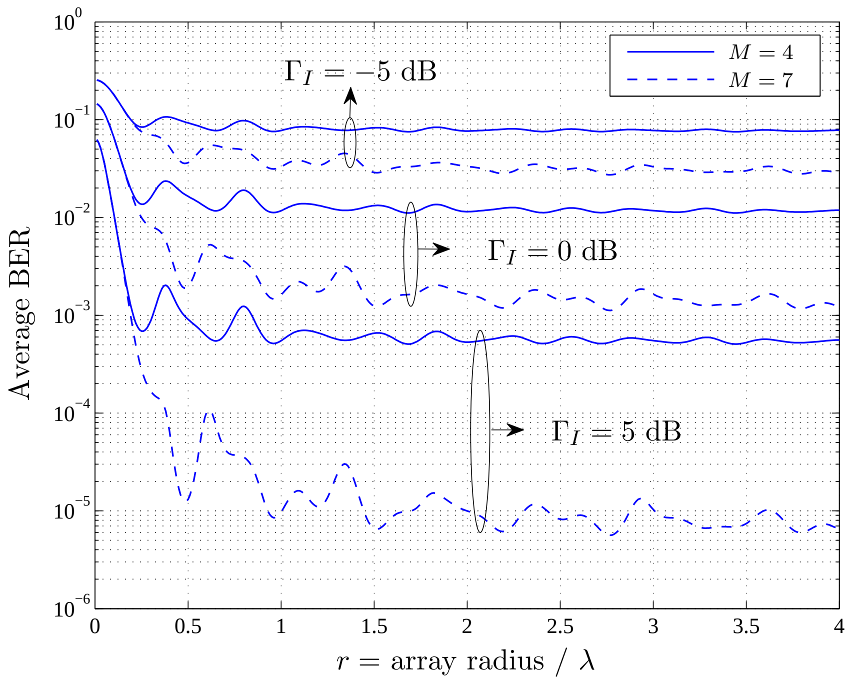


Figure 2.5. The average BER of MRC using a circular array versus the radius of the array normalized by the wavelength. The CCI is synchronous and 2D-omnidirectional reception is assumed for both the desired signal and the interference signals.

deployed, using more antennas would result in a better performance since the optimal detector takes the advantage of the correlation between the interference terms in the received signals in order to better cancel the interference. However, from the viewpoint of the MRC receiver, a highly packed and correlated array is only seen as a single antenna system. In this case, adding more antennas will only increase the system complexity while the achieved performance enhancement is negligible. Second, for any value of M and Γ_I, there is an error floor corresponding to an array of infinite radius where all of the antenna elements in the array are mutually uncorrelated. This is the best achievable BER performance given a fixed number of antennas and at a given average SIR. As expected, this error floor shifts downward as M and/or Γ_I is increased. Finally, the BER curves do not monotonically decrease to their asymptotic limit. This is a consequence of the fact that the magnitude of the correlation between each pair of the antennas (i.e., $\mathrm{J}_0(2\pi d)$) is not a decreasing function of the spacing between the two antennas as seen in [5, Fig.

1.3-6]. Moreover, one can observe that the pattern of the oscillations (i.e., the location of the local minima and maxima) of the BER curve is different for different values of M. However, it is independent of the value of the average SIR when M is fixed.

As seen in Fig. 2.5, a reasonably low BER, almost as low as the asymptotic BER floor, can be achieved by setting the value of the normalized radius r to $r^*(M)$ where the first minimum of the BER occurs. The value of $r^*(M)$ represents a "best" operational value of great practical interest since it determines the normalized radius of the smallest circular array of M antennas that can approximately achieve the best performance among all circular arrays of M antennas. From Fig. 2.5, the best normalized radius for $M = 4$ is $r^*(4) = 0.255$. This corresponds to a normalized distance of 0.36 between the adjacent antennas in the array and that of 0.51 between the antennas on the opposite corners. These normalized distances correspond to correlation coefficients of 0.075 and -0.321, respectively. This disproves the claim of [13] and [14] which assert that in a finite uniform circular array of $M = 4$ antennas (or as phrased in [13], in a square array of $M = 4$ antennas), the best performance is achieved when the distance between the adjacent antennas is such that they are uncorrelated. The best normalized radius for $M = 7$ is given by $r^*(7) = 0.485$. This corresponds to a correlation coefficient of -0.1 between the adjacent antennas, that of -0.248 between the second nearest pairs, and that of 0.142 between the farthest pairs in the array. Again, it can be seen that at the best radius, none of the antenna pairs (particularly, the adjacent antennas) are uncorrelated.

2.4 Chapter Summary

In this chapter, new simple closed-form expressions for the average BER of BPSK using MRC in the presence of correlated Rayleigh fading, CCI, and AWGN were derived. The final solutions are explicitly expressed in terms of the eigenvalues of the correlation matrices of the desired user's channel and that of the interfering users' channels, and are valid for an arbitrary number of interfering users and an arbitrary number of receiver antennas. The derived expressions were then used to calculate the BER performance of MRC systems in some example scenarios. Particularly, it was shown that for a fixed number of antennas in a uniform circular array, there is a best radius for the array corresponding to the smallest uniform circular array that almost achieves its best possible BER performance, as if its radius was infinite (i.e., the totally uncorrelated array). Counter-intuitively, at this radius, the adjacent antennas in the array are not uncorrelated.

In the next chapter, we derive a new expression for the average output SINR of OC under correlated Rayleigh fading and CCI. Then, using the derived expression, we study the the performance of OC as more antennas are introduce to the array while the total physical dimensions of the array are fixed.

Chapter 3

Performance Analysis of OC for Dense Multiple Antenna Reception Under Rayleigh Fading

In this chapter[1], an analytical expression is derived for the average output SINR of optimum combining for a spatially correlated array of antennas in the presence of a single interferer, Rayleigh fading, and AWGN. Using the derived expression and based on an asymptotic analysis of the eigenvalues of dense correlation matrices, the asymptotic performance of optimum combining is evaluated as the number of the antennas increases while the total physical size of the array is fixed. Two different scenarios are considered, namely, fixed average received power per antenna and fixed total average received power. It is shown that in the former scenario, the average output SINR is asymptotically a linear function of the number of the antennas while in the latter scenario it eventually saturates at a certain value. The slope of the asymptote in the former scenario as well as the value of the saturation limit in the latter scenario are derived in terms of the point spectrum of the underlying array correlation function. The case of multiple interferers is examined by simulation and is shown to exhibit similar asymptotic behaviour to the case of one interferer. Here, we provide an asymptotic analysis for OC where, similar to [19] and

[1] *A version of this chapter has been published in part in Proceedings of IEEE Global Telecommunications Conference (GLOBECOM), 2009, pp. 1-8, and in IEEE Transactions on Communications, vol. 58, no. 7, pp. 2014-2022, 2010.*

[22], we assume only one interferer is present. However, in contrast with [19] and [22], we consider a fully correlated model in which both the desired user's channel vector and the interferer's channel vector are correlated. The effect of AWGN is also taken into account in our analysis. A similar asymptotic analysis has been done for MRC by Beaulieu and Zhang [24], [25].

3.1 System Model

After matched-filtering and sampling, the equivalent baseband signal received by an $M-$element array of antennas in the presence of an interfering user is given by

$$\mathbf{y} = \sqrt{P_0(M)}\mathbf{g_0}d_0 + \sqrt{P_1(M)}\mathbf{g_1}d_1 + \mathbf{n} \tag{3.1}$$

where the first term on the right of (3.1) represents the received signal of the desired user, the second term represents the interference, and the third term represents additive noise. The parameters d_0 and d_1 chosen from a complex signal constellation of unit average energy, are i.i.d. transmitted data symbols of the desired user and the interfering user, respectively. The parameters $\mathbf{g_0}$ and $\mathbf{g_1}$ are independent M by 1 zero-mean complex symmetric Gaussian random vectors representing the Rayleigh channel of the desired user and that of the interfering user, respectively. The channel vectors $\mathbf{g_0}$ and $\mathbf{g_1}$ both have the same correlation matrix $\mathbf{R_a} = \mathrm{E}\{\mathbf{g_0}\mathbf{g_0}^H\} = \mathrm{E}\{\mathbf{g_1}\mathbf{g_1}^H\}$. The diagonal elements of the "array correlation matrix" $\mathbf{R_a}$ are all equal to one corresponding to normalized unit variance channel gains. Vector $\mathbf{n}$ denotes the complex AWGN vector whose elements are independent each with variance N_0. The quantities $P_0(M)$ and $P_1(M)$ are the average power per branch received from the desired user and the interfering user, respectively. They are represented as functions of M, the number of antennas, to include two different scenarios considered here. In one scenario the average input power per branch is fixed as the number of antennas increases. This is equivalent to setting $P_0(M) = P_0$ and $P_1(M) = P_1$ where P_0 and P_1 are some constant values. As stated in [35], this scenario would apply, for example, as small fixed-size dipole antennas are added to a sparse array of antennas. In the other scenario, the total average input power is fixed as M increases. This is modeled by setting $P_0(M) = P_0^t/M$ and $P_1(M) = P_1^t/M$ where P_0^t and P_1^t are the total average power input to the array received from the desired user and the interfering user, respectively. This power scenario corresponds to a situation where a single antenna with an effective electromagnetic area A is replaced by two antennas each having an effective area of $A/2$ or by three antennas each having an effective area of $A/3$ and so on. Let

$$\mathbf{R_a} = M\mathbf{Q}\mathbf{A}\mathbf{Q}^H \tag{3.2}$$

denote the EVD of $\mathbf{R_a}/M$ where $\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_M]$ is a unitary matrix containing the eigenvectors of $\mathbf{R_a}$ in its columns and $\mathbf{A} = \mathrm{diag}(\alpha_{1,M}, \alpha_{2,M}, \ldots, \alpha_{M,M})$ is a diagonal matrix containing the eigenvalues of $\mathbf{R_a}/M$. The correlation matrix $\mathbf{R_a}$ is nonnegative definite. Therefore, the $\alpha_{i,M}$'s are all nonnegative. Moreover, $\sum_{i=1}^{M} \alpha_{i,M} = \mathrm{tr}(\mathbf{R_a}/M) = 1$ where $\mathrm{tr}(\cdot)$ indicates the trace operator. Thus, the $\alpha_{i,M}$'s are all less than or equal to one. Without loss of generality, we assume that the eigenvalues are sorted in descending order, i.e., $1 \geqslant \alpha_{1,M} \geqslant \alpha_{2,M} \geqslant \ldots \geqslant \alpha_{M,M} \geqslant 0$.

The optimum combining receiver combines the vector of received signals $\mathbf{y}$ to obtain the decision variable $x = \mathbf{w}^H\mathbf{y}$. The vector of weights $\mathbf{w}$ is given by [2, p. 683]

$$\mathbf{w} = \mathbf{R^{-1}g_0} \tag{3.3a}$$

where

$$\mathbf{R} = N_0\mathbf{I}_M + P_1(M)\mathbf{g_1}\mathbf{g_1}^H \tag{3.3b}$$

is the short-term correlation matrix of the interference-plus-noise component in the received signal. By short-term, we mean that the expectation is taken within the coherence time of the channel where the channel vectors remain fixed but the data symbols and the AWGN are random variables. The short term average SINR at the output of the optimum combiner can be expressed as [2, p. 683]

$$\gamma(M) = P_0(M)\mathbf{g_0}^H\mathbf{R^{-1}g_0} \tag{3.4}$$

with $\mathbf{R}$ defined in (3.3b). In the next section, we derive an exact analytical expression for the long term average of the output SINR in (3.4), i.e., the ensemble average of $\gamma(M)$ over the fading vectors $\mathbf{g_0}$ and $\mathbf{g_1}$.

3.2 Average Output SINR

Using the matrix inversion lemma [31, p. 50], the inverse of $\mathbf{R}$ in (3.3b) can be calculated as

$$\mathbf{R}^{-1} = \frac{1}{N_0}\left(\mathbf{I}_M - \frac{\mathbf{g_1}\mathbf{g_1}^H}{N_0/P_1(M) + \mathbf{g_1}^H\mathbf{g_1}}\right). \tag{3.5}$$

Thus, the expected value of γ in (3.4) is given by

$$\begin{aligned}
\mathrm{E}\{\gamma(M)\} &= P_0(M)\mathrm{E}\{\mathbf{g_0}^H\mathbf{R^{-1}}\mathbf{g_0}\} \\
&= P_0(M)\mathrm{E}\{\mathrm{tr}(\mathbf{g_0}\mathbf{g_0}^H\mathbf{R^{-1}})\} \\
&= P_0(M)\mathrm{tr}\left(\mathbf{R_a}\mathrm{E}\{\mathbf{R^{-1}}\}\right).
\end{aligned} \tag{3.6}$$

In deriving (3.6), we have made use of the fact that $\mathbf{g_0}$ is independent of $\mathbf{g_1}$ and, therefore, of $\mathbf{R}$. Inserting (3.5) in (3.6), after some manipulations one obtains

$$\mathrm{E}\{\gamma(M)\} = \bar{\gamma}_0(M)\left(M - \mathrm{E}\left\{\frac{\mathbf{g_1}^H\mathbf{R_a}\mathbf{g_1}}{\frac{\bar{\gamma}_1(M)}{\bar{\gamma}_0(M)} + \mathbf{g_1}^H\mathbf{g_1}}\right\}\right) \tag{3.7}$$

where we have used the fact that $\mathrm{tr}(\mathbf{R_a}) = M$. In (3.7), $\bar{\gamma}_0(M) = P_0(M)/N_0$ is the average input SNR per branch and $\bar{\gamma}_1(M) = P_0(M)/P_1(M)$ is the average input SIR per branch. Based on the EVD of $\mathbf{R_a}$ in (3.2), one can express $\mathbf{g_1}$ as

$$\mathbf{g_1} = \sqrt{M}\mathbf{Q}\mathbf{A}^{\frac{1}{2}}\mathbf{h} \tag{3.8}$$

where $\mathbf{h} = [h_1, h_2, \ldots, h_M]^T$ is a vector of i.i.d. zero-mean complex symmetric Gaussian random variables with correlation matrix $\mathrm{E}\{\mathbf{h}\mathbf{h}^H\} = \mathbf{I}_M$ and $(\cdot)^T$ is the matrix transpose operator. Substituting (3.2) and (3.8) in (3.7), we get

$$\mathrm{E}\{\gamma(M)\} = M\bar{\gamma}_0(M)\left(1 - \mathrm{E}\left\{\frac{X}{Y}\right\}\right) \tag{3.9a}$$

where

$$X = \sum_{i=1}^{M} \alpha_{i,M}^2 |h_i|^2 \tag{3.9b}$$

$$Y = \frac{\bar{\gamma}_1(M)}{M\bar{\gamma}_0(M)} + \sum_{i=1}^{M} \alpha_{i,M}|h_i|^2. \tag{3.9c}$$

The expected value of the ratio of X and Y in (3.9) can be calculated using the joint MGF of X and Y. In fact, if $\mathcal{M}_{X,Y}(s_1, s_2) = \mathrm{E}\{e^{s_1 X + s_2 Y}\}$ is the joint MGF of X and Y, it can be shown that [36]

$$\mathrm{E}\left\{\frac{X}{Y}\right\} = \int_0^\infty \mathcal{M}_{X,Y}^{(1,0)}(0, -s)ds \tag{3.10a}$$

where

$$\mathcal{M}_{X,Y}^{(1,0)}(s_1, s_2) = \frac{\partial}{\partial s_1}\mathcal{M}_{X,Y}(s_1, s_2). \tag{3.10b}$$

The random variables $|h_i|^2$, $i = 1, \ldots, M$, in (3.9) are i.i.d. unit variance chi-square RV's with two degrees of freedom. Therefore, the joint MGF of X and Y can be written as [37, p. 154]

$$\mathcal{M}_{X,Y}(s_1, s_2) = \exp\left(\frac{\bar{\gamma}_1(M)s_2}{M\bar{\gamma}_0(M)}\right)\prod_{i=1}^{M}\frac{1}{1 - \alpha_{i,M}^2 s_1 - \alpha_{i,M} s_2}. \tag{3.11}$$

Using (3.11) in (3.10), one obtains

$$\mathrm{E}\left\{\frac{X}{Y}\right\} = \int_0^\infty \mathcal{I}_M(s)ds \tag{3.12a}$$

where

$$\mathcal{I}_M(s) = \left(\prod_{i=1}^{M}\frac{1}{1+\alpha_{i,M}s}\right)\left(\sum_{i=1}^{M}\frac{\alpha_{i,M}^2}{1+\alpha_{i,M}s}\right)\exp\left(-\frac{\bar{\gamma}_1(M)}{M\bar{\gamma}_0(M)}s\right). \tag{3.12b}$$

Inserting (3.12) in (3.9) gives the final expression for the average output SINR. This expression is used in Section 3.3 to analyze the asymptotic behavior of the average output SINR as the number of antennas increases while the total physical size of the array is fixed.

3.3 Asymptotic Analysis of Average Output SINR

In this section, the asymptotic behavior of the $\mathrm{E}\{X/Y\}$ in (3.12) for large values of M is studied. In order to do so, we must first understand the asymptotic behavior of the eigenvalues, i.e., the $\alpha_{i,M}$'s as M goes to infinity.

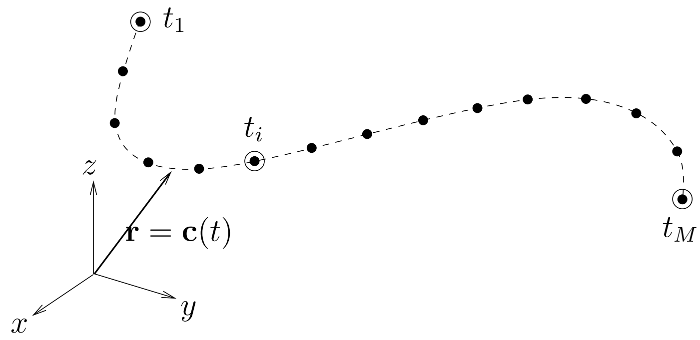


Figure 3.1. A one-dimensional array of M antennas located on a curve with parametric representation $\mathbf{r} = \mathbf{c}(t)$ in a 3D space.

3.3.1 Eigenvalues of Dense One-Dimensional Arrays

Let the 3×1 vector $\mathbf{c}(t) = [c_x(t), c_y(t), c_z(t)]^T$ denote the parametric representation of a one-dimensional (1D) curve in a three-dimensional (3D) space where $t \in [0, 1]$ (Fig. 3.1). Now, assume that an array of M isotropic antennas is distributed along $\mathbf{c}(t)$ where the position of the i'th antenna is given by $\mathbf{c}(t_i)$ for some $t_i \in [0, 1]$, $i = 1, \ldots, M$. The indexing of the antennas is such that $t_1 < t_2 < \ldots < t_M$. The correlation coefficient between the i'th and the j'th antennas in the array can be written as $R_g(\mathbf{c}(t_i), \mathbf{c}(t_j))$ where $R_g(\mathbf{r_1}, \mathbf{r_2}) = \mathrm{E}\{g(\mathbf{r_1})g(\mathbf{r_2})^*\}$ is the spatial autocorrelation function of the underlying channel gain $g(\mathbf{r})$ represented as a complex Gaussian spatial random process. The parameter $\mathbf{r}$ is the 3×1 position vector and $*$ denotes complex conjugate. Note that $R_g(\mathbf{r_1}, \mathbf{r_2})$ is independent of the array geometry and is only a function of the properties of the scattering environment. Define $R_a(x, y) \triangleq R_g(\mathbf{c}(x), \mathbf{c}(y))$ for $(x, y) \in [0, 1] \times [0, 1]$ where $\times$ denotes the Cartesian product. Then, for any given M, $R_a(t_i, t_j)$ is the (i, j)'th element of the array correlation matrix $\mathbf{R_a}$ defined in Section 3.1. The "array correlation function" $R_a(x, y)$ is normalized such that $R_a(x, x) = 1$ for any $x \in [0, 1]$. From its definition, $R_a(x, y)$ is Hermitian, i.e., $R_a(x, y) = R_a^*(y, x)$. It is also square summable over $[0, 1] \times [0, 1]$. In fact, since the magnitude of $R_a(x, y)$ is always less than or equal to 1, one has

$$\int_0^1 \int_0^1 |R_a(x, y)|^2 dx dy \leqslant 1. \tag{3.13}$$

Moreover, we assume that $R_a(x, y)$ is a continuous function over $[0, 1] \times [0, 1]$ as is the case in a practical system. Let $\mathbf{R_a q}_k = M\alpha_{k,M}\mathbf{q}_k$, $k = 1, \ldots, M$, represent the eigenequations of the matrix $\mathbf{R_a}/M$ and define the functions $q_k(t)$, $k = 1, \ldots, M$, on $[0, 1]$ such that $q_k(t_i) = \mathbf{q}_k(i)$, where $\mathbf{q}_k(i)$ denotes the i'th element of $\mathbf{q}_k$. Then, the eigenequations of $\mathbf{R_a}/M$ can be expanded as

$$\frac{1}{M}\sum_{j=1}^{M} R_a(t_i, t_j)q_k(t_j) = \alpha_{k,M}q_k(t_i) \qquad k = 1, \ldots, M. \tag{3.14}$$

If the antennas are uniformly distributed in the parameter space, that is $t_i = (i-1)/M$ for $i = 1, \ldots, M$, the set of nonzero eigenvalues of $\mathbf{R_a}/M$ (i.e., the set of nonzero $\alpha_{k,M}$'s) converges to the point spectrum (i.e., eigenvalues in this case) of a nonnegative definite self-adjoint integral operator with kernel $R_a(x, y)$ as M goes to infinity [35], [38, p. 248]. In the sequel, whenever we need to refer to "the integral operator whose kernel is $R_a(x, y)$", we will simply refer to the function $R_a(x, y)$ itself. The sorted eigenvalues of $R_a(x, y)$ are denoted by $\tilde{\alpha}_1 \geqslant \tilde{\alpha}_2 \geqslant \ldots$ and are determined through the set of integral equations

$$\int_0^1 R_a(x, y)q_k(y)dy = \tilde{\alpha}_k q_k(x) \qquad k = 1, 2, \ldots \tag{3.15}$$

where the $q_k(x)$'s are the orthonormal eigenfunctions of $R_a(x, y)$. The $\tilde{\alpha}_k$'s are all positive and less than or equal to one. Furthermore, they are either finite or countably infinite in number. In the latter case, they form a decreasing sequence tending to zero [35], [39]. Since $R_a(x, y)$ is Hermitian and square summable, it can be expanded in terms of its eigenfunctions as [39, p. 243]

$$R_a(x, y) = \sum_{k=1}^{\infty} \tilde{\alpha}_k q_k(x)q_k^*(y). \tag{3.16}$$

Moreover, according to Mercer's theorem [39, p. 245] the series in (3.16) is uniformly convergent since $R_a(x, y)$ is nonnegative definite and continuous. This justifies term by term integration in (3.16). As a result, using the expansion in (3.16), one can show

$$\sum_{k=1}^{\infty} \tilde{\alpha}_k = \int_0^1 R_a(x, x)dx = 1 \tag{3.17}$$

$$\sum_{k=1}^{\infty} \tilde{\alpha}_k^2 = \int_0^1 \int_0^1 |R_a(x,y)|^2 dxdy. \tag{3.18}$$

The convergence of the set of nonzero $\alpha_{k,M}$'s to their limit values $\tilde{\alpha}_k$'s is in fact uniform and there exists a uniform error bound of the form $|\alpha_{k,M} - \tilde{\alpha}_k| < C/M$ for any $k = 1, \ldots, M$ where C is some positive constant independent of k [35]. Let $\mathcal{N}_a(M)$ denote the number of nonzero eigenvalues of $\mathbf{R_a}/M$. As stated in [35], $\mathcal{N}_a(M)$ is $o(M)$, i.e., $\lim_{M\to\infty} \mathcal{N}_a(M)/M = 0$. Using this result along with the uniform convergence of the eigenvalues, the asymptotic behavior of $\mathrm{E}\{X/Y\}$ in (3.12) is analyzed in the sequel.

3.3.2 Asymptotic Analysis of $\mathrm{E}\{X/Y\}$

The limit of $\mathrm{E}\{X/Y\}$ in (3.12) as M tends to infinity is calculated in two steps. First, it is shown that under certain conditions on the correlation function $R_a(x,y)$, the integrand in (3.12) converges to a finite function $\mathcal{I}(s)$ for any value of $s \in [0,\infty)$ where $\mathcal{I}(s)$ can be expressed in terms of the $\tilde{\alpha}_i$'s. Then, it is shown that the limit of the integral in (3.12) is equal to the integral of the limit of the integrand, that is, the integral of $\mathcal{I}(s)$. We start with the following lemma.

Lemma 3.1. *Let $f : [0,1] \to [0,\infty)$ be a differentiable function with a bounded derivative, that is $|f'(x)| < B$ for any $x \in [0,1]$ and for some positive number B. Then, one has*

$$\lim_{M\to\infty} \sum_{i=1}^{\mathcal{N}_a(M)} [f(\alpha_{i,M}) - f(\tilde{\alpha}_i)] = 0 \tag{3.19}$$

where $\mathcal{N}_a(M)$, $\alpha_{i,M}$'s, and $\tilde{\alpha}_i$'s are defined in Section 3.3.1.

The proof of this lemma is not straightforward since the number of terms in the summation, i.e., $\mathcal{N}_a(M)$ can, in general, grow with M. A proof of Lemma 3.1 is given in Appendix A. The integrand in (3.12b) can be rewritten as follows

$$\mathcal{I}_M(s) = \exp\left(-\sum_{i=1}^{\mathcal{N}_a(M)} \pi_s(\alpha_{i,M})\right)\left(\sum_{i=1}^{\mathcal{N}_a(M)} \sigma_s(\alpha_{i,M})\right)\exp\left(-\frac{\bar{\gamma}_1(M)}{M\bar{\gamma}_0(M)}s\right) \tag{3.20a}$$

where

$$\pi_s(x) = \ln(1 + sx) \tag{3.20b}$$

$$\sigma_s(x) = \frac{x^2}{1 + sx}. \tag{3.20c}$$

Note that the summations in (3.20a) are only taken over the $\mathcal{N}_a(M)$ nonzero eigenvalues since the zero eigenvalues have no effect on the value of $\mathrm{E}\{X/Y\}$ in (3.12). It is easy to see that $\pi_s(x)$ and $\sigma_s(x)$ in (3.20b) and (3.20c) satisfy the conditions of Lemma 3.1 with $B = s$ and $B = 2$, respectively. Thus, Lemma 3.1 applies to them. Moreover, as shown in Appendix B, the limits $\lim_{M\to\infty}\sum_{i=1}^{\mathcal{N}_a(M)} \pi_s(\tilde{\alpha}_i)$ and $\lim_{M\to\infty}\sum_{i=1}^{\mathcal{N}_a(M)} \sigma_s(\tilde{\alpha}_i)$ exist. Since these limits exist and Lemma 3.1 applies to both $\pi_s(x)$ and $\sigma_s(x)$, one has

$$\lim_{M\to\infty} \sum_{i=1}^{\mathcal{N}_a(M)} \pi_s(\alpha_{i,M}) = \sum_{i=1}^{\infty} \pi_s(\tilde{\alpha}_i) \tag{3.21}$$

and

$$\lim_{M\to\infty} \sum_{i=1}^{\mathcal{N}_a(M)} \sigma_s(\alpha_{i,M}) = \sum_{i=1}^{\infty} \sigma_s(\tilde{\alpha}_i). \tag{3.22}$$

This means that the sequence of functions $\{\mathcal{I}_M(s)\}_{M=1}^{\infty}$ in (3.20a) converges to a function $\mathcal{I}(s)$ for any $s \in [0, \infty)$ where $\mathcal{I}(s)$ is given by

$$\mathcal{I}(s) = \exp\left(-\sum_{i=1}^{\infty} \pi_s(\tilde{\alpha}_i)\right)\left(\sum_{i=1}^{\infty} \sigma_s(\tilde{\alpha}_i)\right)\exp(-\beta s) \tag{3.23a}$$

and where

$$\beta = \lim_{M\to\infty} \frac{\bar{\gamma}_1(M)}{M\bar{\gamma}_0(M)} \tag{3.23b}$$

with $\pi_s(x)$ and $\sigma_s(x)$ defined in (3.20b) and (3.20c), respectively. The next step is to show that the limit of the integral in (3.12a) as M tends to infinity is equal to the integral of $\mathcal{I}(s)$ defined in (3.23). This is the subject of the following theorem.

Theorem 3.2.

$$\lim_{M\to\infty} \int_0^{\infty} \mathcal{I}_M(s)ds = \int_0^{\infty} \mathcal{I}(s)ds. \tag{3.24}$$

The proof of Theorem 3.2 is given in Appendix C. Using Theorem 3.2 and the expression for $\mathcal{I}(s)$ in (3.23), the asymptotic value of $\mathrm{E}\{X/Y\}$ is obtained as

$$\lim_{M\to\infty} \mathrm{E}\left\{\frac{X}{Y}\right\} = \int_0^\infty \left(\prod_{i=1}^{\infty} \frac{1}{1+\tilde{\alpha}_i s}\right)\left(\sum_{i=1}^{\infty} \frac{\tilde{\alpha}_i^2}{1+\tilde{\alpha}_i s}\right)\exp(-\beta s)ds \qquad (3.25)$$

where β is defined in (3.23b). As shown in (3.9), the average output SINR is proportional to $1-\mathrm{E}\{X/Y\}$. Since the SINR is a positive random variable, one can immediately conclude that $\mathrm{E}\{X/Y\} < 1$ and, therefore, $\lim_{M\to\infty} \mathrm{E}\{X/Y\}$ given in (3.25) is always less than or equal to one. However, in order to fully understand the asymptotic behavior of the average SINR in (3.9), it is important to determine the conditions under which $\lim_{M\to\infty} \mathrm{E}\{X/Y\}$ is strictly less than one. This is particularly important in the fixed power per branch scenario where the factor $M\bar{\gamma}_0(M) = MP_0/N_0$ in (3.9) grows linearly with M. The next theorem provides a sufficient condition.

Theorem 3.3. *Assume that $R_a(x,y)$ is nonseparable, i.e., one cannot find any function $k(\cdot)$ such that $R_a(x,y)$ can be written as $R_a(x,y) = k(x)k^*(y)$. Then, one has*

$$\lim_{M\to\infty} \mathrm{E}\left\{\frac{X}{Y}\right\} < 1. \qquad (3.26)$$

The proof of Theorem 3.3 is given in Appendix D. At this point, we are ready to state our main result. Based on the above analysis of the asymptotic behavior of $\mathrm{E}\{X/Y\}$, the asymptotic performance of a fixed length one-dimensional array of antennas as the number of antennas goes to infinity is summarized through Theorem 3.4.

Theorem 3.4. *Consider an optimum combining receiver employing the array of antennas introduced in Section 3.3.1. If the underlying array correlation function $R_a(x,y)$ defined in Section 3.3.1 is continuous and nonseparable (see Theorem 3.3), the asymptotic performance of the receiver is as follows.*

- *Fixed average power per branch: The average output SINR for large* *values of M grows linearly with M with a positive slope given by*

$$\lim_{M\to\infty} \frac{\mathrm{E}\{\gamma(M)\}}{M} = \Gamma_0 \left\{ 1 - \int_0^\infty \left(\prod_{i=1}^{\infty} \frac{1}{1+\tilde{\alpha}_i s} \right) \left(\sum_{i=1}^{\infty} \frac{\tilde{\alpha}_i^2}{1+\tilde{\alpha}_i s} \right) ds \right\} \tag{3.27}$$

where $\Gamma_0 = P_0/N_0$ is the average input SNR per branch and the $\tilde{\alpha}_i$'s are defined in Section 3.3.1.

- *Fixed total average power: As M increases, the average output SINR converges to a positive limit value given by*

$$\lim_{M\to\infty} \mathrm{E}\{\gamma(M)\} = \Gamma_0^t \left\{ 1 - \int_0^\infty \left(\prod_{i=1}^{\infty} \frac{1}{1+\tilde{\alpha}_i s} \right) \left(\sum_{i=1}^{\infty} \frac{\tilde{\alpha}_i^2}{1+\tilde{\alpha}_i s} \right) \exp\left(-\frac{\Gamma_1^t}{\Gamma_0^t} s \right) ds \right\} \tag{3.28}$$

where $\Gamma_0^t = P_0^t/N_0$ is the total average input SNR, $\Gamma_1^t = P_0^t/P_1^t$ is the total average input SIR, and the $\tilde{\alpha}_i$'s are defined in Section 3.3.1.

Proof. The proof is immediate using Theorem 3.3, (3.9), (3.25), (3.23b), and noting that $\bar{\gamma}_0(M) = P_0/N_0$, and $\bar{\gamma}_1(M) = P_0/P_1$ for the fixed average power per branch scenario whereas $\bar{\gamma}_0(M) = P_0^t/(MN_0)$, and $\bar{\gamma}_1(M) = (P_0^t/M)/(P_1^t/M) = P_0^t/P_1^t$ for the fixed total average power scenario. □

The identities in (3.27) and (3.28) are valid even for the case of a separable array correlation function $R_a(x, y)$. However, in this case, one can show that the limit in (3.25) is equal to unity. This is because, for a separable correlation function, one has $\tilde{\alpha}_1 = 1$ and $\tilde{\alpha}_i = 0$, for all $i > 1$ (see the proof of Theorem 3.3 for more details). Therefore, although the identity in (3.27) is still valid, it does not imply a linear growth of the average output SINR with M since the expression on the right side of (3.27) will evaluate to zero. The main contribution of Theorem 3.4 lies in the fact that it provides a mathematical proof of the asymptotic behavior of OC for two different power scenarios and for a broad subset of correlation functions. As a second result, Theorem 3.4 also provides the exact numerical values of the asymptotes of the SINR for both power scenarios. These numerical values depend on the eigenvalues of the underlying array correlation function determined by (3.15). Solving the eigenvalue problem in (3.15), in general, lies in the realm of functional (operator) analysis and spectral theory [39]. There are many numerical methods in the mathematics literature for calculation of the eigenvalues of an integral operator (e.g., see [38], [40], and the references therein). However, in some special cases, the abovementioned eigenvalue problem has a nice closed-form solution. One example is when $R_a(x, y)$ in (3.15) is a circulant kernel, i.e., when it can be written as $R_a(x, y) = \tilde{R}_a(z)$ where $z = x - y$ and $\tilde{R}_a(z)$ is some periodic function with period 1. A practical scenario resulting in a circulant array correlation function is investigated in Section 3.4.

3.4 Numerical Results and Discussion

The analytical results presented in Section 3.3, and particularly Theorem 3.4, deal with asymptotic scenarios in which the number of antennas goes to infinity. However, as shown through a numerical example in this section, the number of antennas that is required in order to put a practical receiver in the asymptotic regime is relatively small. Consider a uniform circular array of isotropic antennas located in the horizontal plane. Assume a scattering environment, where the azimuth angle of arrival is uniformly distributed over $[0, 2\pi)$ while the elevation angle of arrival has an arbitrary distribution. In this case, the spatial correlation function $R_g(\mathbf{r_1}, \mathbf{r_2})$ defined in Section 3.3.1 is only a function of the distance between $\mathbf{r_1}$ and $\mathbf{r_2}$ as a result of the rotational symmetry in the horizontal plane. The parametric representation of the array geometry defined in Section 3.3.1 is given by $\mathbf{c}(t) = [r\cos(2\pi t), r\sin(2\pi t), 0]$, $t \in [0, 1]$ where r is the array radius. Then, one can show that the array correlation function $R_a(x, y)$ defined in Section 3.3.1 is only a function of $z = x - y$ and can be written as $R_a(x, y) = \tilde{R}_a(z)$ for some function $\tilde{R}_a(z)$ where $z \in [-1, 1]$. Moreover, due to the circular structure of the array, $\tilde{R}_a(z)$ is a periodic function of z with period 1. Thus, using (3.15), it can be shown that the set of eigenfunctions $\{q_k(x)\}_{k=1}^{\infty}$ is given by $\{\exp(j2\pi kx)\}_{k=-\infty}^{\infty}$ and that the set of eigenvalues $\{\tilde{\alpha}_k\}_{k=1}^{\infty}$ is equal to the set of Fourier series co-

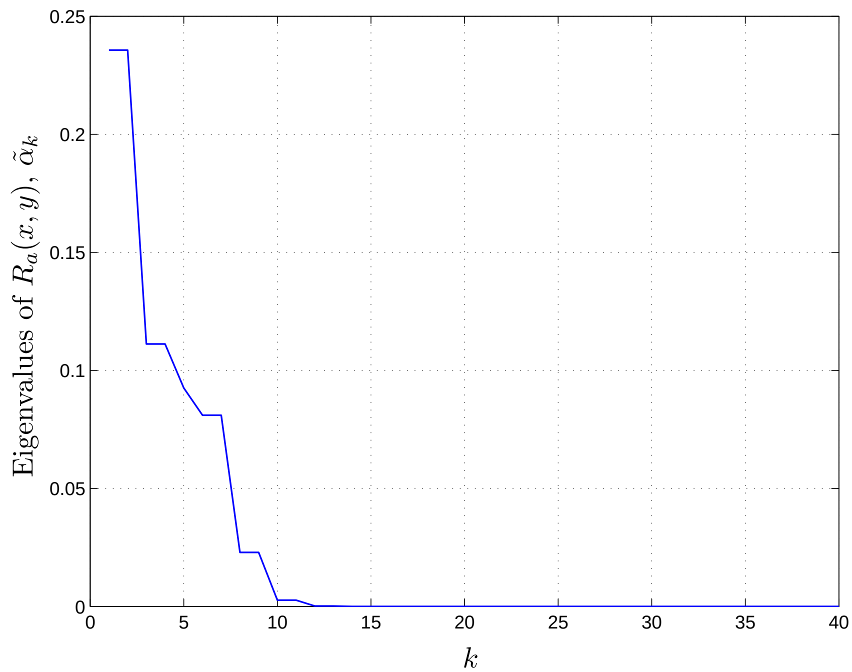


Figure 3.2. The sorted Fourier series coefficients of $\tilde{R}_a(z) = \mathrm{J}_0(\frac{4\pi}{\lambda} r\sin(\pi|z|))$ (i.e., the eigenvalues of $R_a(x,y)$) for $r = 0.5\lambda$.

efficients of $\tilde{R}_a(z)$, that is $\{\int_0^1 \tilde{R}_a(z)\exp(-j2\pi kz)dz\}_{k=-\infty}^{\infty}$. In the case of a 2D-omnidirectional scattering where the distribution of the elevation angle of arrival is modeled as a delta function in the horizontal plane, the spatial correlation function is given by $R_g(\mathbf{r_1}, \mathbf{r_2}) = \mathrm{J}_0(\frac{2\pi}{\lambda}\|\mathbf{r_1} - \mathbf{r_2}\|)$ where λ is the carrier wavelength [5]. Thus, the array correlation function can be written as $R_a(x,y) = \tilde{R}_a(z) = \mathrm{J}_0(\frac{4\pi}{\lambda} r\sin(\pi|z|))$ where $z = x - y$. It can be easily shown that $R_a(x,y)$ in this case is a nonseparable function of x and y as defined in Theorem 3.3. Furthermore, the set of Fourier series coefficients of $\tilde{R}_a(z)$ is given by $\{\mathrm{J}_k^2(\frac{2\pi}{\lambda} r)\}_{k=-\infty}^{\infty}$ where $\mathrm{J}_k(\cdot)$ is the Bessel function of the first kind and order k [26]. Fig. 3.2 shows the first 40 sorted Fourier coefficients of $\tilde{R}_a(z)$ (i.e., $\tilde{\alpha}_1$ to $\tilde{\alpha}_{40}$) for $r = 0.5\lambda$. Figs. 3.3 and 3.4 depict the average output SINR versus the number of antennas M for the fixed power per branch scenario and the fixed total power scenario, respectively. The analytical values of the average output SINR for different values of M are calculated using (3.9) and (3.12). The asymptotic values are computed from (3.27) and (3.28). As seen in Fig. 3.2, the sequence of eigenvalues rapidly converges to zero. As a re-

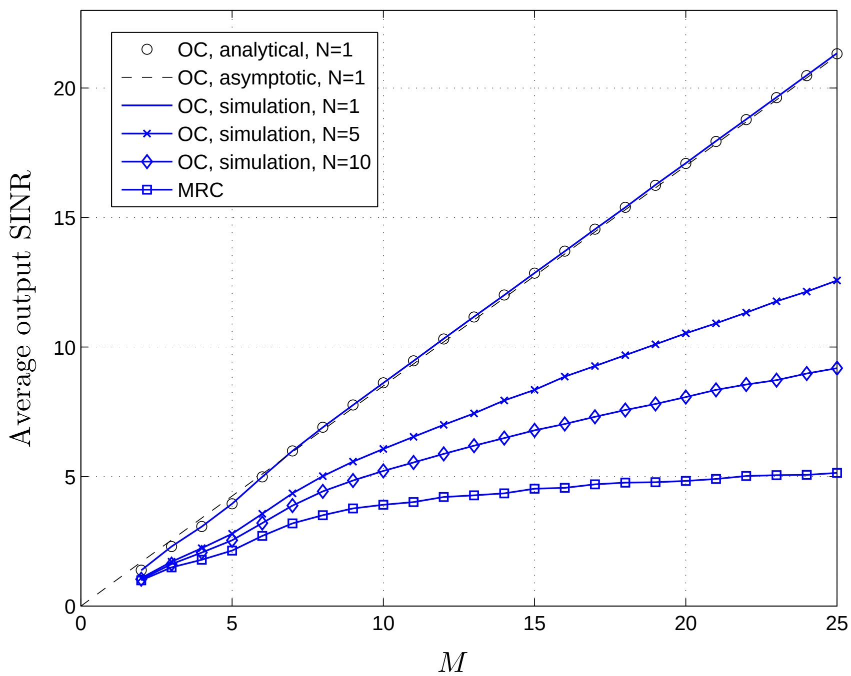


Figure 3.3. The average output SINR of OC and MRC (in linear scale) for a circular array of radius $r = 0.5\lambda$ in the presence of N interfering users under Rayleigh fading and 2D-omnidirectional scattering. The average input SNR and SIR per branch are fixed as M increases and are given by $\Gamma_0 = 0$ dB and $\Gamma_1 = 0$ dB, respectively.

sult, in calculation of (3.27) and (3.28), only the first 15 dominant eigenvalues are used. As seen in Figs. 3.3 and 3.4, the analytical results are in excellent agreement with the simulation results.

The average output SINRs of an OC receiver in the presence of $N = 5$ and $N = 10$ interfering users obtained by simulation are also depicted in Figs. 3.3 and 3.4 where, in both scenarios, the sum of the powers of the N interfering users is kept fixed and equal to that of the single interfering user case. It can be seen that even in the case of more than one interferer, the OC receiver exhibits the same asymptotic behavior. In other words, as shown in Fig. 3.3 when the average power per branch is fixed the average output SINR of OC asymptotically becomes a linear function of M; however, the line's slope decreases as N becomes larger. On the other hand, as seen in Fig. 3.4 when the total power is fixed, the average output SINR of OC quickly reaches a limit as M increases. It is also seen that the asymptotic limiting performance decreases as N increases. Note that the SINR curves in Fig. 3.3 are plotted in

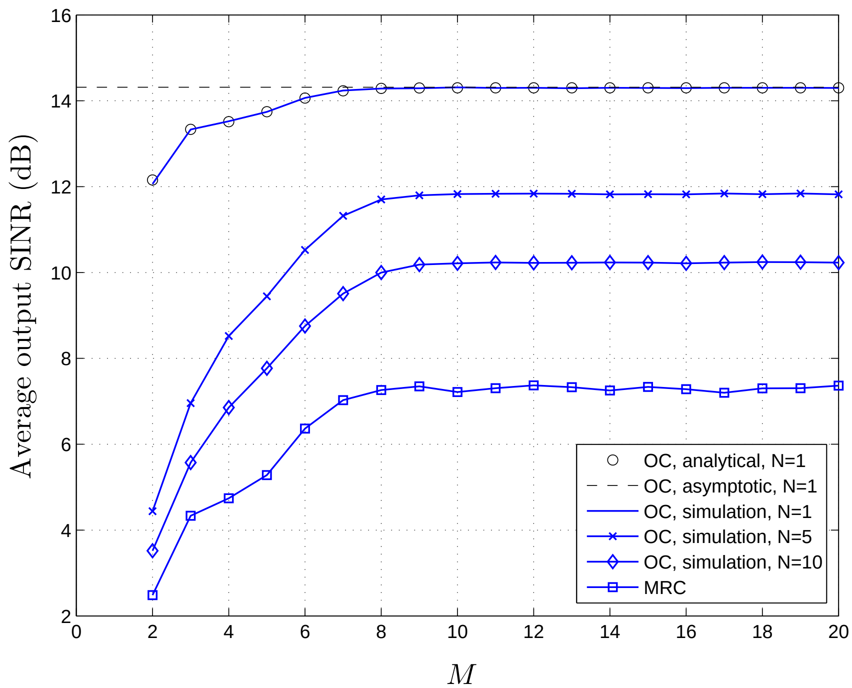


Figure 3.4. The average output SINR of OC and MRC for a circular array of radius $r = 0.5\lambda$ in the presence of N interfering users under Rayleigh fading and 2D-omnidirectional scattering. The total average input SNR and SIR are fixed as M increases and are given by $\Gamma_0^t = 15$ dB and $\Gamma_1^t = 0$ dB, respectively.

a linear scale in order to illustrate the linear asymptotic behavior of OC. The results in Figs. 3.3 and 3.4 indicate that, in all cases, the OC receiver enters the asymptotic regime for a relatively small number of antennas, here being $M = 9$. This confirms the applicability of the asymptotic analysis provided in the previous section to practical scenarios.

The SINR performance of an MRC receiver for both power constraint scenarios is also illustrated in Figs. 3.3 and 3.4 for comparison purposes. In contrast to the OC receiver, as long as the sum of the powers of the interfering users is fixed, the performance of MRC does not change with the number of interfering users N. This is because an MRC receiver does not distinguish the interference from the AWGN. In the case of fixed power per branch, the average SINR of MRC in the asymptotic region seems to slightly increase with M. However, other simulation results (not shown here) as well as the approximate analytical results of [24], [25] suggest, that in both power scenarios, the average output SINR of MRC eventually saturates as M increases. From

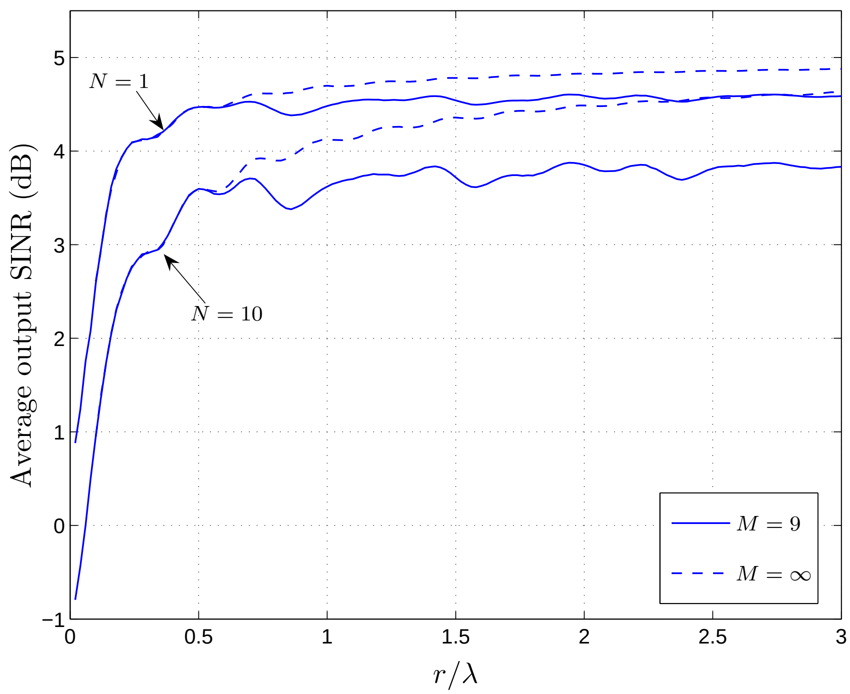


Figure 3.5. The average output SINR of OC for a circular array versus the normalized array radius r/λ in the presence of N interfering users under Rayleigh fading and 2D-omnidirectional scattering. The total average input SNR and SIR are fixed for all values of M and r/λ and are given by $\Gamma_0^t = 5$ dB and $\Gamma_1^t = 0$ dB, respectively.

Figs. 3.3 and 3.4, one can see the dramatic superiority of OC over MRC in an interference dominated environment. Particularly, in the case of fixed total power, an SINR gain of 7 dB, 4.5 dB, and 3 dB is achieved by OC over MRC for $N = 1$, $N = 5$, and $N = 10$ interfering users, respectively.

Next, the effect of the array radius on the output SINR performance of OC for the above circular array is illustrated in Fig. 3.5 where the average output SINR is plotted versus the normalized array radius r/λ while the total average SNR and SIR input to the array are kept fixed for all values of M and r. Two different cases for the number of interfering users have been considered, namely, $N = 1$ and $N = 10$. The dashed lines show the asymptotic average output SINR obtained by letting M go to infinity for a given value of r/λ, i.e., the best achievable performance for each value of r/λ. The solid lines show the average output SINR versus r/λ where the number of antennas in the array is fixed and set to $M = 9$. In the case of $N = 1$, the solid curve is computed from (3.9) and (3.12) and the dashed curve from (3.28). In the case of $N = 10$, both curves are obtained by simulation. It can be seen in Fig. 3.5 that for values of r/λ less than 0.5 the dashed curves and the solid curves coincide. This means that for $r/\lambda < 0.5$ increasing the number of antennas beyond $M = 9$ does not provide any additional performance benefit. This is in agreement with the results of Fig. 3.4 where it was shown that for $r = 0.5\lambda$ the array's performance saturates at $M = 9$. However, for the values of r/λ greater than 0.5, a performance loss will be incurred by limiting the number of antennas to $M = 9$. Moreover, as Fig. 3.5 shows, this performance loss increases as r/λ grows. Particularly, at $r/\lambda = 3$, using $M = 9$ antennas results in a performance loss of 0.3 dB for $N = 1$ and that of 0.8 dB for $N = 10$ compared to the best achievable SINR performance.

3.5 Chapter Summary

In this chapter, the asymptotic performance of OC was investigated as the number of antennas in the array increases while the total physical dimensions of the array are fixed. Two different scenarios were considered, namely, fixed average received power per antenna and fixed total average received power. It was shown that in the former scenario, the average output SINR of OC is asymptotically a linear function of the number of the antennas while in the latter scenario it eventually saturates at a certain value. In the case of a single interferer, the slope of the asymptote in the former scenario as well as the value of the saturation limit in the latter scenario were derived in terms of the point spectrum of the underlying array correlation function. It was shown through a numerical example that the receiver actually exhibits its asymptotic behavior for a practically small number of antennas. This justifies the applicability of the asymptotic approach used in this chapter to the design of OC receivers implemented on size constrained communication devices.

In the next chapter, we turn our focus to the optimum design of size constrained antenna arrays used in OC and/or MRC receivers. To this goal, we find a simple yet accurate approximation to the average output SINR of OC derived in this chapter. Then, we attempt to maximize the value of this approximate expression over the number and the positions of the antennas in the array. We will start with the case of linear 1D arrays and consequently generalize our formulation to the case of 2D arrays.

Chapter 4

Optimum Antenna Arrangement

The focus of our research work presented up to this point has been mainly on the *performance analysis* of multiple antenna receivers in the presence of correlated fading and CCI. In this chapter[1], however, we propose an *array design* problem for multiple antenna receivers based on the results of the previous chapters on the performance analysis of those systems.

As we saw in Chapter 2 and Chapter 3, the average BER performance of MRC and the average SINR performance of OC for a correlated array of antennas are functions of the eigenvalues of the array correlation matrix (see Section 2.2 and Section 3.2). The array correlation matrix itself is a function of the underlying spatial correlation function (see Section 3.3.1) as well as the relative positions of the antennas in the array. The former, i.e., the spatial correlation function is out of the designer's control and is solely determined from the statistics of the scattering environment. To be more precise, it is obtained through the distribution of the direction of arrival (DOA) of the electromagnetic waves received by the array [5]. The latter, however, provides the designer with a degree of freedom that can be efficiently adjusted. In other words, one can properly arrange the antennas in the array such that a certain performance measure of the system is optimized. This is the subject of the research work presented in this chapter.

[1] *A version of this chapter has been submitted for publication in IEEE Transactions on Information Theory and has been published in part in Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), 2011, pp. 1402-1407.*

We adopt the average output SINR of OC derived in Chapter 3 as a measure of the array's performance and attempt to maximize it by properly arranging the antennas while assuming that the total physical dimensions of the array are fixed. We will also show that maximizing the average output SINR of OC is equivalent to maximizing two of the performance measures defined in [25] for MRC in the presence of correlated Rayleigh fading, CCI, and AWGN.

4.1 System Model

The vector of complex equivalent baseband signals received by an array of M antennas in the presence of cochannel interference and noise is given by

$$\mathbf{y} = \sqrt{\frac{P_0}{M}}\mathbf{g_0}d_0 + \sum_{k=1}^{N}\sqrt{\frac{P_k}{M}}\mathbf{g_k}d_k + \mathbf{n} \tag{4.1}$$

where the first term on the right of (4.1) represents the received signal of the desired user, the second term represents the interference, and the third term represents additive noise. The number of interfering users is N, the d_k's ($k = 0, \ldots, N$) are i.i.d. transmitted data symbols of different users each chosen from a signal constellation of unit average energy, and the $\mathbf{g_k}$'s ($k = 0, \ldots, N$) are independent M by 1 zero-mean complex symmetric Gaussian random vectors representing the Rayleigh channel between each user and the array elements in the receiver. The average received power per antenna for User k is denoted by P_k/M. The total average power P_k of User k input to the array is fixed as a result of the limitation on the total physical dimensions of the array. The vector $\mathbf{n}$ denotes AWGN whose elements are i.i.d. each with variance N_0. The average SNR of the desired user is defined by $\Gamma_0 = P_0/N_0$ and the average interference to noise ratios (INRs) of the interferers by $\Gamma_k = P_k/N_0$ ($k = 1, \ldots, N$).

For each $k = 0, \ldots, N$, the elements of the channel vector $\mathbf{g_k}$ are considered to be the samples of the zero-mean Gaussian spatial random process (random field) $g_k(\mathbf{r})$ where $\mathbf{r}$ is the position vector in a given coordinate system. The $g_k(\mathbf{r})$'s are wide sense stationary (WSS) random processes that are independent of each other and all have a common autocovariance (in this case, autocorrelation) function defined by $R_g(\mathbf{r}_1 - \mathbf{r}_2) = \mathrm{E}\{g_k(\mathbf{r}_1)g_k(\mathbf{r}_2)^*\}$. We refer to $R_g(\cdot)$ as the "spatial correlation function". The spatial correlation function is conjugate symmetric, i.e., $R_g(-\mathbf{r}) = R_g(\mathbf{r})^*$. Thus, it has a real value at $\mathbf{r} = 0$. We normalize its value at zero to unity, i.e., $R_g(0) = 1$. The function $R_g(\cdot)$ is also a positive definite function (a.k.a a function of positive type) [41]. Therefore, its magnitude for any value of $\mathbf{r}$ is always less than or equal to $R_g(0) = 1$.

Two scenarios regarding the arrangement of the antennas are considered. In the first scenario, we assume a 1D linear array where the antennas are located on a line segment. The length of the line segment is denoted by L and the position of the i'th antenna, $i = 1, 2, \ldots, M$, by $\mathbf{r}_i = x_i \in [0, L]$ where, without loss of generality, we assume that $0 \leq x_1 \leq x_2 \leq \ldots \leq x_M \leq L$. In this case, the spatial correlation function is a single variable function denoted by $R_g(x)$. In the second scenario, we assume a 2D array where the antennas are located inside a bounded two-dimensional (2D) region $\mathcal{D}$ in $\mathbb{R}^2$ where $\mathbb{R}$ is the set of real numbers. The region $\mathcal{D}$ can, in general, consist of a finite number of disjoint connected subregions. The boundary of $\mathcal{D}$, denoted by $\partial\mathcal{D}$, is assumed to be a smooth simple curve (i.e., it does not intersect itself). The position of the i'th antenna (for some given ordering of the antennas) is denoted by $\mathbf{r}_i = (x_i, y_i) \in \mathcal{D}$. The spatial correlation function, in this case, is a two-variable function denoted by $R_g(x, y)$. The dimensions of length and area in both scenarios and in all subsequent analysis are electromagnetic dimensions, i.e., the unit of length is defined to be equal to one wavelength of the electromagnetic wave in the free space at the carrier frequency.

Define the common correlation matrix of the channel vectors $\mathbf{g_k}$ by $\mathbf{R_a} = \mathrm{E}\{\mathbf{g_k}\mathbf{g_k}^H\}$. We refer to $\mathbf{R_a}$ as the "array correlation matrix". In both scenarios, the (i, j)'th element of $\mathbf{R_a}$ is given by $R_g(\mathbf{r}_i - \mathbf{r}_j)$. In particular, the diagonal elements of $\mathbf{R_a}$ are all equal to unity since $R_g(0) = 1$. Note that $R_g(\cdot)$ is independent of the array geometry and is only determined by the statistics of the scattering environment around the receiver [5] whereas the array correlation matrix $\mathbf{R_a}$ depends on both. Let $\alpha_{1,M}, \alpha_{2,M}, \ldots, \alpha_{M,M}$ denote the eigenvalues of $\mathbf{R_a}/M$. The eigenvalues are all real and nonnegative since $\mathbf{R_a}$

is a Hermitian positive semi-definite matrix. Moreover,

$$\sum_{i=1}^{M} \alpha_{i,M} = \mathrm{tr}(\frac{1}{M}\mathbf{R_a}) = 1 \tag{4.2a}$$

and

$$\sum_{i=1}^{M} \alpha_{i,M}^2 = \mathrm{tr}(\frac{1}{M^2}\mathbf{R_a}^2). \tag{4.2b}$$

As a result of (4.2a), the $\alpha_{i,M}$'s are all less than or equal to one. Without loss of generality, we assume that the eigenvalues are sorted in descending order, i.e,

$$1 \geqslant \alpha_{1,M} \geqslant \alpha_{2,M} \geqslant \ldots \geqslant \alpha_{M,M} \geqslant 0. \tag{4.3}$$

The vector of received signals $\mathbf{y}$ in (4.1) is combined according to the OC scheme to obtain the decision variable $\mathbf{w}^H\mathbf{y}$ based on which the desired user's signal will be detected. The vector of combining coefficients is given by [3]

$$\mathbf{w} = \mathbf{R^{-1}g_0} \tag{4.4a}$$

where

$$\mathbf{R} = N_0\mathbf{I}_M + \sum_{k=1}^{N} \frac{P_k}{M}\mathbf{g_k g_k}^H. \tag{4.4b}$$

The SINR at the output of the OC receiver is given by [2, p. 683]

$$\gamma = \frac{P_0}{M}\mathbf{g_0}^H\mathbf{R^{-1}g_0} \tag{4.5}$$

where $\mathbf{R}$ is defined in (4.4b).

Note that the signal model in (4.1) also represents a single user MIMO channel. This can be seen by rewriting (4.1) as

$$\mathbf{y} = \frac{1}{\sqrt{M}}\mathbf{G}\mathbf{P}^{\frac{1}{2}}\mathbf{d} + \mathbf{n} \tag{4.6}$$

where $\mathbf{G} = [\mathbf{g_0}\ \mathbf{g_1}\ \ldots\ \mathbf{g_N}]$ is the $M \times (N+1)$ MIMO channel matrix, $\mathbf{d} = [d_0\ d_1\ \ldots\ d_N]^T$ is the transmitter's signal vector, and $\mathbf{P} = diag(P_0, P_1, \ldots, P_N)$ is a diagonal matrix containing the average powers transmitted by the antennas in the transmitter. The model in (4.6) corresponds to a MIMO system operating in the spatial multiplexing mode where $N+1$ independent data streams are transmitted to the receiver. The columns of $\mathbf{G}$ are independent, modeling a transmitter array at the base station with a large spatial separation, whereas the rows of $\mathbf{G}$ are correlated due to the limitation on the dimensions of the mobile device. It can be shown that the vector MMSE receiver of $\mathbf{d}$ in (4.6) corresponds to $N+1$ OC receivers, one for each individual element d_k of $\mathbf{d}$ [42]. Therefore, the SINR of each data stream at the output of the MIMO MMSE receiver is given by an expression similar to (4.5). In particular, (4.5) represents the SINR of data Stream 0. In the sequel, we will only focus on OC bearing in mind that all subsequent analyses are equally applicable to the above mentioned MIMO case as well.

4.2 Array Design

In this section, we study the effect of the number of antennas as well as their positions in the receiver array on the performance of the system defined in Section 4.1. In particular, we are interested in answering the following questions. What is the maximum number of antennas that can be usefully deployed within a physical space of fixed dimensions? What is the best arrangement of those antennas in the given space? We consider the average over fading of the output SINR in (4.5) as a measure of the system's performance and attempt to maximize it over all possible arrangements of the antennas in the array. This performance measure allows to formulate the problem in the form of a convex quadratic programming which can always be solved numerically using the well known methods of convex optimization [43]. Moreover, in the special case of an exponential correlation model for linear arrays, we are able to find an analytical solution to this optimization problem.

4.2.1 The Objective Function

As shown in Chapter 3, the average of the output SINR in (4.5) over the fading channel vectors $\mathbf{g_0}$ and $\mathbf{g_1}$, in the case of one interfering user ($N=1$), can be expressed as

$$\mathrm{E}\{\gamma\} = \Gamma_0(1-q) \tag{4.7a}$$

where

$$q = \mathrm{E}\left\{ \frac{\sum_{i=1}^{M} \alpha_{i,M}^2 |h_i|^2}{\frac{1}{\Gamma_1} + \sum_{i=1}^{M} \alpha_{i,M} |h_i|^2} \right\}. \tag{4.7b}$$

The parameters Γ_0 and Γ_1 are defined in Section 4.1. The h_i's are i.i.d zero-mean complex symmetric Gaussian random variables with unit variance, and the $\alpha_{i,M}$'s are the eigenvalues of $\mathbf{R_a}/M$. As mentioned in Section 4.1, the array correlation matrix $\mathbf{R_a}$ is a function of the array structure and so are its eigenvalues. Therefore, the average output SINR in (4.7) is a function of both the number of the antennas M and their positions in the array. To maximize the $\mathrm{E}\{\gamma\}$ in (4.7), one has to minimize q, given in (4.7b), over M and the positions of the antennas. Unfortunately, finding a closed-form solution for q in terms of the $\alpha_{i,M}$'s is not straightforward. It was expressed in a single integral form in Chapter 3. However, the expression given there is rather complicated and does not provide a mathematically tractable objective function for our optimization purposes. To circumvent this issue, we attempt to approximate q with a more tractable expression through the following lines. The expected value of the second term in the denominator of q in (4.7b) is equal to 1 according to (4.2a). Thus, if the value of the INR, Γ_1, is large enough compared to unity, e.g., $\Gamma_1 \geq 10$ (i.e., 10 dB) which is typical in an interference dominant cellular network, the parameter q in (4.7b) can be tightly upper-bounded as

$$q \leq \mathrm{E}\left\{ \frac{X}{Y} \right\} \tag{4.8a}$$

where

$$X = \sum_{i=1}^{M} \alpha_{i,M}^2 |h_i|^2 \tag{4.8b}$$

and

$$Y = \sum_{i=1}^{M} \alpha_{i,M} |h_i|^2. \tag{4.8c}$$

The inequality in (4.8a) becomes an equality as Γ_1 tends to infinity. Now, since X/Y is a ratio of two quadratic forms in Gaussian random variables, its first moment can be accurately approximated by [44]

$$\mathrm{E}\left\{ \frac{X}{Y} \right\} \approx \frac{\mathrm{E}\{X\}}{\mathrm{E}\{Y\}} \tag{4.9a}$$

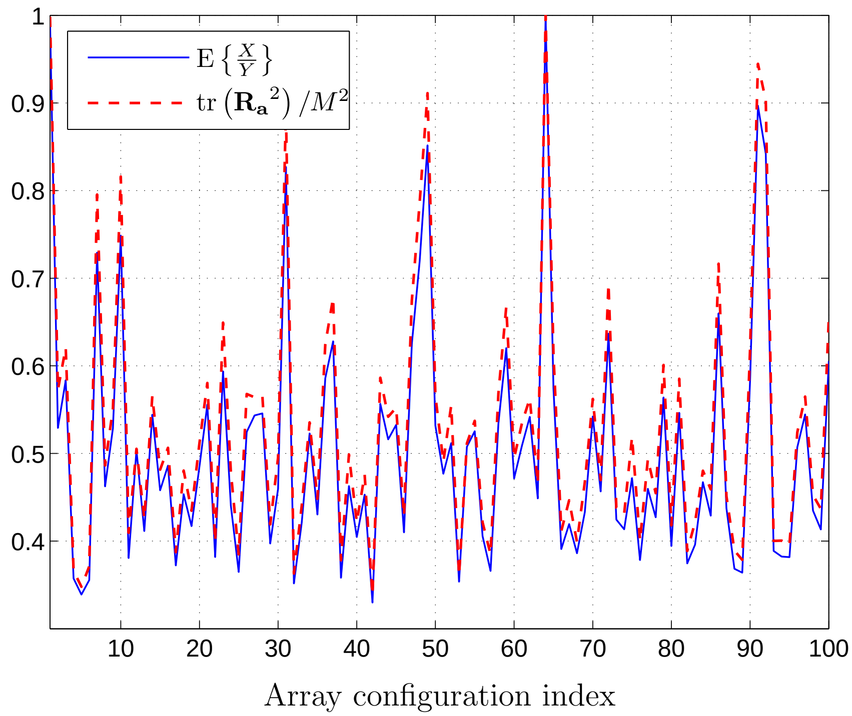


Figure 4.1. A comparison of $\mathrm{E}\left\{\frac{X}{Y}\right\}$ in (4.8) with its approximation in (4.9) for 100 random realizations of a linear array of antennas with length $L = 1$.

where

$$\frac{\mathrm{E}\{X\}}{\mathrm{E}\{Y\}} = \frac{\frac{1}{M^2}\mathrm{tr}\left(\mathbf{R_a}^2\right)}{\frac{1}{M}\mathrm{tr}(\mathbf{R_a})} = \frac{1}{M^2}\mathrm{tr}\left(\mathbf{R_a}^2\right) \tag{4.9b}$$

as a result of (4.2). Note that since $\mathbf{R_a}$ is a Hermitian matrix, the right side of (4.9b) is nothing but the square of the Frobenius norm of the normalized array correlation matrix $\mathbf{R_a}/M$. Figs. 4.1 and 4.2, generated by computer simulation, are two examples that demonstrate the accuracy of the approximation in (4.9a). To produce each of these figures, 100 random realizations of the antenna array are generated. Then, for each realization of the array (i.e., for a fixed $\mathbf{R_a}$ and a fixed set of $\alpha_{i,M}$'s), the left side of (4.9a) is calculated numerically and is plotted together with the value of the right side of (4.9b) calculated for that particular realization of the array. Fig. 4.1 corresponds to a linear 1D array of length $L = 1$. To produce a random array realization, in this case, the number of antennas M is randomly chosen from the set $\{2, 3, \ldots, 8\}$ and the position of each antenna x_i, $i = 1, 2, \ldots, M$, is independently chosen according to a uniform distribution over $[0, L]$. Fig. 4.2

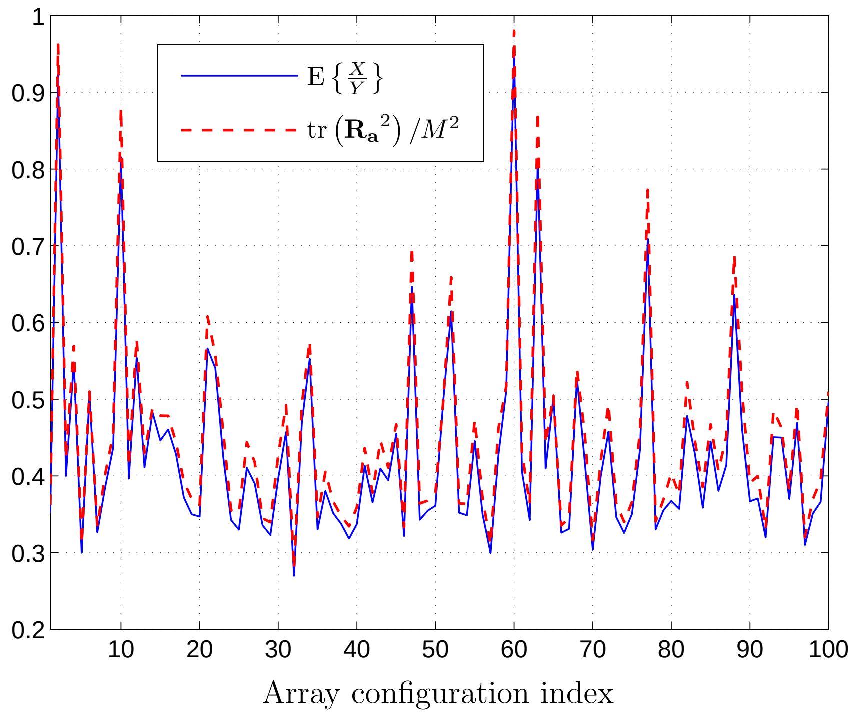


Figure 4.2. A comparison of $\mathrm{E}\left\{\frac{X}{Y}\right\}$ in (4.8) with its approximation in (4.9) for 100 random realizations of a 2D rectangular array of antennas with dimensions $L_x = 1$ and $L_y = 0.5$.

corresponds to a rectangular 2D array where $\mathcal{D} = [0, L_x] \times [0, L_y]$ with $L_x = 1$ and $L_y = 0.5$. For each random array realization, the number of antennas M is randomly chosen from the set $\{2, 3, \ldots, 10\}$ and the position of each antenna (x_i, y_i), $i = 1, 2, \ldots, M$, is independently chosen according to a uniform distribution over $\mathcal{D}$. A 2D-omnidirectional scattering environment is assumed in generating Figs. 4.1 and 4.2 where the spatial correlation function is given by $R_g(x) = \mathrm{J}_0(2\pi x)$ [5]. As Figs. 4.1 and 4.2 show, the graph of $\mathrm{tr}(\mathbf{R_a}^2)/M^2$ closely approximates that of $\mathrm{E}\{X/Y\}$. Moreover, it can be seen that the former is not only an accurate approximation but also an upper bound for the latter, i.e.,

$$\mathrm{E}\left\{\frac{X}{Y}\right\} \leq \frac{1}{M^2}\mathrm{tr}\left(\mathbf{R_a}^2\right). \tag{4.10}$$

Numerous other simulation results (not shown here for the sake of brevity) obtained for various 1D and 2D array shapes and various spatial correlation models also exhibit this inequality relationship. In fact, based on a large number of simulation results, we claim an even more general inequality relation as follows.

Conjecture 4.1. *Let $z_1, z_2, \ldots, z_M$ be a set of M i.i.d. real and positive random variables. Then, for any set of real and positive numbers $a_1, a_2, \ldots, a_M$, one has*

$$\mathrm{E}\left\{\frac{\sum_{i=1}^{M} a_i^2 z_i}{\sum_{i=1}^{M} a_i z_i}\right\} \leq \frac{\sum_{i=1}^{M} a_i^2}{\sum_{i=1}^{M} a_i}. \tag{4.11}$$

The proof of Conjecture 4.1 for $M = 1$ is obvious. A proof for the case of $M = 2$ is given in Appendix E. However, we have not been able to prove it for $M > 2$ and have only resorted to computer simulation to confirm its validity. The inequality in (4.10) is a special case of Conjecture 4.1 with $z_i = |h_i|^2$ and $a_i = \alpha_{i,M}$, $i = 1, 2, \ldots, M$.

Using (4.7a), (4.8a), and (4.10), the average output SINR of OC, in the case of a single interferer with $\Gamma_1 \gg 1$, can be tightly lower-bounded as

$$\mathrm{E}\{\gamma\} \geq \Gamma_0 \left[1 - \frac{1}{M^2}\mathrm{tr}\left({\mathbf{R}_{\mathbf{a}}}^2\right)\right]. \tag{4.12}$$

On the other hand, if $\Gamma_1 \ll 1$, the first term in the denominator of q in (4.7b) will be dominant and, therefore, neglecting the second term, the $\mathrm{E}\{\gamma\}$ can be tightly lower-bounded as

$$\mathrm{E}\{\gamma\} \geq \Gamma_0 \left[1 - \frac{\Gamma_1}{M^2}\mathrm{tr}\left({\mathbf{R}_{\mathbf{a}}}^2\right)\right]. \tag{4.13}$$

From (4.12) and (4.13), one arrives at two important observations. First, using the fact that (4.12) and (4.13) provide accurate approximations of the $\mathrm{E}\{\gamma\}$, one can conclude that, in the case of one interferer with either $\Gamma_1 \gg 1$ or $\Gamma_1 \ll 1$, the problem of maximizing $\mathrm{E}\{\gamma\}$ over the number and the positions of the antennas can be well approximated by the problem of minimizing $\mathrm{tr}\left({\mathbf{R}_{\mathbf{a}}}^2\right)/M^2$ over the number and the antenna positions. Second, since (4.12) and (4.13) provide lower bounds for the $\mathrm{E}\{\gamma\}$, one can use them to calculate the minimum attainable average output SINR performance of any given array of antennas and particularly that of the optimum array configuration. Moreover, the first of the two observations above can be generalized to the case of $N > 1$ interferers with arbitrary INR values, again by verifying it through the computer simulation. To check that, we generated several graphs similar to those of Figs. 4.1 and 4.2. This time, however, we numerically computed the value of $1 - \mathrm{E}\{\gamma\}/\Gamma_0$ from the definition of γ in (4.5) for many random configurations of the antennas and compared the resulting plot with the plot of $\mathrm{tr}\left(\mathbf{R_a}^2\right)/M^2$ for those array configurations. We found that, for any values of the parameters $N \geq 1$ and Γ_k's, $k = 1, \ldots, N$, used in our simulations, the locations of the global minima of the two plots always coincide, i.e., for any set of randomly generated array configurations, the configuration that minimized $1 - \mathrm{E}\{\gamma\}/\Gamma_0$ always minimized $\mathrm{tr}\left(\mathbf{R_a}^2\right)/M^2$ as well. This is similar to what is seen in Figs. (4.1) and (4.2) for the case of $N = 1$ and $\Gamma_1 = \infty$ where, in Fig. 4.1, the global minimum for both plots occurs at Configuration 42 and, in Fig. 4.2, it occurs at Configuration 32.

Interestingly, minimizing $\mathrm{tr}\left(\mathbf{R_a}^2\right)/M^2$ is also tantamount to maximizing some average performance measures of MRC in the presence of CCI and AWGN. This can be seen by looking at the expressions for two different performance metrics derived for MRC in [25]. The first performance metric is the long term average signal-power-to-interference-plus-noise-power ratio (SINRP). Using our notation, the SINRP can be expressed as [25]

$$\gamma_{SINRP} = \Gamma_0 \left(\frac{1 + l_2}{1 + (\sum_{k=1}^{k=N} \Gamma_k) l_2} \right) \tag{4.14a}$$

where

$$l_2 = \frac{\mathrm{tr}\left(\mathbf{R_a}^2\right)}{M^2}. \tag{4.14b}$$

It is easy to check that the γ_{SINRP} is a decreasing function of l_2 if $\sum_{k=1}^{k=N} \Gamma_k > 1$, i.e., if $\sum_{k=1}^{k=N} P_k > N_0$. In other words, in an interference dominant environment which is usually the case for the cellular networks, maximizing γ_{SINRP} is equivalent to minimizing l_2. The second performance metric, defined in [25], is the long term average signal-amplitude-to-square-root-interference-plus-noise-power ratio (SAINPR). The expression for this performance metric is given by [25]

$$\gamma_{SAINPR} = \sqrt{\frac{\Gamma_0}{1 + (\sum_{k=1}^{k=N} \Gamma_k) l_2}} \tag{4.15}$$

where l_2 is defined in (4.14b). Again, one can see that minimizing l_2 will maximize γ_{SAINPR}.

In the sequel, we develop a method to find the optimal array configuration that minimizes the quantity $\operatorname{tr}\left(\mathbf{R_a}^2\right)/M^2$ subject to the constraints imposed by the limited physical dimensions of the array. We start with the case of 1D linear arrays and subsequently extend our results to the case of 2D arrays.

4.2.2 Linear One Dimensional Arrays

Recalling from Section 4.1, the (i,j)'th element of $\mathbf{R_a}$, in this case, is equal to $R_g(x_i - x_j)$. Thus, one has

$$\operatorname{tr}\left(\mathbf{R_a}^2\right) = \operatorname{tr}\left(\mathbf{R_a}\mathbf{R_a}^H\right) = \sum_{i=1}^{M}\sum_{j=1}^{M} |R_g(x_i - x_j)|^2 = \sum_{i=1}^{M}\sum_{j=1}^{M} f(x_i - x_j) \quad (4.16)$$

where $f(x) \triangleq |R_g(x)|^2$. Therefore, based on the discussion in Section 4.2.1, our optimization problem can be formulated as

$$\begin{aligned} &\text{minimize} \quad \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} f(x_i - x_j) \\ &\text{over } M \text{ and } x_1, \ldots, x_M \text{ subject to} \\ &M \in \mathbb{N} \text{ and } x_1, \ldots, x_M \in [0, L] \end{aligned} \quad (4.17)$$

where $\mathbb{N}$ is the set of natural numbers. Solving the constrained optimization problem defined in (4.17) is not an easy task, in general, due to the discrete nature of M and also the nonlinear form of the objective function. In particular, it can be shown that, even for a fixed value of M, the objective function in (4.17) is not a convex function of $x_1, \ldots, x_M$. To circumvent this issue, we rewrite the objective function of (4.17) in an integral form by defining the function $c(t)$ over the interval $\mathcal{T} = [0,1]$ as follows.

$$c(t) = \sum_{i=1}^{M} x_i \mathbb{I}_{\mathcal{T}_i}(t), \qquad t \in \mathcal{T} = [0,1] \quad (4.18)$$

where $\mathcal{T}_i$ is the interval $((i-1)/M, i/M)$, $i = 1, \ldots, M$, and $\mathbb{I}_{\mathcal{A}}(\cdot)$ for some set $\mathcal{A}$ is the indicator function, i.e., its value is equal to one for any point inside $\mathcal{A}$ and zero anywhere outside $\mathcal{A}$. The function $c(t)$ is a nondecreasing simple function that maps a uniform partitioning of $[0,1]$ with M points to a generally nonuniformly distributed set of points $x_1, \ldots, x_M$ in $[0, L]$. Using the definition of $c(t)$ in (4.18), the objective function in (4.17) can be expressed as

$$\frac{1}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} f(x_i - x_j) = \int_0^1 \int_0^1 f(c(t) - c(t'))dtdt'. \qquad (4.19)$$

Let $\mathcal{S}$ denote the class of all functions of the form defined in (4.18) with a finite M. Hence, our optimization problem changes to minimizing the right side of (4.19) over all functions in $\mathcal{S}$. The new optimization problem in its current form is not any more tractable than the original one as it still suffers from the noncontinuous discrete nature of $c(t)$ which prevents us from further analytical development of the problem. However, we can proceed if we assume that $c(t)$ is a continuously differentiable function with positive derivative and attempt to minimize the right side of (4.19) over this new class of functions. The problem in its new form is then given by

$$\begin{aligned} &\text{minimize} \int_0^1 \int_0^1 f(c(t) - c(t'))dtdt' \\ &\text{over } c(t) \in \mathcal{C}([0,1]) \text{ subject to} \\ &\frac{dc(t)}{dt} > 0, \quad \forall t \in [0,1] \\ &\text{range}(c(t)) = [a,b] \subset [0,L] \end{aligned} \qquad (4.20)$$

where $\mathcal{C}([0,1])$ is the set of continuously differentiable functions over $[0,1]$. Any function in $\mathcal{S}$ can be thought of as a limit of a sequence of functions in $\mathcal{C}([0,1])$ with positive derivative. Therefore, if an optimal solution of the form (4.18) with finite M exists for the original problem, a numerical algorithm solving (4.20) should be able to get as close as desired to the optimal function. In other words, if by ever increasing the numerical precision of the optimization algorithm, the algorithm's output $c_{opt}(t)$ looks more and more like a simple function with M_{opt} different levels, one concludes that M_{opt} is the optimal number of antennas and adding more antennas will actually degrade the performance. Moreover, the M_{opt} different levels of $c_{opt}(t)$ provide the optimal positions of the antennas. As we will see later, this is the case for a 2D-omnidirectional scattering correlation model [5]. On the other hand, if there is an optimal solution $c_{opt}(t)$ to (4.20) with a continuously differentiable part, one concludes that the optimal number of antennas is infinite, i.e., adding more and more antennas will always improve the performance. In this case, for a practical system with M maximum available antennas, the values of $c_{opt}(t)$ at $\hat{t}_i = (i-1)/M$, $i = 1, \ldots, M$, can be taken as a practical solution for the optimal positions of the antennas. We will see that this is the case for an exponential correlation model.

Since $c(t)$ in (4.20) is continuously differentiable and has a positive derivative, it has an inverse function $c^{-1}(x)$ defined on $x \in [a, b]$ where $c^{-1}(x)$ is also continuously differentiable and has a positive derivative for every $x \in [a, b]$. Let $p(x) \triangleq \frac{d}{dx} c^{-1}(x)$ for $x \in [a, b]$ and extend the domain of $p(x)$ to $[0, L]$ by defining it to be zero anywhere outside the interval $[a, b]$. Then, one has $p(x) \geq 0$ and

$$\int_0^L p(x)dx = \int_a^b p(x)dx = c^{-1}(b) - c^{-1}(a) = 1 - 0 = 1. \tag{4.21}$$

Therefore, by changing the integration variables in (4.20) to $x = c(t)$ and $x' = c(t')$, the optimization problem in (4.20) reduces to

$$\begin{aligned} &\text{minimize} \int_0^L \int_0^L f(x - x')p(x)p(x')dxdx' \\ &\text{over } p(x) \text{ subject to} \\ &p(x) \geq 0 \text{ and} \int_0^L p(x)dx = 1. \end{aligned} \tag{4.22}$$

Note that the function $p(x)$ in (4.22) is, in fact, a probability density function describing the density of the antenna distribution over the interval $[0, L]$. Due to the continuity of $c(t)$ in (4.20), the $p(x)$ in (4.22) can be zero only over some intervals of the form $[0, a]$ and $[b, L]$ for some numbers a and b such that $0 \leq a < b \leq L$. The $p(x)$ must be positive everywhere between a and b. We relax this assumption for the new problem defined in (4.22) by letting $p(x)$ assume any arbitrary nonnegative values over the entire interval $[0, L]$. By doing this, we will automatically include a broader class of functions $c(t)$ in our optimization problem, i.e., the class of all bounded piecewise continuous functions with positive derivative. Moreover, to include the case of discrete antenna distributions, we further extend the class of functions $p(x)$ to functions of the form $p(x) = p_c(x) + \sum_{i=1}^{M} a_i \delta(x - x_i)$ where $p_c(x)$ is a nonnegative continuous function over $[0, L]$, $\delta(\cdot)$ is the Dirac delta function, M is some finite natural number, and $x_1, \ldots, x_M \in [0, L]$. This is equivalent to relaxing the "positive derivative" constraint on $c(t)$ in (4.20) to a "nonnegative derivative" constraint. The above relaxations allow us to add the class of all functions of the form (4.18) to the search domain of our optimization problem.

The problem in (4.22) is a quadratic convex optimization problem over the infinite dimensional space of the density functions $p(x)$. The convexity is a result of $f(x)$ being a positive definite function. To show this, recall that $R_g(x)$ is a positive definite function by definition (since it is a correlation function). Therefore, according to Bochner's theorem [41, p. 13], its Fourier transform $\mathcal{R}_g(\omega)$ is a real positive function. But, the Fourier transform of $f(x) = |R_g(x)|^2$ is given by $\mathcal{F}(\omega) = \mathcal{R}_g(\omega) \star \mathcal{R}_g^*(-\omega)$ where $\star$ denotes convolution. Thus, $\mathcal{F}(\omega)$ is also a real positive function and, consequently, its inverse Fourier transform $f(x)$ is a positive definite function. In the case of Rayleigh fading, considered in this work, it is well known that the power correlation coefficients are equal to the square of the magnitude of the channel correlation coefficients. Thus, the function $f(x) = |R_g(x)|^2$ is, in fact, the autocovariance function of the spatial process $|g_k(\mathbf{r})|^2$ (i.e., the instantaneous channel power process) where $g_k(\mathbf{r})$ is defined in Section 4.1. In other words,

$$f(\mathbf{r}_1 - \mathbf{r}_2) = \mathrm{Cov}(|g_k(\mathbf{r}_1)|^2, |g_k(\mathbf{r}_2)|^2), \qquad \forall k \in \{0, \ldots, N\} \tag{4.23}$$

where the position vector $\mathbf{r}$ is just the x-coordinate in the case of a 1D array.

The convex problem in (4.22) can always be solved numerically using the well known methods of convex optimization, for example, using the interior point method [43]. This can be done by rectangular approximation of each integral in (4.22) with n points. This reduces the infinite dimensional problem in (4.22) to the n dimensional convex problem

$$\begin{aligned} &\text{minimize } \mathbf{p}^T \mathbf{F} \mathbf{p} \\ &\text{subject to} \\ &\mathbf{p} \geq 0 \text{ and } \mathbf{1}_n^T \mathbf{p} = n \end{aligned} \tag{4.24}$$

where $\mathbf{p}$ is an $n \times 1$ vector in $\mathbb{R}^n$, the $\mathbf{1}_n$ is an $n \times 1$ vector whose elements are all equal to 1, and the matrix $\mathbf{F}$ is an $n \times n$ positive definite matrix whose

(i, j)'th entry is equal to $f(x_i - x_j)$ where $x_i = (i-1)L/n$ for $i = 1, \ldots, n$. The value of n can be chosen according to the desired accuracy in approximating $p(x)$ from $\mathbf{p}$. A larger n results in a better accuracy but at the same time increases the computational complexity of the problem in (4.24).

For the special case of an exponential correlation function, we are able to solve (4.22) analytically. This is stated in the following theorem.

Theorem 4.2. *Given an exponential spatial correlation function* $R_g(x) = \exp(-k|x|)$*, a solution of (4.22) is given by*

$$p_{opt}(x) = \frac{k}{1+kL} + \frac{1}{2+2kL}\left(\delta(x) + \delta(x-L)\right). \tag{4.25}$$

The proof of Theorem 4.2 can be found in Appendix F. The optimal function $c_{opt}(t)$ corresponding to the $p_{opt}(x)$ in (4.25) is obtained by integrating (4.25) with respect to x and calculating the inverse function. The result is

$$c_{opt}(t) = \begin{cases} 0 & 0 \leq t < t_0 \\ \dfrac{1+kL}{k} \cdot (t - t_0) & t_0 \leq t < 1 - t_0 \\ L & 1 - t_0 \leq t \leq 1 \end{cases} \tag{4.26a}$$

where

$$t_0 = \frac{1}{2+2kL}. \tag{4.26b}$$

The physical interpretation of the expression for $p_{opt}(x)$ in (4.25) and its corresponding $c_{opt}(t)$ in (4.26) regarding the arrangement of the antennas is explained in Section 4.3.

4.2.3 Two Dimensional Arrays

The case of a 2D array can also be solved by following similar steps to those followed in the 1D array scenario, that is, by converting the original discrete non-convex problem to a continuous quadratic convex problem. However, the intermediate steps, in this case need to be examined in more detail. Similar to the 1D case, we start with expressing our objective function in terms of the samples of the spatial correlation function, i.e.,

$$\begin{aligned}\operatorname{tr}\left(\mathbf{R_a}^2\right) = \operatorname{tr}\left(\mathbf{R_a}\mathbf{R_a}^H\right) &= \sum_{i=1}^{M}\sum_{j=1}^{M}\left|R_g(x_i - x_j, y_i - y_j)\right|^2 \\ &= \sum_{i=1}^{M}\sum_{j=1}^{M} f(x_i - x_j, y_i - y_j)\end{aligned} \tag{4.27}$$

where $f(x, y) \triangleq |R_g(x, y)|^2$ is the two-dimensional autocovariance function of the instantaneous channel power process (see (4.23) and note that $\mathbf{r} = (x, y)$ in the case of 2D arrays). Our optimization problem is then given by

$$\begin{aligned}&\text{minimize} \quad \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} f(x_i - x_j, y_i - y_j) \\ &\text{over } M \text{ and } (x_1, y_1), \ldots, (x_M, y_M) \text{ subject to} \\ &M \in \mathbb{N} \text{ and } (x_1, y_1), \ldots, (x_M, y_M) \in \mathcal{D}.\end{aligned} \tag{4.28}$$

Next, similar to the 1D case, we attempt to rewrite the summations in the objective function in (4.28) in terms of integrals by introducing an appropriate mapping. In the 1D scenario, we did so by partitioning the unit-length interval $\mathcal{T} = [0, 1]$ to M equal length open subintervals $\mathcal{T}_i$, $i = 1, \ldots, M$, and then defining a simple function $c(t)$ that maps all the points inside the subinterval $\mathcal{T}_i$ to x_i. In the 2D scenario, this can be done by letting $\mathcal{T}$ be a bounded unit-area subset of the plane that is partitioned into M open equal area subsets $\mathcal{T}_i$, $i = 1, \ldots, M$ and by defining a two-dimensional mapping $\mathbf{c}(s, t) = (c_x(s, t), c_y(s, t))$ from $\mathcal{T}$ to $\mathcal{D}$ that maps every point (s, t) inside $\mathcal{T}_i$ to the point (x_i, y_i) in $\mathcal{D}$. Using indicator function notation, the mapping $\mathbf{c}(s, t)$ can be written as

$$\mathbf{c}(s, t) = \sum_{i=1}^{M} \mathbf{r}_i \mathbb{I}_{\mathcal{T}_i}(s, t), \qquad (s, t) \in \mathcal{T} \tag{4.29}$$

where $\mathbf{r}_i = (x_i, y_i)$. Now, the objective function in (4.28) can be expressed in terms of the $\mathbf{c}(s, t)$ as follows.

$$\begin{aligned}&\frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} f(x_i - x_j, y_i - y_j) = \\ &\qquad \int_{\mathcal{T}}\int_{\mathcal{T}} f(c_x(s, t) - c_x(s', t'), c_y(s, t) - c_y(s', t'))ds dt ds' dt'\end{aligned} \tag{4.30}$$

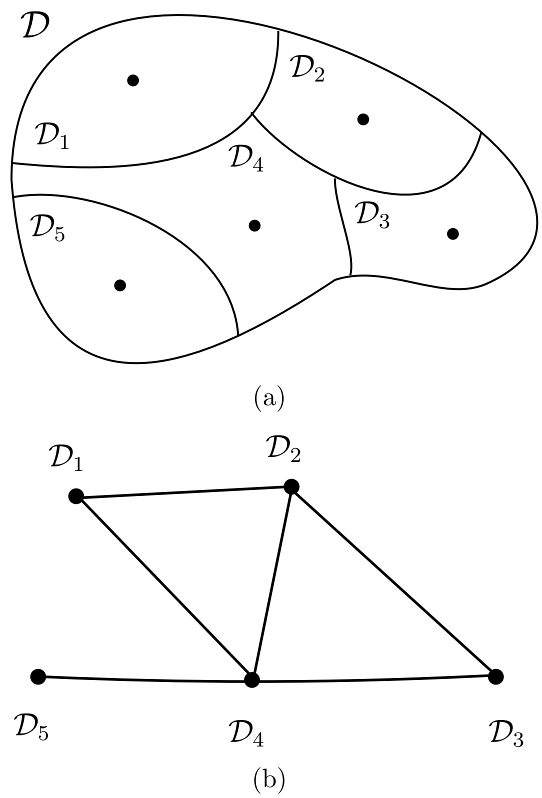


Figure 4.3. (a) An example partitioning of $\mathcal{D}$ into $M = 5$ regions. The locations of the antennas are shown with small circles. (b) The corresponding graph $\mathcal{G}$ of this partitioning of $\mathcal{D}$.

where each integral in (4.30) is taken over the set $\mathcal{T}$. In the 1D case, we made use of the fact that the positions of the antennas x_i's can be sorted in an ascending (descending) order. This enabled us to interpret the noncontinuous function $c(t)$ defined in (4.18) as the limit of a sequence of increasing (decreasing) continuously differentiable functions. The monotonicity of these functions implies that they are injective and, thus, invertible mappings. As a result, we are able to apply the change of variables theorem to transform the integrals to the desired quadratic forms. To extend this method to the 2D case, one has to show that it is possible to think of the noncontinuous mapping $\mathbf{c}(s,t)$ in (4.29) as the limit of a sequence of 2D bijective (hence, invertible) continuously differentiable mappings , a.k.a. diffeomorphisms, from $\mathcal{T}$ to $\mathcal{D}$. Here is how one can make such a sequence of diffeomorphisms. Let $\mathcal{A}$ denote the set of M points $(x_i, y_i) \in \mathcal{D}$, $i = 1, \ldots, M$, representing the locations of

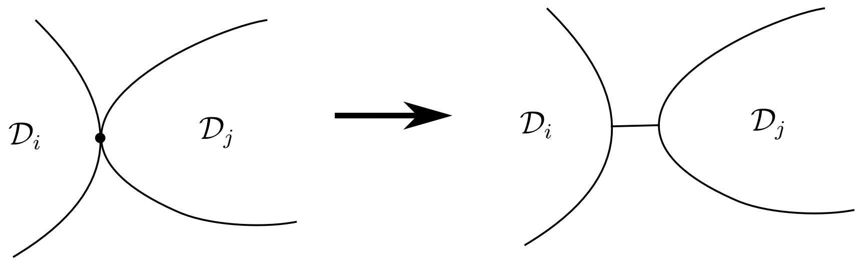


Figure 4.4. A transformation of the regions in a partitioning of $\mathcal{D}$ to avoid an isolated point in the intersection of the closures of the regions.

the antennas. Partition $\mathcal{D}$ into M open subsets $\mathcal{D}_i$ with smooth boundaries such that each subset contains exactly one point of $\mathcal{A}$, for example by forming the Voronoi regions of the points in $\mathcal{A}$. Fig. 4.3(a) shows an example of such a partitioning for $M = 5$. It is always possible to make the partitioning in such a way that the intersection of the closures of any two $\mathcal{D}_i$ and $\mathcal{D}_j$ is either empty or a curve with nonzero length. Because, if the closures of $\mathcal{D}_i$ and $\mathcal{D}_j$ have only a number of isolated points in common, those points can be easily removed using the transformation shown in Fig. 4.4. This is possible because $\partial\mathcal{D}$ is assumed to be a simple curve. To such a partitioning of $\mathcal{D}$, we assign a planar graph $\mathcal{G}$ whose vertices are the regions $\mathcal{D}_i$, $i = 1, \ldots, M$, and two vertices $\mathcal{D}_i$ and $\mathcal{D}_j$ are connected by an edge if and only if the intersection of the closures of $\mathcal{D}_i$ and $\mathcal{D}_j$ is nonempty. Fig. 4.3(b) shows the graph corresponding to the partitioning of $\mathcal{D}$ shown in Fig. 4.3(a). Now, for some partitioning of $\mathcal{T}$ into $\mathcal{T}_i$'s, define $\hat{\mathbf{c}}(s,t)$ to be a mapping from $\mathcal{T}$ to $\mathcal{D}$ that maps $\mathcal{T}_i$ to $\mathcal{D}_i$ and, when its domain is restricted to $\mathcal{T}_i$ ($i = 1, \ldots, M$), is continuously differentiable and bijective. For this mapping to be globally continuous and injective over the entire set $\mathcal{T}$, it is necessary for the partitioning graph of $\mathcal{T}_i$'s to be isomorphic to that of the $\mathcal{D}_i$'s. Taking the $\mathcal{T}_i$'s to be scaled versions of the $\mathcal{D}_i$'s (so that their areas sum to unity) may seem to be an appropriate choice for the $\mathcal{T}_i$'s. However, this is not the case since the $\mathcal{T}_i$'s generated in this way do not necessarily have equal areas. Thus, the question is whether one can find a partitioning of $\mathcal{T}$ into equal area regions such that the corresponding partition graph is isomorphic to that of the $\mathcal{D}_i$'s. The following theorem answers this question.

Theorem 4.3. *For any arbitrary partitioning of $\mathcal{D}$ with partition graph $\mathcal{G}$, there exists a partitioning of $\mathcal{T}$ into equal area regions with a partition graph isomorphic to $\mathcal{G}$.*

The proof of Theorem 4.3 is given in Appendix G. Once an equal area partitioning of $\mathcal{T}$ with a graph isomorphic to that of $\mathcal{D}$ is found, it is possible to make a diffeomorphism $\hat{\mathbf{c}}(s,t)$ from $\mathcal{T}$ to $\mathcal{D}$ that approximates the mapping $\mathbf{c}(s,t)$ in (4.29). The continuous mapping $\hat{\mathbf{c}}(s,t)$ made in this way can be made as close as desired to the noncontinuous mapping $\mathbf{c}(s,t)$ in (4.29) by having it map every point in $\mathcal{T}_i$ whose distance from the boundary of $\mathcal{T}_i$ is more than a small number ϵ to an ϵ-neighbourhood of the point (x_i, y_i) in $\mathcal{D}_i$, $i = 1, \ldots, M$.

Up to this point, we have shown that any mapping $\mathbf{c}(s,t)$ of the form (4.29) can be approached by a sequence of diffeomorphisms from $\mathcal{T}$ to $\mathcal{D}$. Thus, using a reasoning similar to that employed in the 1D case, one can replace the original optimization problem in (4.28) by

$$\begin{aligned}
&\text{minimize} \quad \int_{\mathcal{T}}\int_{\mathcal{T}} f(c_x(s,t)-c_x(s',t'), c_y(s,t)-c_y(s',t'))dsdtds'dt' \\
&\text{over } \mathbf{c}(s,t) = (c_x(s,t), c_y(s,t)) \text{ subject to} \\
&\mathbf{c}(s,t) \text{ is a diffeomorphism from } \mathcal{T} \text{ to } \mathcal{D}
\end{aligned} \tag{4.31}$$

which is the 2D equivalent of (4.20). Since $\mathbf{c}(s,t)$ in (4.31) is a diffeomorphism, it has an inverse $\mathbf{c}^{-1}(x,y)$ that is also a diffeomorphism. Thus, one can apply the change of variables theorem for multiple integrals [45, p. 421] to convert the optimization problem in (4.31) to

$$\begin{aligned}
&\text{minimize} \quad \int_{\mathcal{D}}\int_{\mathcal{D}} f(x-x', y-y')p(x,y)p(x',y')dxdydx'dy' \\
&\text{over } p(x,y) \text{ subject to} \\
&p(x,y) \geq 0 \text{ and } \int_{\mathcal{D}} p(x,y)dxdy = 1
\end{aligned} \tag{4.32}$$

where $p(x,y) = |J_{\mathbf{c}^{-1}}(x,y)|$ is the absolute value of the Jacobian determinant [45] of the inverse mapping $\mathbf{c}^{-1}(x,y)$. The last constraint on $p(x,y)$ is a consequence of the fact that the area of $\mathcal{T}$ is equal to one and

$$\int_{\mathcal{D}} p(x,y)dxdy = \int_{\mathcal{D}} |J_{\mathbf{c}^{-1}}(x,y)|dxdy = \int_{\mathcal{T}} dsdt = Area(\mathcal{T}) = 1. \tag{4.33}$$

The function $p(x, y)$ is, in fact, a probability density function that describes the density of the antenna distribution over the region $\mathcal{D}$. Similar to the 1D case, we extend the class of functions $p(x, y)$ in (4.32) to include the Dirac delta functions. In contrast with the 1D case, however, once the optimum density function $p_{opt}(x, y)$ is found by solving the optimization problem in (4.32), one cannot find a unique mapping $\mathbf{c}_{opt}^{-1}(x, y)$ corresponding to $p_{opt}(x, y)$. This is because the Jacobian determinant of an invertible 2D mapping does not uniquely define the mapping itself. However, since we are only interested in finding the optimum density distribution of the antennas, i.e., the $p_{opt}(x, y)$, finding a mapping $\mathbf{c}_{opt}^{-1}(x, y)$ corresponding to $p_{opt}(x, y)$ is, in fact, unnecessary. In other words, $p_{opt}(x, y)$ contains all the information about the optimum arrangement of the antennas.

The optimization problem in (4.32) is a quadratic convex problem since $f(x, y)$ is a two-variable positive definite function[46]. The positive definiteness of $f(x, y)$ is a result of (4.23) for the 2D case, i.e., where the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ in (4.23) are given by (x, y) and (x', y'), respectively. To solve the problem in (4.32) numerically, we can approximate it by a finite dimensional convex problem similar to (4.24). This can be done by covering $\mathcal{D}$ with a rectangular grid with n cells and assuming that the integrands in (4.32) remain approximately constant within each cell. In this case, the (i, j)'th element of the matrix $\mathbf{F}$ in (4.24) is given by $f(x_i - x_j, y_i - y_j)$ where i and j run from 1 to n and (x_i, y_i) is a point inside the i'th cell for some ordering of the cells in the grid. Once the problem is converted to the finite dimensional convex programming of the form (4.24), it can be solved efficiently using the well known numerical methods available for convex optimization [43].

4.3 Numerical Results and Discussion

In this section, we provide a number of examples for both the 1D and the 2D antenna arrangement scenarios where we numerically solve the convex optimization problems in (4.22) and (4.32) using the finite dimensional approximation in (4.24). We also discuss the interpretation of the optimal solutions and

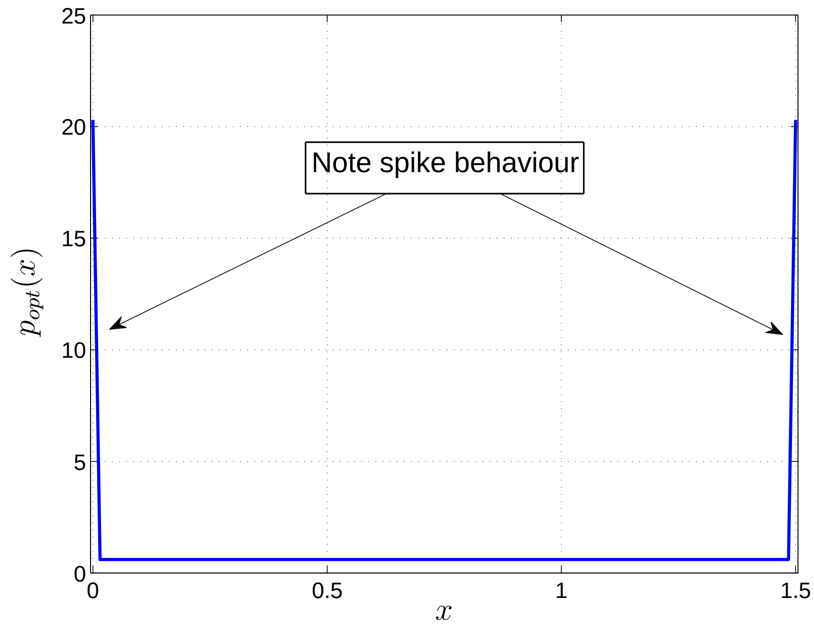
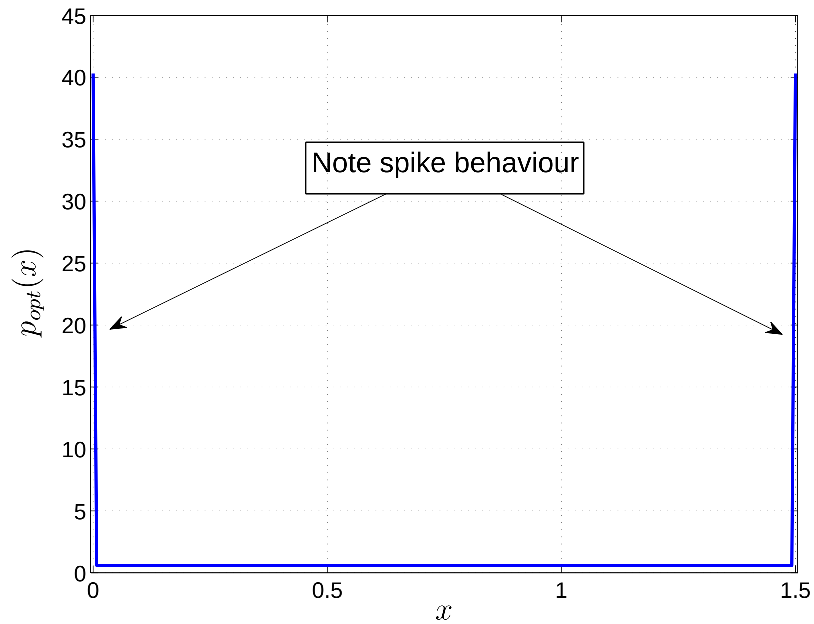


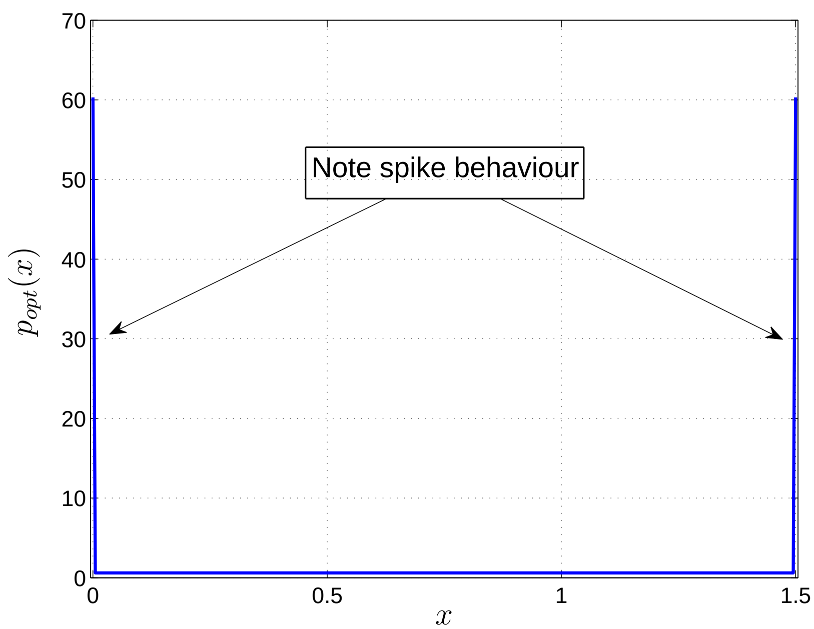
Figure 4.5. A solution of the optimization problem in (4.24) for the linear array in Example 4.1 with $n = 100$

their relation to the actual antenna placements. All the numerical solutions are obtained using the *interior point convex optimization algorithm* available in MATLAB's optimization toolbox. The first two examples consider the optimal design of a linear 1D array for two different correlation models while the last three examples focus on the optimal design of a 2D array for regions with different shapes. The physical dimensions of the arrays in all examples are in the unit of wavelength. Finally, at the end of this section, we compare the performance of the optimal arrangement solution obtained in one of the example scenarios with those of a few suboptimal designs to find out how much benefit can be achieved by deploying the optimal antenna arrangement.

Example 4.1. Consider a linear 1D array whose normalized length is equal to $L = 1.5$. Assume an exponential correlation model where the spatial correlation function is given by $R_g(x) = \exp(-|x|)$. We solve the optimization problem in (4.24) with the approximation dimension $n = 100$. Fig. 4.5 shows a plot of the elements of the optimum vector $\mathbf{p}_{opt}$ versus their indices which approximates the plot of $p_{opt}(x)$, the optimal solution to (4.22). It can be



(a)



(b)

Figure 4.6. Solutions of the optimization problem in (4.24) for the linear array in Example 4.1 for two different values of the approximation dimension n. (a) the approximation dimension $n = 200$. (b) the approximation dimension $n = 300$.

seen that except for two sharp spikes on both ends, the plot of $\mathbf{p}_{opt}$ is approximately constant over $[0, 1.5]$. Executing the optimization algorithm for values

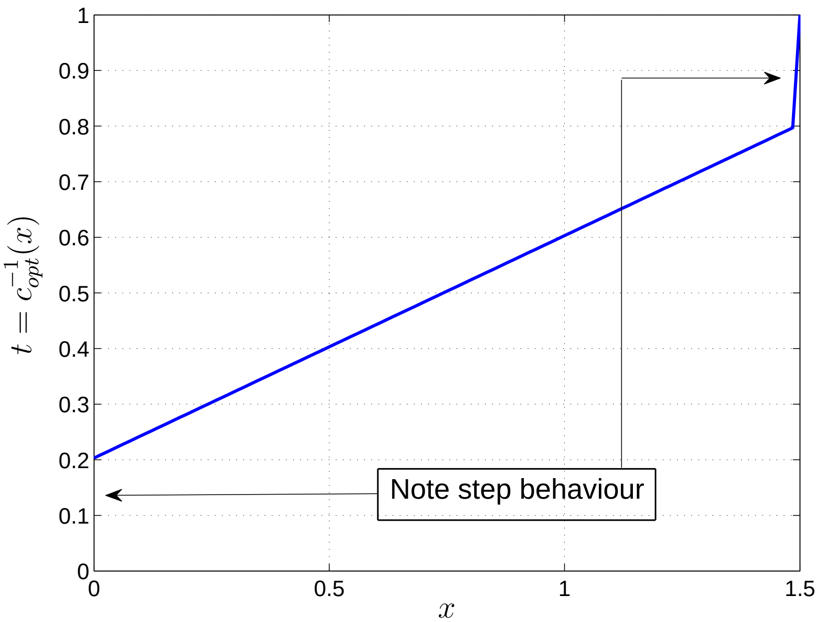


Figure 4.7. The inverse of the optimum antenna arrangement function in Example 4.1, i.e., the integral of $p_{opt}(x)$ in Fig. 4.5.

of n larger than 100 will result in even sharper and higher spikes in the plot of $\mathbf{p}_{opt}$ as seen in Fig. 4.6(a) and 4.6(b). This suggests that the observed spikes may actually tend to two Dirac delta functions as n goes to infinity. This can be confirmed by integrating the plot of $\mathbf{p}_{opt}$, i.e., by calculating the partial summations of the elements of $\mathbf{p}_{opt}$ multiplied by $1/n$. Fig. 4.7 shows the plot of the integral of $\mathbf{p}_{opt}$ which approximates the plot of $c_{opt}^{-1}(x)$. The two steps in Fig. 4.7 on both ends of the interval correspond to the two spikes in the plot of $\mathbf{p}_{opt}$. Other numerical results (not shown here) show that the heights of these steps do not change as n increases. The heights of the steps are in fact the weights of the corresponding Dirac delta functions in $p_{opt}(x)$. A closer look at the plots of Figs. 4.5 – 4.7 reveals that they are in excellent agreement with the analytical solutions in (4.25) and (4.26).

The optimal solution for $c(t)$ in Fig. 4.7 can be interpreted as follows. First, since the optimal solution is not a simple function of the form (4.18), the optimal number of antennas is infinite, meaning that adding more and more antennas in this case will always improve the performance. Second, for a given finite number of antennas M, a reasonable choice for a near optimal

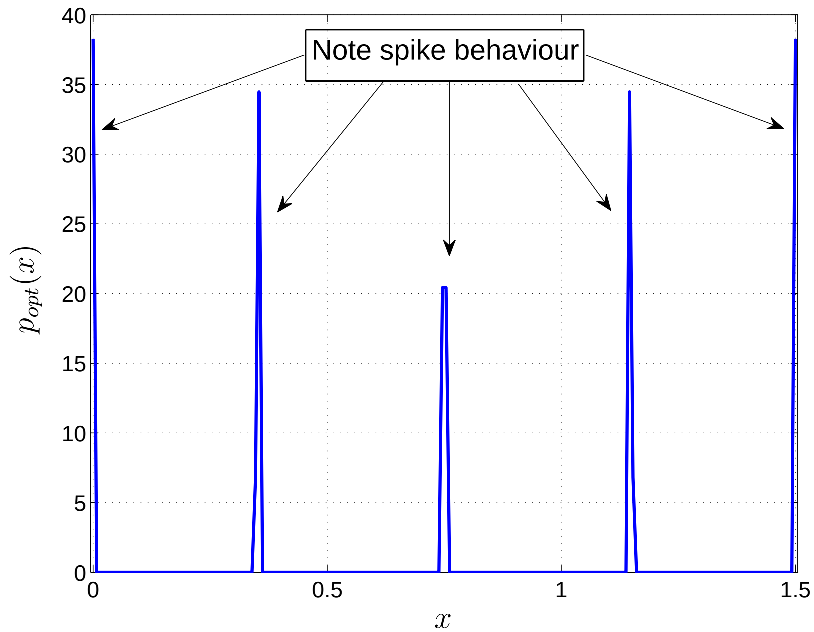


Figure 4.8. A solution of the optimization problem in (4.24) for the linear array in Example 4.2.

arrangement of the antennas is obtained by uniformly sampling the vertical axis of Fig. 4.7 at points $\hat{t}_i = (i-1)/M$, $(i = 1, \ldots, M)$, and finding the corresponding points $x_i = c_{opt}(\hat{t}_i)$ on the horizontal axis. For $M = 2$ and $M = 3$, this translates to a uniform placement of antennas over $[0, 1.5]$. However, for $M \geq 4$ the antenna distribution is not uniform at the end points. Moreover, when $1/(M-1) \leq 0.2$ or equivalently when $M \geq 6$, the optimum placement results in putting more than one antenna at each end. Particularly, as M grows larger, Fig. 4.7 suggests that the optimal placement is to put 20 percent of the antennas at each end and distribute the remaining 60 percent of them uniformly over $[0, 1.5]$. Placing more than one antenna at a given point of space may sound practically redundant. However, it can be easily shown that two co-located antennas each receiving an average power P can be replaced with a single antenna with an average received power $2P$ without changing the value of the average output SINR. Therefore, what Fig. 4.7 means in practice is that for the exponential correlation function of our example the optimal array structure is the one that captures 20 percent of the total electromagnetic

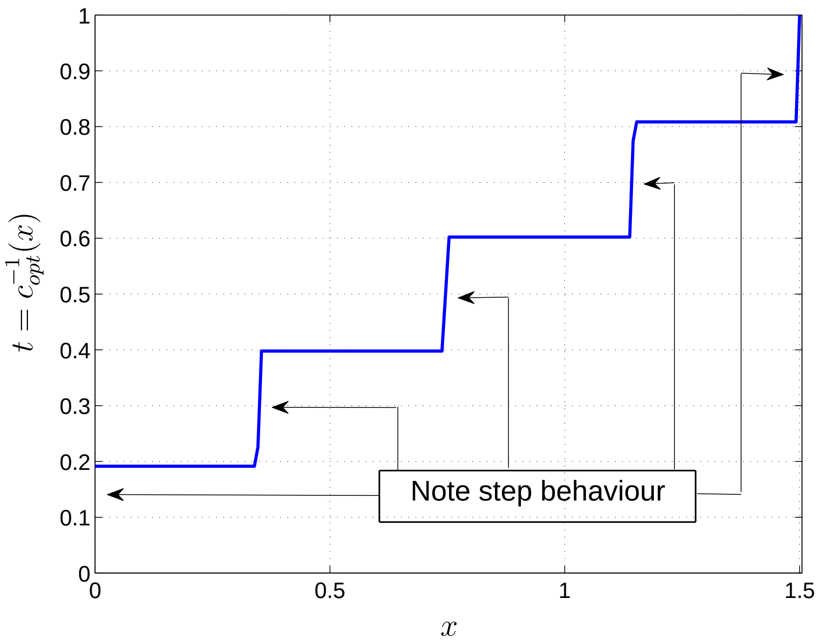


Figure 4.9. The inverse of the optimum antenna arrangement function in Example 4.2, i.e., the integral of $p_{opt}(x)$ in Fig. 4.8.

power incident to the array at each end-point and captures the remaining 60 percent of the total power uniformly over $[0, 1.5]$. In general, the antenna distribution density $p(x)$ introduced in Section 4.2.2 can be interpreted as the distribution density of the power captured by the array within the interval $[0, L]$. The same interpretation can be used for the two-dimensional arrays, i.e., the density function $p(x, y)$ defined in Section 4.2.3 can be interpreted as the distribution density of the electromagnetic power captured by the array within the region $\mathcal{D}$.

Example 4.2. Again, consider a linear 1D array of length $L = 1.5$. However, this time, assume a 2D-omnidirectional scattering model where the spatial correlation function is given by $R_g(x) = \mathrm{J}_0(2\pi x)$ [5]. We solve the optimization problem in (4.24) with $n = 200$. The resulting $p_{opt}(x)$ and its integral $c_{opt}^{-1}(t)$ are plotted in Figs. 4.8 and 4.9, respectively. Fig. 4.8 suggests that the optimal density, in this case, consists of 5 Dirac delta functions that are almost uniformly positioned between 0 and 1.5. The optimal density does not have a continuous part and its integral, shown in Fig. 4.9, is a simple function

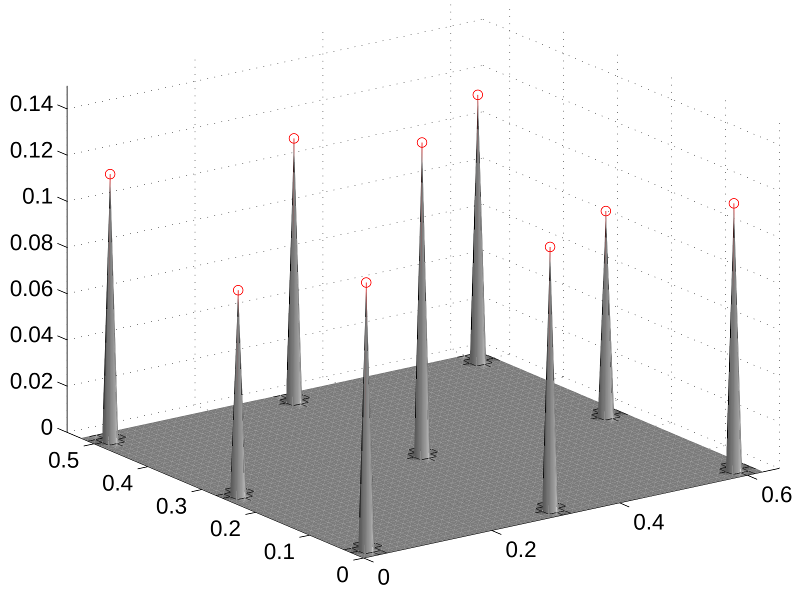


Figure 4.10. A solution of the optimization problem in (4.24) for the 2D array in Example 4.3.

with 5 different levels. Therefore, the optimal number of antennas is $M = 5$. Moreover, from Fig. 4.9, one can observe that the heights of the steps (i.e., the weights of the delta functions in Fig. 4.8) are all equal to 0.2. Therefore, one can conclude that the optimal arrangement, in this case, is an array consisting of $M = 5$ identical antennas uniformly positioned over $[0, 1.5]$

Example 4.3. In this example, we consider a 2D array where the available space $\mathcal{D}$ for arranging the antennas is a 0.5×0.6 rectangle. Again, we assume a 2D-omnidirectional scattering model. Under this model, the two-variable spatial correlation function is given by $R_g(x, y) = \mathrm{J}_0(2\pi\sqrt{x^2 + y^2})$. The optimization problem in (4.24) is solved for a rectangular grid with $n = 1833$ cells (see Section 4.2.3 for more details). Fig. 4.10 shows the resulting optimal density $p(x, y)$ normalized by n. As shown in this figure, the optimal solution consists of 9 delta functions corresponding to $M = 9$ antennas located on a rectangular lattice within the available space $\mathcal{D}$. Although the optimal solution is uniform in terms of the positions of the antennas, the optimal powers that must be captured by different antennas, i.e, the weights of different delta functions in Fig. 4.10 are not all equal. The values of the weights of the delta

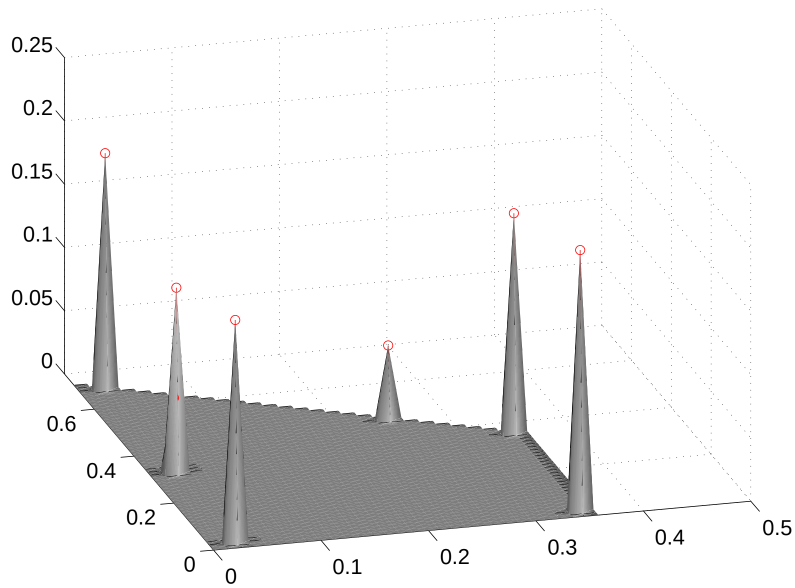


Figure 4.11. A solution of the optimization problem in (4.24) for the 2D array in Example 4.4.

functions in Fig. 4.10 are shown with a stem plot marked with small circle markers. One can observe that the heights of the stems are not all equal.

Example 4.4. The available region $\mathcal{D}$ for antenna arrangement, in Example 4.3, had a symmetric shape (a rectangle). However, in a practical application, the region on the receiver board (or antenna mount) available for antenna placement may have a more complicated shape. In this example, under the same correlation model as in Example 4.3, we consider a trapezoidal region $\mathcal{D}$ as shown in Fig. 4.11. The number of cells in the rectangular grid covering $\mathcal{D}$, i.e., the dimension of the optimization problem in (4.24) is chosen to be $n = 1107$. Fig. 4.11 shows that the optimal solution places $M = 6$ antennas on the boundary of the region. Moreover, the optimal distribution of the power absorbed by the antennas is not uniform.

Example 4.5. The optimization problem formulated in (4.32) can be virtually applied to any two-dimensional region with a smooth boundary. The region $\mathcal{D}$ neither needs to be a *simply connected* region (i.e., a region with no holes) nor even a connected region. To show the versatility of the optimization problem in (4.32), we choose $\mathcal{D}$ to be a disk with a hole inside it as shown in Fig. 4.12.

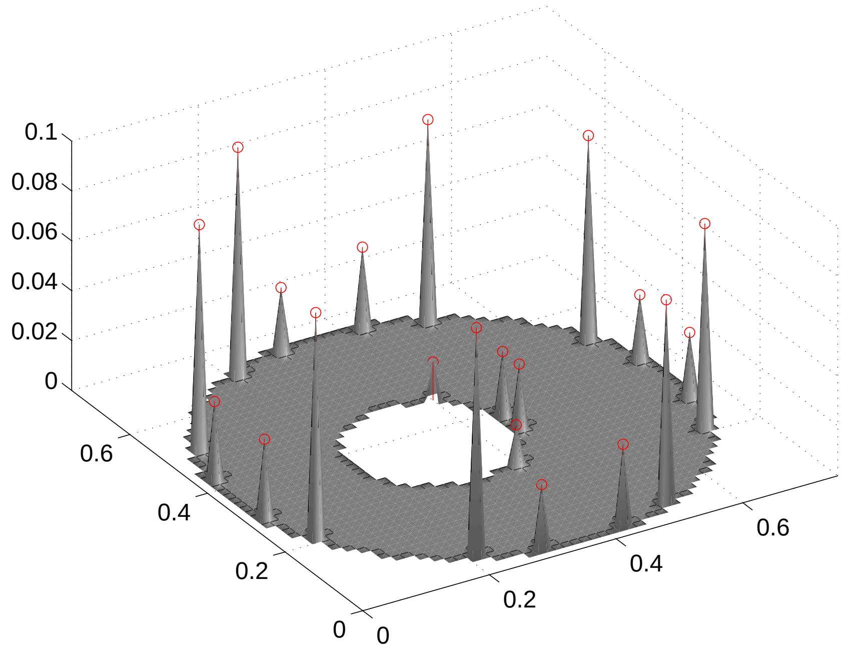


Figure 4.12. A solution of the optimization problem in (4.24) for the 2D array in Example 4.5.

We assume a 3D-omnidirectional scattering model where the spatial correlation function is given by $R_g(x, y) = \mathrm{sinc}(2\pi\sqrt{x^2 + y^2})$ [34]. The number of cells in the approximation grid is $n = 1558$. Fig. 4.12 shows that the optimum solution, in this example, consists of 20 impulses corresponding to $M = 20$ antennas located on the boundaries of the region.

The above array design examples clearly show the applicability of the formulation derived in Section 4.2 to the design of a wide range of 1D and 2D arrays under various spatial correlation models. However, a natural question that arises at this point is that, for a given array shape, how much improvement one can achieve by using the optimal arrangement solution over a suboptimal arrangement. We answer this question by examining the performance of the optimal solution in Example 4.3 and comparing it with those of three different suboptimal arrangements. The first suboptimal arrangement is one with four identical antennas located on the four corners of the available 0.5×0.6 rectangular region. The second suboptimal arrangement consists of 16 identical antennas located on the lattice points of a 4×4 rectangular lattice covering the available area. In the third suboptimal arrangement, we consider the lim-

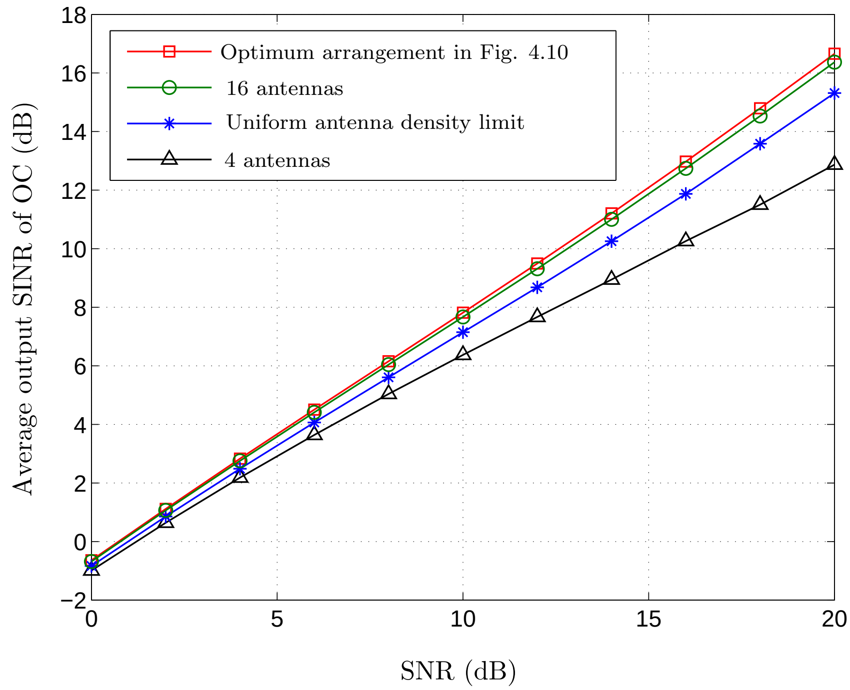


Figure 4.13. A comparison of the optimal arrangement in Example 4.3 with three suboptimal arrangements in terms of their achievable average output SINR when used in an OC receiver.

iting case of a very large number of identical antennas uniformly distributed over the available rectangular region. This is equivalent to a uniform received power density $p(x, y)$ where the array captures the incident electromagnetic power uniformly over the entire surface of the available region. In our simulations, this limiting case is implemented by placing an antenna in every cell of the approximation grid used in Example 4.3. The three suboptimal arrangements are compared with the optimal solution in Example 4.3 based on three different performance measures, namely, the average output SINR of OC, the outage probability of OC, and the average BER of MRC in the presence of CCI. Fig. 4.13 shows the average output SINR of OC as a function of the average input SNR for the optimal arrangement as well as the three suboptimal arrangements. The graphs in Fig. 4.13 are generated by computer simulation. The number of the interfering users is set to $N = 3$ and the total average input SIR is set to 0 dB where the total average input SIR, in this case, is defined as the ratio of the desired user's power to the sum of the powers of N equal-power interferers. One can see that the optimal arrangement out-

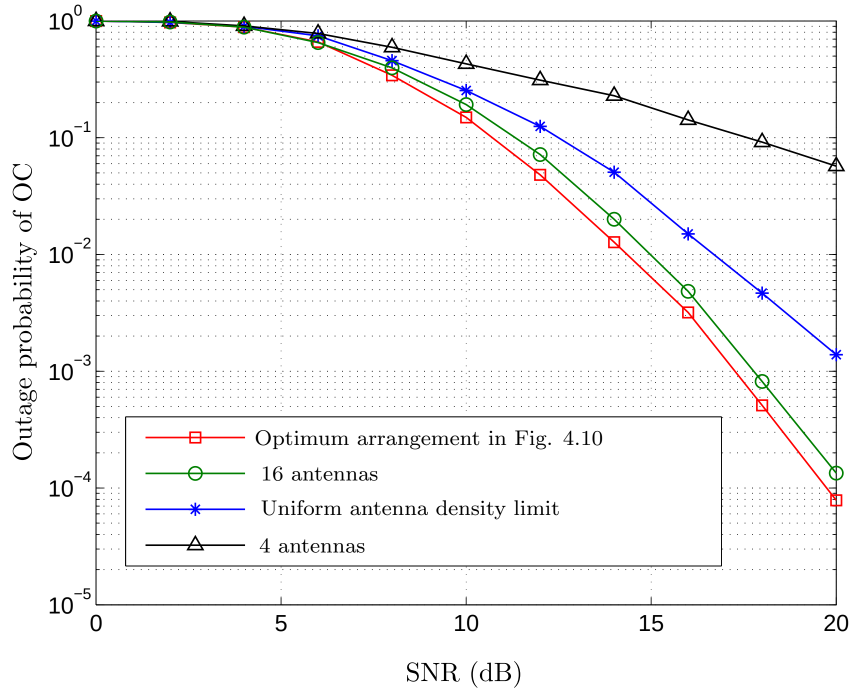


Figure 4.14. A comparison of the optimal arrangement in Example 4.3 with three suboptimal arrangements in terms of their achievable outage probability when used in an OC receiver.

performs all the suboptimal designs at all values of the input SNR. Moreover the improvement achieved by the optimal design over the suboptimal ones increases as the average input SNR gets larger. For example, at an input SNR level of 20 dB, the optimal design with 9 antennas provides an average output SINR that is about 4 dB larger than that of the suboptimal design with 4 antennas. An important insight is that adding more antennas than the optimal number 9 results in a decrease in the average output SINR. As seen in Fig. 4.13, at an input SNR value of 20 dB, the design with 16 antennas is slightly worse than the optimal solution while the design with infinite number of antennas (i.e., the uniform antenna density limit) is about 1.5 dB worse than the optimal solution. Significant implementation and equipment costs can be spared by applying the methods and results presented in this chapter. The next performance measure considered is the outage probability of OC, i.e., the probability that the instantaneous output SINR of OC in (4.5) drops below a certain threshold level. Fig. 4.14 shows the outage probability of OC as a

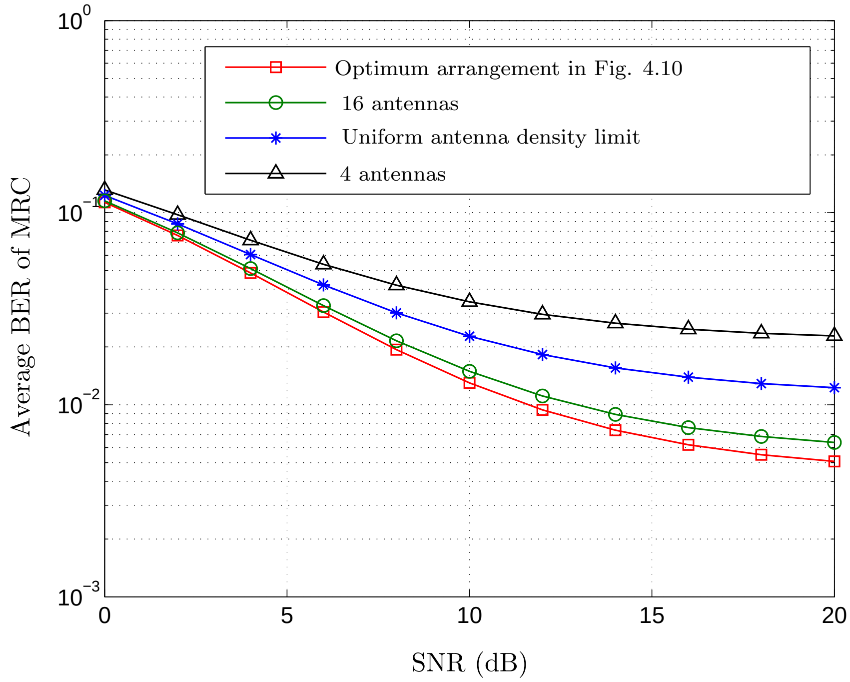


Figure 4.15. A comparison of the optimal arrangement in Example 4.3 with three suboptimal arrangements in terms of their achievable BER in the presence of CCI when used in an MRC receiver.

function of the average input SNR for the four different antenna arrangements. The graphs in Fig. 4.14 are also obtained by computer simulation. The outage threshold level is set to 5 dB, the number of interferers is $N = 3$, and the total average input SIR is 0 dB. It is clear from Fig. 4.14 that the optimum arrangement outperforms all the suboptimal designs in terms of the outage performance too. For example, at outage probability values around 10^{-3}, the optimal design outperforms the 16-antenna design and the infinite-antenna limit by 0.5 dB, and 3 dB, respectively. The last performance measure considered is the average BER of MRC in the presence of CCI. Fig. 4.15 shows the average BER of MRC in the presence of synchronous CCI as a function of the average input SNR for the four different arrangement scenarios. The graphs in Fig. 4.15 are generated using the expressions derived in Section 2.2 where the total average input SIR is set to 0 dB. One can see that the performance of the optimum arrangement is superior to those of the suboptimal designs for this performance measure as well. For example, at an input SNR level of 20 dB, the optimal design provides a BER of about 5×10^{-3} while the BERs of the 16-antenna design, the infinite-antenna limit, and the 4-antenna design are about 6×10^{-3}, 10^{-2}, and 2×10^{-2}, respectively. We observe that the bit error rate floor caused by the cochannel interference can be lowered by using the optimal number of antennas.

Note that the array design problem in Section 4.2 was originally developed based on the optimization of the average output SINR of OC (and also a similar average performance measure of MRC). However, as Figs. 4.14 and 4.15 suggest, the optimum solution based on the objective function defined in Section 4.2 performs well in terms of two other performance measures as well.

4.4 Topic for Future Research: Array Design Based on the Outage Probability of OC (Preliminary Considerations)

In Section 4.2, we used the long term average of the output SINR of OC over the fading as a measure of the array's performance and attempted to maximize it by properly arranging the antennas. This was also equivalent to maximizing some similar long term average performance measures of MRC. Furthermore, through an example design scenario, we observed that the optimum solution obtained based on these long term average performance measures works well in terms of the outage probability of OC too. Since the outage probability is a more meaningful measure of performance for a receiver working in a fading channel, one may adopt the outage probability of OC as the measure of array's performance in the first place. An optimum antenna arrangement, in this case, will be one that minimizes the outage probability. In Section 4.2, we focused on the average SINR performance of OC rather than its outage probability mainly because of the mathematical tractability of the former. As we showed, optimizing the average SINR performance of OC reduces to minimizing the Frobenius norm of the normalized array correlation matrix which can be further formulated as a convex optimization problem. On the other hand, as we will see in this section, minimizing the outage probability of OC over all possible antenna arrangements will result in an optimization problem with a much more complex objective function. Finding the global optimizer of such a problem will require the use of more sophisticated optimization techniques.

There is no closed-form expression for the outage probability of OC in the general case of correlated fading. However, asymptotic expressions have been reported in the literature [47]. In the high SNR regime, the CDF of the output SINR of OC defined in (4.5) can be approximated by its first order polynomial expansion around zero as follows [47].

$$F_\gamma(x) \approx \mathrm{E}\left\{\left|\frac{1}{M}\mathbf{G_I}\boldsymbol{\Gamma}_\mathbf{I}\mathbf{G_I}^H + \mathbf{I}_M\right|\right\} \frac{M^M}{M!|\mathbf{R_a}|}\left(\frac{x}{\Gamma_0}\right)^M \tag{4.34}$$

where $\mathbf{G_I}$ is the $M \times N$ matrix of the interferers' channels obtained by removing the first column of matrix $\mathbf{G}$ in (4.6). The diagonal matrix $\boldsymbol{\Gamma}_\mathbf{I}$ is given by $\boldsymbol{\Gamma}_\mathbf{I} = diag(\Gamma_1, \Gamma_2, \ldots, \Gamma_N)$ where the Γ_k's are the INRs defined in Section 4.1. In (4.34), $|\cdot|$ denotes matrix determinant. In the case of $N = 1$ interferer, the right side of (4.34) can be simplified to

$$F_\gamma(x) \approx (1 + \Gamma_1)\frac{M^M}{M!|\mathbf{R_a}|}\left(\frac{x}{\Gamma_0}\right)^M. \tag{4.35}$$

By definition, the SINR outage probability of OC as a function of Γ_0 and at the threshold level γ_{th} is then given by

$$P_{out}^{\gamma_{th}}(\Gamma_0) = F_\gamma(\gamma_{th}). \tag{4.36}$$

It is clear from (4.34), (4.35), and (4.36) that OC achieves the full diversity order M even in the case of correlated antennas. That is, for fixed values of γ_{th} and $\Gamma_1, \ldots, \Gamma_N$, the outage probability of OC is proportional to $(1/\Gamma_0)^M$ when Γ_0 is sufficiently larger than γ_{th}. However, this does not mean that correlation has no effect on the performance of OC. As seen in the above equations, the asymptotic outage probability is inversely proportional to the determinant of the array correlation matrix. It can be shown that $|\mathbf{R_a}|$ decreases as the antennas in the array become more and more correlated. In other words, correlation will increase the outage probability. It is easy to show that the maximum value of $|\mathbf{R_a}|$ is unity and is achieved when the array is uncorrelated, i.e., when $\mathbf{R_a} = \mathbf{I}_M$. In [47], an asymptotic expression is derived for the symbol error rate (SER) of OC as well. Similar to the outage probability, it is shown that the SER is proportional to $|\mathbf{R_a}|^{-1}(1/\Gamma_0)^M$ in the high SNR regime.

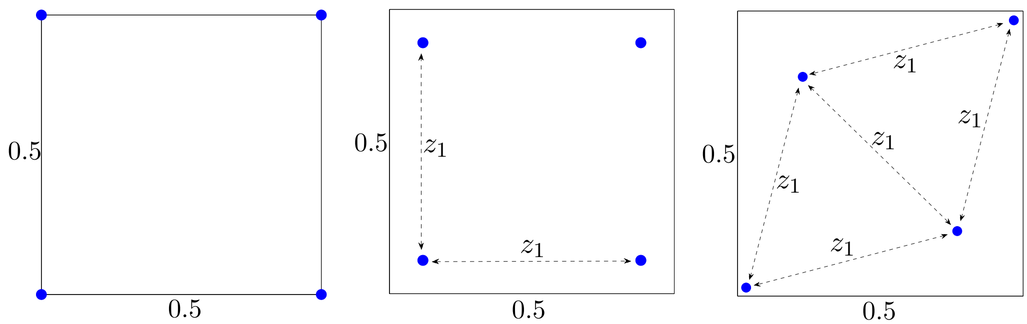
Based on the above discussion, in order to minimize the outage probability and/or the SER of a size limited OC receiver for a given value of M one has to arrange the M antennas within the available space such that the determinant of the array correlation matrix is maximized. This optimization problem in the case of a linear array of length L and for a given value of M can be formulated as

$$\begin{aligned}&\text{maximize } |\mathbf{R_a}| \text{ over } x_1, \ldots, x_M\\ &\text{where } [\mathbf{R_a}]_{(i,j)} = R_g(x_i - x_j)\\ &\text{subject to } x_1, \ldots, x_M \in [0, L]\end{aligned} \tag{4.37}$$

where the operator $[\cdot]_{(i,j)}$ returns the (i,j)'th element of its matrix argument. In the case of a 2D array, the optimization problem can be written as

$$\begin{aligned}&\text{maximize } |\mathbf{R_a}| \text{ over } (x_1, y_1), \ldots, (x_M, y_M)\\ &\text{where } [\mathbf{R_a}]_{(i,j)} = R_g\left(x_i - x_j, y_i - y_j\right)\\ &\text{subject to } (x_1, y_1), \ldots, (x_M, y_M) \in \mathcal{D}\end{aligned} \tag{4.38}$$

where $\mathcal{D}$ is a subset of the plane. Solving the optimization problems defined in (4.37) and (4.38) numerically and perhaps analytically for some special cases is a potential subject for future research work. Note that, contrary to the case of the average SINR optimization discussed in Chapter 4, it is not clear whether the problem of SINR outage optimization can be formulated as a convex optimization problem. In generall, the objective functions in (4.37) and (4.38) may have many local maxima within the allowed ranges of their variables. Therefore, in order to solve the problem numerically, one has to resort to global optimization techniques such as simulated annealing, particle swarm optimization, or genetic algorithms [48] since any local optimization method may get trapped in the local maxima. Now, one may ask whether the performance enhancement achieved by optimal arrangement of the antennas is worth the computational complexity of such optimization methods. In other words, the question is how much benefit one can get by deploying the optimal arrangement over a trivial arrangement, e.g., a uniform distribution of the

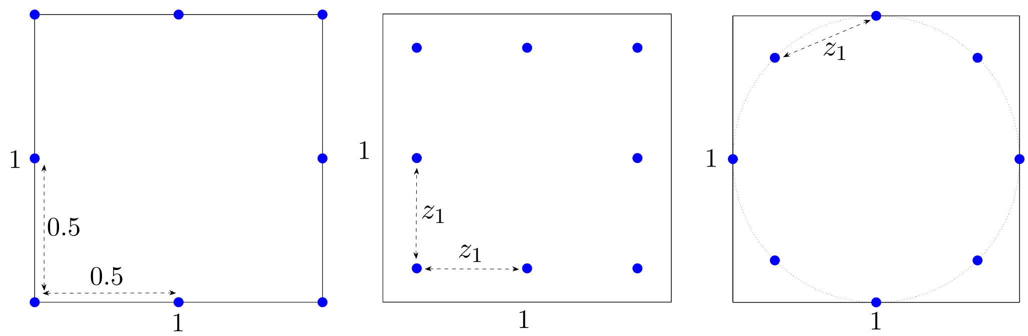


(a) $|\mathbf{R_a}| = 0.132$ (b) $|\mathbf{R_a}| = 0.752$ (c) $|\mathbf{R_a}| = 0.855$

Figure 4.16. Different antenna arrangements discussed in Example 4.6.

antennas within the available space. The following two examples show that a significant improvement may be achieved in the case of a 2D array.

Example 4.6. Consider a 2D-omnidirectional scattering model where the spatial correlation function is given by $R_g(x, y) = \mathrm{J}_0(2\pi\sqrt{x^2 + y^2})$. Assume that a square area of normalized size 0.5 by 0.5 (i.e., the side length is equal to one-half of a wavelength) is available on the receiver and we are going to place $M = 4$ antennas on this surface. Three different antenna arrangements are shown in Fig. 4.16. The determinant of the corresponding array correlation matrix is also shown under each arrangement. Note that $\mathbf{R_a}$, in this example, is a 4 by 4 matrix whose elements are defined in (4.38). In the first arrangement, shown in Fig. 4.16(a), the antennas are located on the 4 corners of the available square area. In the second scenario, shown in Fig. 4.16(b), the antennas form a smaller square within the given area. The side length of this square is equal to z_1 where $z_1 = 0.3827$ is the first positive zero of $\mathrm{J}_0(2\pi x)$. Finally, in the third scenario, shown in Fig. 4.16(c), the 4 antennas are located on the vertices of the fundamental parallelotope of a hexagonal lattice with minimum distance z_1. It can be seen that the arrangement in the third scenario exhibits the largest determinant among the three arrangements in Fig. 4.16. Based on the asymptotic expressions for the CDF (equivalently, the outage probability) of the SINR in (4.34) and (4.35), the quantity $|\mathbf{R_a}|^{1/M}$ can be thought as the array SNR gain. Therefore, for a fixed value of M, the SNR advantage of a particular antenna arrangement over another arrangement can



(a) $|\mathbf{R_a}| = 0.0068$ (b) $|\mathbf{R_a}| = 0.0355$ (c) $|\mathbf{R_a}| = 0.3374$

Figure 4.17. Different antenna arrangements discussed in Example 4.7.

be calculated, in dB units, as the difference between their corresponding values of $\frac{10}{M}\log_{10}(|\mathbf{R_a}|)$. Consequently, the arrangement in Fig. 4.16(c) provides a SNR gain of $\frac{10}{4}(\log_{10}(0.855) - \log_{10}(0.132)) = 2.03$ dB over the arrangement in Fig. 4.16(a). This means that for a desired outage probability in the high SNR regime, the required SNR for the arrangement in the third scenario is about 2 dB smaller than that of the arrangement in the first scenario.

Example 4.7. In this example, we consider a spatial correlation model similar to that of Example 4.6. However, we assume that a square area of normalized size 1 by 1 is available on the receiver where $M = 8$ antennas are to be placed. Fig. 4.17 shows three different antenna arrangements. The determinant of the corresponding array correlation matrix is again shown under each arrangement. In the first scenario, shown in Fig. 4.17(a), the antennas are uniformly located on the perimeter of the square region. In the second scenario, shown in Fig. 4.17(b), the antennas are located on the perimeter of a smaller square where the normalized distance between any two adjacent antennas is equal to z_1 (z_1 is defined in Example 4.6). In the third scenario, shown in Fig. 4.17(c), the antennas are located on the perimeter of a circle where the normalized distance between any two adjacent antennas is equal to z_1. One can see that the antenna arrangement in Fig. 4.17(c) outperforms the one in Fig. 4.17(a) by a SNR gain of $\frac{10}{8}(\log_{10}(0.3374) - \log_{10}(0.0068)) = 2.12$ dB.

Although none of the different antenna arrangements discussed in Example 4.6 and Example 4.7 are obtained by solving the optimization problem in

(4.38), they suggest that a SNR loss of up to at least 2 dB can be incurred if the antennas are not placed optimally. This justifies the importance of the search for the optimal antenna arrangement despite the relatively high computational complexities associated with the optimization algorithms.

4.5 Chapter Summary

In this chapter, the problem of optimizing the antenna arrangement for a size limited multiple antenna receiver was formulated as an infinite dimensional quadratic convex programming. The formulation was originally established for a 1D linear array and then generalized to the case of a 2D array with arbitrary shape. The derived infinite dimensional problem was solved analytically for the special case of linear arrays and under an exponential correlation model. In the general case, it was shown through several numerical examples that the infinite dimensional optimization problem can be solved accurately and efficiently by approximating it with a finite dimensional problem. Moreover, the achievable performance of the optimal antenna arrangement in a 2D array example was compared with those of a few suboptimal array designs. The comparison was made based on three different performance measures, namely, the average output SINR of OC, the outage probability of OC, and the average BER of MRC in the presence of CCI. Although the optimal solution was originally obtained by maximizing the output SINR of OC, it was shown that it outperforms all the suboptimal designs in terms of the other two performance measures as well. Finally, adopting the outage probability and/or SER of OC as the measure of array's performance in the design of size limited arrays was proposed as a possible direction for future research. It was shown that, unlike the array optimization based on the average SINR of OC, the array optimization based on the outage probability of OC results in a much more complicated objective function whose global optimum solution cannot be easily found. However, it was shown through a number of numerical examples that solving this new optimization problem can potentially lead to a significant gain in terms of the asymptotic outage probability at high SNR values.

Chapter 5

Concluding Remarks

The main contributions of the thesis are summarized in this chapter. In Chapter 2, new simple closed-form expressions for the average BER of BPSK using MRC in the presence of correlated Rayleigh fading, CCI, and AWGN were derived. The final solutions were explicitly expressed in terms of the eigenvalues of the correlation matrices of the desired user's channel and that of the interfering users' channels and are valid for an arbitrary number of interfering users and an arbitrary number of receiver antennas. The derived expressions were then used to calculate the BER performance of MRC systems in some example scenarios. Particularly, it was shown that for a fixed number of antennas in a uniform circular array, there is a best radius for the array corresponding to the smallest uniform circular array that almost achieves its best possible BER performance, as if its radius was infinite (i.e., the totally uncorrelated array).

Motivated by the application of multiple antenna receivers on small size mobile units in cellular systems, in Chapter 3, the asymptotic performance of OC was investigated as the number of antennas in the array increases while the total physical dimensions of the array are fixed. Two different scenarios were considered, namely, fixed average received power per antenna and fixed total average received power. It was shown that in the former scenario, the average output SINR is asymptotically a linear function of the number of the antennas while in the latter scenario it eventually saturates at a certain value. In the case of a single interferer, the slope of the asymptote in the former scenario as well as the value of the saturation limit in the latter scenario were derived in terms of the point spectrum of the underlying array correlation function. It was shown through a numerical example that the receiver actually exhibits its asymptotic behavior for a practically small number of antennas.

In Chapter 4, we turned our focus to an array design problem where we attempted to maximize the achievable performance of a size constrained multiple antenna receiver by properly arranging the antennas within the array. Using the average output SINR of OC as a measure of the array's performance, it was shown that the problem of optimizing the antenna arrangement for a size limited multiple antenna receiver can be formulated as an infinite dimensional quadratic convex programming. The formulation was originally established for a 1D linear array and then generalized to the case of a 2D array with arbitrary shape. The derived infinite dimensional problem was solved analytically for the special case of linear arrays and under exponential correlation model. In the general case, it was shown through several numerical examples that the infinite dimensional optimization problem can be solved accurately and efficiently by approximating it with a finite dimensional problem. Moreover, the achievable performance of the optimal antenna arrangement in a 2D array example was compared with those of a few suboptimal array designs. The comparison was made based on three different performance measures, namely, the average output SINR of OC, the outage probability of OC, and the average BER of MRC in the presence of CCI. Although the optimal solution was originally obtained by maximizing the output SINR of OC, it was shown that it outperforms the suboptimal designs in terms of the other two performance measures as well. Finally, adopting the outage probability and/or SER of OC as the measure of array's performance in the the design of size limited arrays was proposed as a possible direction for future research.

Bibliography

[1] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.

[2] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*. New Jersey: John Wiley & Sons, 2005.

[3] J. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE Trans. Veh. Technol.*, vol. 33, pp. 144–155, Aug. 1984.

[4] A. Haimovich, A. Shah, and X. Wu, "Reduced-rank array processing for wireless communications with applications to IS-54/IS-136," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 743–747, May 2000.

[5] W. C. Jakes, *Microwave Mobile Communications*. Piscataway, NJ: IEEE Press, 1974.

[6] A. Shah and A. M. Haimovich, "Performance analysis of maximal ratio combining and comparison with optimum combining for mobile radio communications with cochannel interference," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1454–1463, Jul. 2000.

[7] X. Zhang and N. C. Beaulieu, "Outage probability of MRC with equipower cochannel interferers in correlated Rayleigh fading," *IEEE Commun. Lett.*, vol. 9, pp. 577–579, Jul. 2005.

[8] X. W. Cui, Q. T. Zhang, and Z. M. Feng, "Outage performance for maximal ratio combiner in the presence of unequal-power co-channel interferers," *IEEE Commun. Lett.*, vol. 8, no. 5, pp. 289–291, May 2004.

[9] X. Zhang and N. C. Beaulieu, "Outage probability of MRC with unequal-power cochannel interferers in correlated Rayleigh fading," *IEEE Commun. Lett.*, vol. 10, pp. 7–9, Jan. 2006.

[10] Y. Ma, T. J. Lim, and S. Pasupathy, "Error probability for coherent and differential PSK over arbitrary rician fading channels with multiple cochannel interferers," *IEEE Trans. Commun.*, vol. 50, pp. 429–441, Mar. 2002.

[11] Y. Ma and T. J. Lim, "Bit error probability for MDPSK and NCFSK over arbitrary Rician fading channels," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2179–2189, Nov. 2000.

[12] E. Kreyszig, *Advanced Engineering Mathematics*, 9th ed. John Wiley & Sons, 2006.

[13] X. Zhang and N. C. Beaulieu, "Closed-form BER expressions for BPSK using MRC with multiple interferers in arbitrarily correlated Rayleigh fading," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, Scotland, Jun. 24–28 2007, pp. 6295–6298.

[14] X. Zhang and N. C. Beaulieu, "A closed-form BER expression for BPSK using MRC in correlated CCI and Rayleigh fading," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2249–2252, Dec. 2007.

[15] M. R. McKay, A. Zanella, I. B. Collings, and M. Chiani, "Error probability and SINR analysis of optimum combining in Rician fading," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 676–687, Mar. 2009.

[16] H. Kang, J. S. Kwak, T. G. Pratt, and G. L. Stuber, "Analytical framework for optimal combining with arbitrary-power interferers and thermal noise," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1564–1575, May 2008.

[17] J. S. Kwak and J. H. Lee, "Closed-form expressions of approximate error rates for optimum combining with multiple interferers in a Rayleigh fading channel," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 158–166, Jan. 2006.

[18] M. Chiani, M. Z. Win, A. Zanella, and J. H. Winters, "A Laguerre polynomial-based bound on the symbol error probability for adaptive antennas with optimum combining," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 12–16, Jan. 2004.

[19] A. Shah, A. M. Haimovich, M. K. Simon, and M. S. Alouini, "Exact bit-error probability for optimum combining with a Rayleigh fading Gaussian cochannel interferer," *IEEE Trans. Commun.*, vol. 48, no. 6, pp. 908–912, Jun. 2000.

[20] D. Yue, X. Wang, and F. Xu, "Performance analysis for optimum combining of Rayleigh fading signals with correlated Rayleigh interferers and noise," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 269–272, May 2006.

[21] Q. T. Zhang and X. W. Cui, "Outage probability for optimum combining of arbitrarily faded signals in the presence of correlated Rayleigh interferers," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1043–1051, Jul. 2004.

[22] R. K. Mallik and Q. T. Zhang, "Optimum combining with correlated interference," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2340–2348, Sep. 2005.

[23] A. Shah and A. M. Haimovich, "Performance analysis of optimum combining in wireless communications with Rayleigh fading and cochannel interference," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 473–479, Apr. 1998.

[24] N. C. Beaulieu and X. Zhang, "On the maximum number of receiver diversity antennas that can be usefully deployed in a cochannel interference dominated environment," *IEEE Trans. Signal Process.*, vol. 55, pp. 3349–3359, Jul. 2007.

[25] N. C. Beaulieu and X. Zhang, "On the maximum useful number of receiver antennas for MRC diversity in cochannel interference and noise," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, Jun. 2007, pp. 5103–5108.

[26] T. S. Pollock, T. D. Abhayapala, and R. A. Kennedy, "Antenna saturation effects on MIMO capacity," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 4, May 11–15, 2003, pp. 2301–2305.

[27] K. Cho, Y. Takatori, K. Nishimori, and T. Hori, "Optimum element arrangement of adaptive base station antennas for SDMA in circular cell," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, vol. 1, no. 52, Sep. 2000, pp. 150–156.

[28] I. Haya, B. Petersen, and B. Colpitts, "Optimum 2-D LOS MIMO performance using omni-directional antennas attained through genetic algorithms," in *Communication Networks and Services Research Conference (CNSR 2008)*, May 2008, pp. 331–338.

[29] K. Ishizawa, J. Zhou, S. Sasaki, S. Muramatsu, and H. Kikuchi, "Spatial correlation of a circular array antenna and BER performance investigation," in *Circuits and Systems, 2004. Proceedings. The 2004 IEEE Asia-Pacific Conf. on*, vol. 1, Dec. 6–9, 2004, pp. 385–388.

[30] P. D. Teal, T. D. Abhayapala, and R. A. Kennedy, "Spatial correlation for general distributions of scatterers," *IEEE Signal Process. Lett.*, vol. 9, no. 10, pp. 305–308, Oct. 2002.

[31] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins University Press, 1996.

[32] N. C. Beaulieu and Q. Xie, "An optimal lognormal approximation to lognormal sum distributions," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 479–489, Mar. 2004.

[33] A. A. Basri and T. J. Lim, "Exact average bit-error probability for maximal ratio combining with multiple cochannel interferers and Rayleigh fading," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, Scotland, Jun. 24–28 2007, pp. 1102–1107.

[34] R. H. Clarke and W. L. Khoo, "3-D mobile radio channel statistics," *IEEE Trans. Veh. Technol.*, vol. 46, no. 3, pp. 798–799, Aug. 1997.

[35] S. Wei, D. Goeckel, and R. Janaswamy, "On the asymptotic capacity of MIMO systems with antenna arrays of fixed length," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1608–1621, Jul. 2005.

[36] S. Sorooshyari and D. G. Daut, "Bit error rate analysis of maximal-ratio combining over correlated Gaussian vector channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1128–1133, Apr. 2008.

[37] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*, 4th ed. McGraw-Hill, 2002.

[38] C. T. H. Baker, *The Numerical Treatment of Integral Equations.* Oxford, UK: Clarendon Press, 1977.

[39] F. Riesz and B. Nagy, *Functional Analysis.* New York: Dover Publications, 1990.

[40] P. O. Cubillos, "Eigenvalue problems for Fredholm integral operators," *IMA Journal of Numerical Analysis*, vol. 7, no. 2, pp. 191–204, Apr. 1987.

[41] M. Reed and B. Simon, *Methods of Modern Mathematical Physics.* Academic Press, 1975, vol. 2.

[42] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 271–286, Jan. 2006.

[43] S. Boyd and L. Vandenberghe, *Convex Optimization.* Camebridge University Press, 2004.

[44] O. Lieberman, "A Laplace approximation to the moments of a ratio of quadratic forms," *Biometrika*, vol. 81, no. 4, pp. 681–690, Dec. 1994.

[45] T. M. Apostol, *Mathematical Analysis*, 2nd ed. Addison Wesley, 1974.

[46] I. J. Schoenberg, "Metric spaces and positive definite functions," *Trans. of the American Mathematical Society*, vol. 44, no. 3, pp. 522–536, Nov. 1938.

[47] R. Louie, M. McKay, and I. Collings, "New asymptotic performance results for MIMO and SIMO optimum combining," in *Proc. Australian Communications Theory Workshop (AusCTW 2008)*, Jan. 2008, pp. 69–74.

[48] E. M. Hendrix and B. G.-Tóth, *Introduction to Nonlinear and Global Optimization*, ser. Springer Optimization and Its Applications. Springer New York, 2010, vol. 37.

[49] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[50] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1426–1438, Aug. 2006.

Appendix A

Proof of Lemma 3.1

Proof. If $\lim_{M\to\infty} \mathcal{N}_a(M) < \infty$, the number of summands in (3.19) is finite and the proof is straightforward. The proof for the general case is as follows.

$$\begin{aligned} \forall i \in \{1, \ldots, M\} : \quad |f(\alpha_{i,M}) - f(\tilde{\alpha}_i)| = \left| \int_{\alpha_{i,M}}^{\tilde{\alpha}_i} f'(x)dx \right| &\leq B \left|\alpha_{i,M} - \tilde{\alpha}_i\right| \\ &\leq B\frac{C}{M} \end{aligned} \tag{A.1}$$

where the last inequality is a result of the uniform error bound discussed in Section 3.3.1. Now, one can write

$$\begin{aligned} \left| \sum_{i=1}^{\mathcal{N}_a(M)} \left[f(\alpha_{i,M}) - f(\tilde{\alpha}_i)\right] \right| &\leq \sum_{i=1}^{\mathcal{N}_a(M)} \left|f(\alpha_{i,M}) - f(\tilde{\alpha}_i)\right| \\ &\leq BC\frac{\mathcal{N}_a(M)}{M} \end{aligned} \tag{A.2}$$

where the last inequality is a result of (A.1). Using (A.2) and recalling from Section 3.3.1 that $\mathcal{N}_a(M)$ is $o(M)$ one obtains

$$\lim_{M\to\infty} \left| \sum_{i=1}^{\mathcal{N}_a(M)} \left[f(\alpha_{i,M}) - f(\tilde{\alpha}_i)\right] \right| = 0 \tag{A.3}$$

and the proof is complete. □

Appendix B

Proof of The Existence of Two Limits Used in Section 3.3.2

In this appendix, it is shown that $\lim_{M\to\infty} \sum_{i=1}^{\mathcal{N}_a(M)} \pi_s(\tilde{\alpha}_i)$ and $\lim_{M\to\infty} \sum_{i=1}^{\mathcal{N}_a(M)} \sigma_s(\tilde{\alpha}_i)$ exist.

Proof. If $\lim_{M\to\infty} \mathcal{N}_a(M) < \infty$, the proof is straightforward. In the general case, one has

$$\sum_{i=1}^{\mathcal{N}_a(M)} \pi_s(\tilde{\alpha}_i) = \sum_{i=1}^{\mathcal{N}_a(M)} \ln(1 + s\tilde{\alpha}_i) \leqslant s \sum_{i=1}^{\mathcal{N}_a(M)} \tilde{\alpha}_i \tag{B.1a}$$

$$\leqslant s \sum_{i=1}^{\infty} \tilde{\alpha}_i \tag{B.1b}$$

$$= s \tag{B.1c}$$

and

$$\sum_{i=1}^{\mathcal{N}_a(M)} \sigma_s(\tilde{\alpha}_i) = \sum_{i=1}^{\mathcal{N}_a(M)} \frac{\tilde{\alpha}_i^2}{1 + s\tilde{\alpha}_i} \leqslant \sum_{i=1}^{\mathcal{N}_a(M)} \tilde{\alpha}_i^2 \tag{B.2a}$$

$$\leqslant \sum_{i=1}^{\infty} \tilde{\alpha}_i^2 \tag{B.2b}$$

$$\leqslant 1. \tag{B.2c}$$

The inequality in (B.1a) comes from the inequality $\ln(1 + x) \leqslant x$ [49] and the equality in (B.1c) is based on (3.17). Also, the inequality in (B.2c) is an immediate result of (3.13) and (3.18). Now, Since the series in (B.1) and (B.2) are both bounded from above and noting that $\pi_s(\tilde{\alpha}_i) \geqslant 0$ and $\sigma_s(\tilde{\alpha}_i) \geqslant 0$ for all i's, one obtains that both series are convergent and the proof is complete. □

Appendix C

Proof of Theorem 3.2

Proof. Define $\mathcal{U}(s)$ as follows

$$\mathcal{U}(s) = \begin{cases} \dfrac{1}{(1+\frac{\tilde{\alpha}_1}{2}s)s}, & s > s_0 \\ 1, & 0 \leqslant s \leqslant s_0 \end{cases} \tag{C.1}$$

where s_0 is some positive number. It is easy to see that the integral of $\mathcal{U}(s)$ over $[0, \infty)$ is finite. Now, recalling the expression for $\mathcal{I}_M(s)$ in (3.12b), one has

$$\mathcal{I}_M(s) \leqslant \sum_{i=1}^{M} \alpha_{i,M}^2 \leqslant \sum_{i=1}^{M} \alpha_{i,M} = 1. \tag{C.2}$$

Moreover, since $\alpha_{1,M}$ approaches $\tilde{\alpha}_1$ as M grows, for a sufficiently large M one has $\alpha_{1,M} > \tilde{\alpha}_1/2$. Hence, using the definition of $\mathcal{I}_M(s)$ in (3.12b), one can write

$$\mathcal{I}_M(s) < \frac{1}{1+\frac{\tilde{\alpha}_1}{2}s}\left(\frac{1}{s}\sum_{i=1}^{M}\alpha_{i,M}\right) = \frac{1}{(1+\frac{\tilde{\alpha}_1}{2}s)s}. \tag{C.3}$$

From (C.2) and (C.3), one obtains $\mathcal{I}_M(s) \leqslant \mathcal{U}(s)$ for any $M \geqslant 1$ and $s \in [0, \infty)$. Also, as shown in Section 3.3.2, the sequence of functions $\{\mathcal{I}_M(s)\}_{M=1}^{\infty}$ has a limit $\mathcal{I}(s)$ given in (3.23). Applying the Lebesgue's dominated convergence theorem [45, p. 270] completes the proof. □

Appendix D

Proof of Theorem 3.3

Proof. From (3.12) and recalling that $\alpha_{1,M} = \max(\{\alpha_{i,M}\}_{i=1}^{M})$, one has

$$\begin{aligned}
\mathrm{E}\left\{\frac{X}{Y}\right\} &\leqslant \alpha_{1,M} \int_0^\infty \left(\prod_{i=1}^{M} \frac{1}{1+\alpha_{i,M}s}\right)\left(\sum_{i=1}^{M} \frac{\alpha_{i,M}}{1+\alpha_{i,M}s}\right) ds \\
&= \alpha_{1,M} \left[\prod_{i=1}^{M} \frac{-1}{1+\alpha_{i,M}s}\right]_0^\infty \\
&= \alpha_{1,M}.
\end{aligned} \tag{D.1}$$

Thus,

$$\lim_{M\to\infty} \mathrm{E}\left\{\frac{X}{Y}\right\} \leqslant \lim_{M\to\infty} \alpha_{1,M} = \tilde{\alpha}_1. \tag{D.2}$$

Now, if we show that $\tilde{\alpha}_1 < 1$, the proof is complete. Let us assume this is not the case. Then $\tilde{\alpha}_1 = 1$ (since one cannot have $\tilde{\alpha}_1 > 1$). But, from (3.17), we must have $\tilde{\alpha}_i = 0$ for $i \geqslant 2$. Hence, (3.16) reduces to $R_a(x,y) = q_1(x)q_1(y)^*$ which contradicts the nonseparability of $R_a(x,y)$ and Theorem 3.3 is proved by contradiction. □

Appendix E

Proof of Conjecture 4.1 for $M = 2$

Proof. If we consider the expression inside the expectation operator in (4.11) as a function $g(z_1, z_2, \ldots, z_M)$ of the z_i's, the right side of (4.11) is equal to $g(\mathrm{E}(z_1), \mathrm{E}(z_2), \ldots, \mathrm{E}(z_M))$, provided that $\mathrm{E}\{z_i\}$ exists. This suggests the possibility of the use of Jensen's inequality [43, p. 77] as a method of proof. However, it cannot be applied directly since $g(z_1, z_2, \ldots, z_M)$ is not a concave function of the z_i's for $M > 1$. But, in the case of $M = 2$, it is possible to convert it to a concave function $h(t)$ as follows.

$$g(z_1, z_2) = \frac{a_1^2 z_1 + a_2^2 z_2}{a_1 z_1 + a_2 z_2} = h(t) \tag{E.1a}$$

where

$$t \triangleq \frac{z_1 - z_2}{z_1 + z_2} \tag{E.1b}$$

and

$$h(t) \triangleq \frac{a_1^2 + a_2^2 + (a_1^2 - a_2^2)t}{a_1 + a_2 + (a_1 - a_2)t}, \qquad t \in (-1, 1). \tag{E.1c}$$

Since z_1 and z_2 are positive random variables, the absolute value of the new variable t defined in (E.1b) is always less than 1. Thus, the domain of $h(t)$ is $(-1, 1)$. The function $h(t)$ is concave over its domain since its second derivative, given by

$$h''(t) = \frac{-4a_1 a_2 (a_1 - a_2)^2}{(a_1 + a_2 + (a_1 - a_2)t)^3} \tag{E.2}$$

is negative for any $t \in (-1, 1)$. Therefore, we can apply Jensen's inequality to $h(t)$ to get

$$\mathrm{E}\{h(t)\} \leq h(\mathrm{E}\{t\}). \tag{E.3}$$

But, the expected value of t can be calculated as

$$\mathrm{E}\{t\} = \mathrm{E}\left\{\frac{z_1 - z_2}{z_1 + z_2}\right\} = \mathrm{E}\left\{\frac{z_1}{z_1 + z_2}\right\} - \mathrm{E}\left\{\frac{z_2}{z_1 + z_2}\right\} = 0 \qquad \text{(E.4)}$$

where the last equality is due to the symmetry arising from the fact that z_1 and z_2 are identically distributed. The proof is immediate using (E.1a), (E.1c),(E.3), and (E.4).

Note that, we did not make use of independence of the random variables z_1 and z_2 in the proof. However, the independence condition seems to be necessary for the conjecture to be true in the case of $M > 2$. $\square$

Appendix F

Proof of Theorem 4.2

Proof. Using the method of Lagrange multipliers the objective functional can be written as

$$\mathcal{J}(p) = \int_0^L \int_0^L f(x-x')p(x)p(x')dxdx' - 2\int_0^L \lambda(x)p(x)dx - 2\nu \int_0^L p(x)dx \tag{F.1}$$

where $\lambda(x)$ is a positive function over $[0, L]$ whose value at point x represents the Lagrange multiplier corresponding to the inequality constraint $p(x) \geq 0$ at point x. The parameter ν is the Lagrange multiplier corresponding to the equality constraint $\int_0^L p(x)dx = 1$. The multiplication by 2 of the constraint terms in (F.1) is just a normalization of the Lagrange multipliers that has no effect on the solution and is merely intended to simplify the form of the expressions to be derived in the following. The next step is to find the gradient (a.k.a. the variational derivative) of the objective functional in (F.1) with respect to variations in function $p(\cdot)$. The gradient of $\mathcal{J}(p)$ is defined through the usual directional derivative formula

$$\langle \nabla \mathcal{J}(p), v \rangle = \frac{d}{dt}\mathcal{J}(p+tv)\big|_{t=0} \tag{F.2}$$

where $v(\cdot)$ can be any function over $[0, L]$ and the inner product $\langle \cdot, \cdot \rangle$ for a pair of functions $f(\cdot)$ and $g(\cdot)$ is defined by

$$\langle f, g \rangle = \int_0^L f(x)g(x)dx. \tag{F.3}$$

The function $v(\cdot)$ represents an arbitrary variation in function $p(\cdot)$ and the inner product $\langle \nabla \mathcal{J}(p), v \rangle$ represents the variation of the functional $\mathcal{J}(p)$ in the direction of $v(\cdot)$. After replacing $p(x)$ and $p(x')$ in (F.1) with $p(x)+tv(x)$ and $p(x')+tv(x')$, respectively, and applying (F.2), one obtains

$$\langle \nabla \mathcal{J}(p), v \rangle = \int_0^L \int_0^L f(x-x')(p(x)v(x')+v(x)p(x'))dxdx' - 2\int_0^L (\lambda(x)+\nu)v(x)dx. \tag{F.4}$$

But, note that $f(x) = |R_g(x)|^2$ where $R_g(x)$ is a conjugate symmetric function as a result of being a correlation function. Therefore, $f(x)$ is an even function (i.e., $f(x) = f(-x)$). Thus, (F.4) can be further simplified to

$$\langle \nabla \mathcal{J}(p), v \rangle = \int_0^L 2\left[\int_0^L f(x-x')p(x')dx' - \lambda(x) - \nu\right] v(x)dx. \tag{F.5}$$

Using the definition of the inner product given in (F.3) and the fact that (F.5) holds for all functions $v(\cdot)$ over $[0, L]$, one can conclude

$$\nabla \mathcal{J}(p) = 2\left[\int_0^L f(x-x')p(x')dx' - \lambda(x) - \nu\right]. \tag{F.6}$$

The Karush-Kuhn-Tucker (KKT) conditions of optimality are given by [50]

$$\begin{aligned} &\int_0^L p(x)dx = 1 \\ &p(x) \geq 0, && \forall x \in [0, L] \\ &\lambda(x) \geq 0, && \forall x \in [0, L] \\ &\lambda(x)p(x) = 0, && \forall x \in [0, L] \\ &\nabla \mathcal{J}(p) = 0, && \forall x \in [0, L] \end{aligned} \tag{F.7}$$

where $\nabla \mathcal{J}(p)$ is given in (F.6). Next, we show that, for an exponential spatial correlation function where $f(x) = |R_g(x)|^2 = \exp(-2k|x|)$, the function $p_{opt}(x)$ given in (4.25) along with $\lambda_{opt}(x) = 0$ and $\nu = \frac{1}{1+kL}$ satisfy the KKT conditions in (F.7). Checking the first four conditions in (F.7) is straightforward. Moreover, using the expression for $\nabla \mathcal{J}(p)$ in (F.6), the last condition in (F.7) is equivalent to

$$\int_0^L e^{-2k|x-x'|}\left[\frac{k}{1+kL} + \frac{1}{2+2kL}\left(\delta(x') + \delta(x'-L)\right)\right]dx' = \frac{1}{1+kL}, \ \forall x \in [0, L] \tag{F.8}$$

which can be easily proved by direct calculation of the integral on the left side. The KKT conditions are, in general, necessary conditions for the optimality of a solution. However, since the optimization problem in (4.22) is convex and its constraints satisfy the Slater's constraint qualification condition [43], the KKT conditions are also sufficient for optimality. This completes the proof. $\square$

Appendix G

Proof of Theorem 4.3

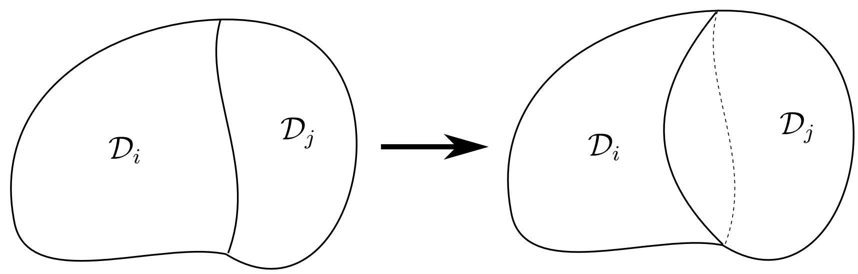


Figure G.1. Passing weight from the region $\mathcal{D}_i$ to its neighboring region $\mathcal{D}_j$.

Proof. First, we assume that $\mathcal{G}$ is a connected graph. The case of a disconnected graph is considered afterwards. Define a weight function $w(\cdot)$ over the nodes of $\mathcal{G}$ where $w(\mathcal{D}_i)$ is the area of the region $\mathcal{D}_i$, $i = 1, \ldots, M$. A node $\mathcal{D}_i$ can *send* a portion $w_{i,j}$ of its weight to a neighbouring node $\mathcal{D}_j$. This will update the weight of $\mathcal{D}_i$ to $w(\mathcal{D}_i) - w_{i,j}$ and that of $\mathcal{D}_j$ to $w(\mathcal{D}_j) + w_{i,j}$. The equivalent of this weight passing operation on the partition regions is to move the common boundary of $\mathcal{D}_i$ and $\mathcal{D}_i$ toward the interior of $\mathcal{D}_i$ such that the area of $\mathcal{D}_i$ is reduced by $w_{i,j}$ and that of $\mathcal{D}_j$ is increased by $w_{i,j}$ as shown in Fig. G.1. The two end points of the common boundary curve are kept fixed during this operation. This weight passing operation does not change the partition graph $\mathcal{G}$, i.e., the adjacent nodes remain adjacent and the non-adjacent nodes remain non-adjacent. Since $\mathcal{G}$ is connected, any two nodes in the graph can pass weights to each other through some intermediate nodes. Let $\mathcal{V}_0 = \{\mathcal{D}_1, \ldots, \mathcal{D}_M\}$ denote the set of the vertices (i.e., nodes) of $\mathcal{G}$ and define $\bar{w}$ to be the arithmetic mean of the set $w\{\mathcal{V}_0\} = \{w(\mathcal{D}_1), \ldots, w(\mathcal{D}_M)\}$.

Executing the following weight passing algorithm will result in a partitioning of $\mathcal{D}$ into equal area $\mathcal{D}_i$'s while the corresponding partition graph $\mathcal{G}$ remains unaltered.

1: $\mathcal{V} \leftarrow \mathcal{V}_0$
2: $i \leftarrow \text{max\_index}\{w(\mathcal{V})\}$
3: $j \leftarrow \text{min\_index}\{w(\mathcal{V})\}$
4: Send $\bar{w} - w(\mathcal{D}_j)$ from Node $\mathcal{D}_i$ to Node $\mathcal{D}_j$
5: $\mathcal{V} \leftarrow \mathcal{V} - \{\mathcal{D}_j\}$
6: If $|\mathcal{V}| > 1$ then go to Step 2
7: return

The max_index (min_index) function returns the index of the maximum (minimum) element of the set in its argument and $|\mathcal{V}|$ denotes the number of elements in the set $\mathcal{V}$. Once the above algorithm is completed, the $\mathcal{T}_i$ regions can be obtained by scaling the $\mathcal{D}_i$ regions with a common scaling factor such that their areas add up to unity. Now, we consider the case where $\mathcal{G}$ is a disconnected graph. In this case, $\mathcal{G}$ can be broken into a number of connected subgraphs $\mathcal{G}_k$'s, $k = 1, \ldots, r$ for some number r. Since each $\mathcal{G}_k$ is a connected graph, one can execute the above algorithm over it. After executing the algorithm for all of the subgraphs, the partition corresponding to each $\mathcal{G}_k$ will have equal area regions. Therefore, one can obtain the $\mathcal{T}_i$ regions by properly scaling the partitions corresponding to different $\mathcal{G}_k$s with different scaling factors such that all the regions in all partitions have the same area and the sum of their areas is unity. This completes the proof. $\square$