A Random Matrix Model of Communication Via Antenna Arrays

Ralf R. Müller, *Associate Member, IEEE*

***Abstract*—A random matrix model is introduced that probabilistically describes the spatial and temporal multipath propagation between a transmitting and receiving antenna array with a limited number of scatterers for mobile radio and indoor environments. The model characterizes the channel by its *richness delay profile* which gives the number of scattering objects as a function of the path delay. Each delay is assigned the eigenvalue distribution of a random matrix that depends on the number of scatterers, receiving antennas, and transmitting antennas. The model allows one to calculate signal-to-interference-and-noise ratios (SINRs) and channel capacities for large antenna arrays analytically and quantifies to what extent rich scattering improves performance.**

***Index Terms*—Antenna arrays, channel models, eigenvalue distribution, fading channels, information rates, intersymbol interference, multiuser communications, random matrices, Stieltjes transform.**

I. Introduction

COMMUNICATION via antenna arrays allows a significant increase in spectral efficiency—the information rate per communication link [1], [2]. While several recent proposals [1], [3]–[5] aim to utilize this advantage, it is still not sufficiently understood how the physical properties of these channels translate into achievable *signal-to-interference-and-noise ratio*s (SINRs) and therefore the supported information rates. On the physical side, channel models are based on propagation measurements. They provide statistics of the propagation between a pair of transmitter arrays and receiver arrays in terms of delays, received powers, and directions of arrival and departure. Statements about the information rates capable in the channel, however, are given in terms of the eigenvalues related to the matrix-algebraic description of the communication link. This work aims to build a bridge between propagation scenarios and the eigenvalues of the covariance matrices of the channel in order to allow for predictions of channel capacity based on the morphology of the physical medium.

It is natural to describe a linear time-invariant multiple-input–multiple-output (MIMO) system by its matrix-valued impulse response. The matrix-valued taps of the impulse response of the antenna array channel depend on various parameters such as the exact locations of all antenna elements and all scattering objects, which are usually modeled as random variables in mobile communications. The quality of the communication link, however, is mainly determined by the singular values of these matrix taps.

It is well known [6]–[8] that the singular values of a large class of random matrix ensembles show fewer random fluctuations the larger the matrices are, and become deterministic in the limit of infinite matrix size. In the large-matrix limit, the influences of many properties of the matrix entries are lost, such as the shapes of their distributions, and in some cases even the statistical dependencies among them. Reference [9] is a good example of the influence of statistical dependencies on singular values vanishing in the asymptotic limit.

Though the asymptotic distribution of singular values is only an approximation to the distribution in the case of finite-dimensional matrices, it offers two important advantages.

- In contrast to finite-dimensional matrices, the singular value distribution of asymptotically large random matrices can be calculated analytically in many cases.
- In the asymptotic limit, only those physical parameters survive that show significant influence on the singular value distribution.

With these two properties, the limiting singular value distribution can help to analytically extract which physical parameters of the radio propagation channel largely determine the quality of a MIMO communication link.

Motivated by reasons such as these, random matrix theory was used for analysis of antenna arrays in [10]–[13]. These references modeled the antenna-array channel as memoryless with a channel matrix composed of independent and identically distributed random entries. This simplifies the analysis, but may ignore some important properties of the channel. In particular, measurements have demonstrated that multipath *richness*—a parameter that does not occur in [10]–[13]—heavily influences the singular values of the channel [14]–[16]. In order to include this effect, the recent reference [17] has classified MIMO channels into high-rank and low-rank channels. The present paper handles the more general case with an arbitrary number of scatterers at different delay times. This allows for arbitrary ranks of the channel matrices and also includes multipath propagation that causes intersymbol interference.

Analytical results are given in terms of the distribution of the eigenvalues of the spatio–temporal covariance matrix of the channel as the number of antennas at both ends as well as the number of scattering objects grow large, but their ratios remain fixed. The asymptotic eigenvalue distribution is characterized in

Manuscript received September 6, 2000; revised April 1, 2002. The material in this paper was presented in part at the 38th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, October 2000 and the IEEE International Conference on Communications, Helsinki, Finland, June 2001.

The author is with Forschungszentrum Telekommunikation Wien (ftw.), 1220 Vienna, Austria.(e-mail: mueller@ftw.at).

Communicated by M. L. Honig, Associate Editor for Communications.
Publisher Item Identifier 10.1109/TIT.2002.801467.

two different ways. By its probability density function $\mathrm{f}(x)$ and by its Stieltjes transform[1]

$$G(s) \triangleq \int_{-\infty}^{+\infty} \frac{f(x)}{x+s}\,\mathrm{d}x. \tag{1}$$

This goes along with the need for different descriptions for different engineering aims. The probability density function is useful for an intuitive insight into the behavior of the channel. The Stieltjes transform is a well-known tool for easily calculating performance measures such as SINRs. In presence of additive white Gaussian noise (AWGN) of variance σ^2, for instance, and space–time equalization according to the linear minimum mean-squared error (MMSE) criterion, the SINR can be calculated directly in the Stieltjes domain

$$\text{SINR} = \frac{1}{\sigma^2 G(\sigma^2)} - 1. \tag{2}$$

See Section IV-A.

The paper is organized as follows. Section II introduces the channel model in matrix-algebraic notation and the assumptions that are needed for analytical tractability. Section III derives the asymptotic eigenvalue distribution of the channel. This is used to derive asymptotic expressions for channel capacity and SINRs that are achievable with two kinds of linear space–time processing at the receiver site in Section IV. Section V generalizes the results to multiple users at different locations. Section VI summarizes the main results and conclusions. A discussion about the physical meaning of the assumptions made in Section II and the proofs of various theorems are presented in Appendixes A and B, respectively.

II. Random Matrix Model

Consider a communication channel with T transmitting and R receiving antennas. Let there be $S_{\max}$ scattering objects each corresponding to a propagation path with excess delay τ_κ. Further, allow for the follwing assumption.

Assumption 1: On its way from the transmitter to the receiver, each signal is bounced off a scattering object exactly once.

Assumption 1 ensures that there is no line of sight between the transmitter and the receiver. It also excludes multifold scattering. In order to propose a discrete-time model, quantization in space and time is required.

Assumption 2: The delays of all scattering objects can be expressed in discrete time k.

Assumption 3: All scattering objects are located in such a way that they can be separated in either space or time.

A pair of objects that violates Assumption 3 is simply considered a single scattering object. A more detailed discussion of Assumptions 1–3 is given in Appendix A-A.

[1]Sometimes the Stieltjes transform is also defined with negative sign of s. In our context, the definition (1) that follows [18] turns out more intuitive.

The received signal that is received at antenna ν is given by

$$y_\nu[k] = \sum_{\ell=1}^{L} \sum_{\kappa=1}^{S_\ell} A_{\kappa,\ell} \mathrm{e}^{\mathrm{j}\varphi_{\kappa,\nu,\ell}} \sum_{\mu=1}^{T} \mathrm{e}^{\mathrm{j}\vartheta_{\kappa,\mu,\ell}} x_\mu[k-\ell-1] \tag{3}$$

where $x_\mu[k]$ is the signal transmitted by antenna μ, and $\varphi_{\kappa,\nu,\ell}$, $\vartheta_{\kappa,\mu,\ell}$, $A_{\kappa,\ell}$, and S_ℓ are the relative carrier phases at the νth receive and μth transmit antenna, the attenuation of the κth path, and the number of scattering objects, respectively, all at delay ℓ. Note that each of the relative carrier phases depends on the distance between the individual antenna element and the scattering object.

The propagation coefficient from antenna μ to antenna ν at delay ℓ is given as

$$h_{\nu,\mu,\ell} = \sum_{\kappa=1}^{S_\ell} A_{\kappa,\ell} \mathrm{e}^{\mathrm{j}(\vartheta_{\kappa,\mu,\ell}+\varphi_{\kappa,\nu,\ell})}. \tag{4}$$

The number of scatterers may vary with delay. This effect is modeled by the scatterer count delay profile S_ℓ in addition to the well-known power-delay profile

$$P_\ell \triangleq \frac{1}{S_\ell} \sum_{\kappa=1}^{S_\ell} |A_{\kappa,\ell}|^2. \tag{5}$$

The received signal at time instant k can be written in vector notation as

$$\boldsymbol{y}[k] = \sum_{\ell=1}^{L} \boldsymbol{H}_\ell \boldsymbol{x}[k-\ell-1] \tag{6}$$

where the entries of the $R \times T$ matrix $\boldsymbol{H}_\ell$ are defined in (4). It is obvious from (4) that those entries show strong statistical dependencies even if A_κ, $\mathrm{e}^{\mathrm{j}\vartheta_{\kappa,\mu,\ell}}$, and $\mathrm{e}^{\mathrm{j}\varphi_{\kappa,\nu,\ell}}$ are statistically independent for all κ, μ, and ν. These dependencies are examined in greater detail in the following.

Define the two $S_\ell \times T$ and $S_\ell \times R$ matrices

$$\boldsymbol{\Theta}_\ell \triangleq \begin{bmatrix} \mathrm{e}^{\mathrm{j}\vartheta_{1,1,\ell}} & \cdots & \mathrm{e}^{\mathrm{j}\vartheta_{1,T,\ell}} \\ \vdots & \ddots & \vdots \\ \mathrm{e}^{\mathrm{j}\vartheta_{S_\ell,1,\ell}} & \cdots & \mathrm{e}^{\mathrm{j}\vartheta_{S_\ell,T,\ell}} \end{bmatrix} \tag{7}$$

and

$$\boldsymbol{\Phi}_\ell \triangleq \begin{bmatrix} \mathrm{e}^{-\mathrm{j}\varphi_{1,1,\ell}} & \cdots & \mathrm{e}^{-\mathrm{j}\varphi_{1,R,\ell}} \\ \vdots & \ddots & \vdots \\ \mathrm{e}^{-\mathrm{j}\varphi_{S_\ell,1,\ell}} & \cdots & \mathrm{e}^{-\mathrm{j}\varphi_{S_\ell,R,\ell}} \end{bmatrix} \tag{8}$$

respectively, as well as $\boldsymbol{A}_\ell \triangleq \mathrm{diag}([A_{1,\ell}, \ldots, A_{S_\ell,\ell}])$. Then, $\boldsymbol{H}_\ell$ may be expressed as

$$\boldsymbol{H}_\ell = \boldsymbol{\Phi}_\ell^{\mathrm{H}} \boldsymbol{A}_\ell \boldsymbol{\Theta}_\ell \tag{9}$$

with $\cdot^{\mathrm{H}}$ denoting complex conjugate transpose. Note that the matrix $\boldsymbol{\Theta}_\ell$ describes the propagation from the transmitter array to the scattering objects, while the matrix $\boldsymbol{\Phi}_\ell$ models the propagation from the scattering objects to the receiver array.

The further development of the paper will require the following additional assumptions.

Assumption 4: The entries of the matrices $\boldsymbol{\Theta}_\ell$ and $\boldsymbol{\Phi}_\ell$ are independent and identically distributed random variables with zero mean and unit variance for all delays ℓ.

In Appendix A-B, it is shown that Assumption 4 cannot be true exactly, but it is a reasonable approximation to make, in view of the accuracy of the final results, and the analytical tractability that results from this assumption. Further basis for Assumption 4 is outlined in Appendix A-B. One further assumption is needed to allow analytical tractability of the proposed random matrix model.

Assumption 5: The matrices $\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\rm H}$ and $\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\rm H}$ become unitarily invariant for all delays ℓ, as their sizes grow large.

This assumption is a rather technical one; see Appendix A-C for further discussion.

III. Asymptotic Eigenvalue Distribution

For what follows we condition on the path delay and drop the index ℓ for ease of notation where misunderstanding is unlikely. The performance of communication of a large class of linear channels described by a matrix $\boldsymbol{H}_\ell$ is determined by the T eigenvalues of the normalized covariance matrix

$$\boldsymbol{C}_\ell \triangleq \frac{1}{RS_\ell}\boldsymbol{H}_\ell^{\rm H}\boldsymbol{H}_\ell. \tag{10}$$

In general, not all eigenvalues $\lambda_1, \ldots, \lambda_T$ are nonzero, as

$$\mathrm{rank}(\boldsymbol{H}_\ell) \le \min\{T;\, R;\, S_\ell\}. \tag{11}$$

Their empirical distribution[2]

$$\mathrm{F}_{\boldsymbol{C}}(x) \triangleq \frac{1}{T}\left|\{\lambda_i\colon \lambda_i < x\}\right| \tag{12}$$

denotes the fraction of eigenvalues that fall below a certain threshold x.

Theorem 1: Condition on a particular realization of $\boldsymbol{A}_\ell$ and let Assumptions 4 and 5[3] hold. Then all positive moments of (12) converge almost surely to nonrandom limits as T, R, S_ℓ tend to infinity, but the ratios

$$\zeta \triangleq \frac{T}{S_\ell} \quad \text{and} \quad \rho \triangleq \frac{S_\ell}{R} \tag{13}$$

remain fixed.

Theorem 1 is a special case of a more general result in [8]; see Appendix B-A for details.

The asymptotic limits hopefully serve as good estimates for the eigenvalues in the nonasymptotic case. This has been verified for code-division multiple-access (CDMA) systems in [20] and is assumed to extend to a broader class of communication systems described by large random matrices. In the following, we find the asymptotic distributions of the eigenvalues.

The distribution of the eigenvalues is conveniently represented in terms of its Stieltjes transform (1). It follows from [11] (see also [21]) that the Stieltjes transform of the asymptotic eigenvalue distribution of the $S_\ell \times S_\ell$ matrix $\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\rm H}/S_\ell$ is given by

$$G_{\boldsymbol{\Theta}\boldsymbol{\Theta}^{\rm H}/S}(s) = \frac{1-\zeta}{2s} - \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1+\zeta}{2s} + \frac{(1-\zeta)^2}{4s^2}}. \tag{14}$$

Note that the same formula applies to the $S_\ell \times S_\ell$ matrix $\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\rm H}/S_\ell$ if ζ is replaced by ρ^{-1}. It will turn out helpful for calculation of $G_{\boldsymbol{C}_\ell}(s)$ to consider the $S_\ell \times S_\ell$ matrix

$$\tilde{\boldsymbol{C}}_\ell \triangleq \frac{1}{RS_\ell}\boldsymbol{A}_\ell^{\rm H}\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\rm H}\boldsymbol{A}_\ell\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\rm H} \tag{15}$$

instead of the original $T \times T$ matrix (10) in the following. Because of (69), the nonzero eigenvalues of $\boldsymbol{C}_\ell$ are identical to the nonzero eigenvalues of $\tilde{\boldsymbol{C}}_\ell$. It is shown in Appendix B-B that this implies that their respective eigenvalue densities $\mathrm{f}(x)$ satisfy

$$T\,\mathrm{f}_{\boldsymbol{C}}(x) - T\delta(x) = S_\ell\,\mathrm{f}_{\tilde{\boldsymbol{C}}}(x) - S_\ell\delta(x). \tag{16}$$

In the Stieltjes domain, this reads

$$TG_{\boldsymbol{C}}(s) - \frac{T}{s} = S_\ell G_{\tilde{\boldsymbol{C}}}(s) - \frac{S_\ell}{s}. \tag{17}$$

Next, we will focus on $G_{\tilde{\boldsymbol{C}}}(s)$ and return to $G_{\boldsymbol{C}}(s)$ later via (17).

A. Equal Power Case

The asymptotic eigenvalue distribution depends on the distribution of the powers over the paths with fixed delay. Assuming that path loss is the only means of attenuation, all paths with identical delays are attenuated by the same amount. With (5), this means the following.

Assumption 6:

$$\boldsymbol{A}_\ell = \sqrt{P_\ell}\,\mathbf{I}. \tag{18}$$

Note that this model still includes the independent and identically distributed Gaussian ensemble considered in [11], [12] as a special case for $S_\ell \to \infty$ and $L = 1$.

Assumption 6 gives

$$\tilde{\boldsymbol{C}}_\ell = \rho P_\ell \,\frac{\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\rm H}}{S_\ell}\,\frac{\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\rm H}}{S_\ell}. \tag{19}$$

According to Assumptions 4 and 5, $\tilde{\boldsymbol{C}}_\ell$ is the product of two random matrices that are unitarily invariant in the asymptotic limits. The Stieltjes transforms of the asymptotic densities of the factors in the product are available from (14).

A machinery that yields the Stieltjes transform of a matrix product, given the Stieltjes transforms of the factors, is found in [22, Sec. 3.6] and called *multiplicative free convolution*. It can be applied if the factors of the matrix product form a free family in the large matrix limit. This condition is shown to hold in the proof of Theorem 1. Its key tool is the S-transform $S(z)$ which is defined in terms of Stieltjes transforms as

$$S(z) \triangleq \frac{z+1}{z}\,\Upsilon^{-1}(z) \tag{20}$$

$$\Upsilon(s) \triangleq \frac{G\left(-s^{-1}\right)}{-s} - 1 \tag{21}$$

with the auxiliary function $\Upsilon(s)$ and its inverse with respect to composition $\Upsilon^{-1}(z)$. The auxiliary function $\Upsilon(s)$ also has

[2] $|\cdot|$ denotes the cardinality of a set.

[3] A recent more general proof in [19] shows that Assumption 5 is actually not required. However, the proof given in Appendix B-A makes use of Assumption 5.

some practical meaning; it follows from the definition of the Stieltjes transform (1) and the Taylor expansion of its kernel that

$$M(s) \triangleq \sum_{k=0}^{\infty} \underbrace{\int x^k \, \mathrm{dF}(x)}_{m_k} s^k = \Upsilon(s) + 1 \tag{22}$$

where m_k denotes the kth moment of the eigenvalue density and $M(s^{-1})$ the Z-transform of the sequence of moments.

The S-transforms corresponding to the factors in (19) can be shown with (20) and (21) by straightforward algebra to be

$$S_{\mathbf{\Theta\Theta}^{\mathrm{H}}/S}(z) = \frac{1}{z+\zeta} \tag{23}$$

and

$$S_{\mathbf{\Phi\Phi}^{\mathrm{H}}/S}(z) = \frac{1}{z+\rho^{-1}} \tag{24}$$

respectively. Since the S-transform of a product is the product of the factors' S-transforms (see [22, Theorem 3.6.3]) we get

$$S_{\tilde{\boldsymbol{C}}/(\rho P)}(z) = \frac{1}{(z+\zeta)(z+\rho^{-1})}. \tag{25}$$

Definitions (20) and (21) yield

$$\Upsilon_{\tilde{\boldsymbol{C}}/(\rho P)}(s) = s\Big(\Upsilon_{\tilde{\boldsymbol{C}}/(\rho P)}(s)+1\Big)\Big(\Upsilon_{\tilde{\boldsymbol{C}}/(\rho P)}(s)+\zeta\Big) \cdot \Big(\Upsilon_{\tilde{\boldsymbol{C}}/(\rho P)}(s)+\rho^{-1}\Big) \tag{26}$$

and

$$s^2 G^3_{\tilde{\boldsymbol{C}}/(\rho P)}(s) + s\big(\zeta+\rho^{-1}-2\big)G^2_{\tilde{\boldsymbol{C}}/(\rho P)}(s) + \big(s+(\zeta-1)\big(\rho^{-1}-1\big)\big)G_{\tilde{\boldsymbol{C}}/(\rho P)}(s) - 1 = 0 \tag{27}$$

respectively.

Note that $\mathrm{F}_{\alpha X}(x) = \mathrm{F}_X(x/\alpha)$ for any positive α and any random variable X. Thus, the eigenvalue densities satisfy

$$\mathrm{f}_{\tilde{\boldsymbol{C}}/(\rho P)}(x) = \rho P \mathrm{f}_{\tilde{\boldsymbol{C}}}(\rho P x)$$

which reads in the Stieltjes domain [18, Ch. 14] as

$$G_{\tilde{\boldsymbol{C}}/(\rho P)}(s) = \rho P G_{\tilde{\boldsymbol{C}}}(\rho P s).$$

The previous equations are symmetric in ζ and ρ^{-1}. This property does not occur for $G_{\tilde{\boldsymbol{C}}}(s)$

$$s^2\rho P_\ell G^3_{\tilde{\boldsymbol{C}}}(s) + sP_\ell(\rho\zeta+1-2\rho)G^2_{\tilde{\boldsymbol{C}}}(s) + (s+P_\ell(\zeta-1)(1-\rho))G_{\tilde{\boldsymbol{C}}}(s) - 1 = 0. \tag{28}$$

Returning to $G_{\boldsymbol{C}}(s)$ via (17) finally yields

$$s^2 P_\ell \rho\zeta^2 G^3_{\boldsymbol{C}}(s) + sP_\ell\zeta(1+\rho-2\rho\zeta)G^2_{\boldsymbol{C}}(s) + (s+P_\ell(\rho\zeta-1)(\zeta-1))G_{\boldsymbol{C}}(s) - 1 = 0. \tag{29}$$

Although (29) can be resolved with respect to $G_{\boldsymbol{C}}(s)$ via Cardano's formula, the result is omitted here, as it gives no new insight.

Additional insight into (29) is obtained by defining the ratio

$$\beta \triangleq \frac{T}{R} = \zeta\rho. \tag{30}$$

Note that in the theory of CDMA [23], where the number of receive and transmit antennas correspond to the spreading factor and the number of users, respectively [11], β is known as the *load* of the system. For convenience it is named *system load* in the context of antenna arrays as well. With (30), the parameter ζ may be eliminated from (29), giving

$$s^2 P_\ell \beta^2 G^3_{\boldsymbol{C}}(s) + sP_\ell\beta(\rho+1-2\beta)G^2_{\boldsymbol{C}}(s) + (s\rho+P_\ell(\beta-1)(\beta-\rho))G_{\boldsymbol{C}}(s) - \rho = 0. \tag{31}$$

In (31), the Stieltjes transform is parameterized by the system load β and the *richness* of the channel ρ. The system load gives the size ratio of the channel matrix. The *richness* of the channel is the number of scatterers normalized by the number of receiving antennas. This richness in scattering is associated in the literature [24]–[26] with high channel capacities for communication via antenna arrays.

Rewriting (31) in terms of ζ instead of the richness as

$$s^2 P_\ell \beta\zeta G^3_{\boldsymbol{C}}(s) + sP_\ell(\beta+\zeta-2\beta\zeta)G^2_{\boldsymbol{C}}(s) + (s+P_\ell(\beta-1)(\zeta-1))G_{\boldsymbol{C}}(s) - 1 = 0 \tag{32}$$

discloses a nice symmetry in β and ζ. Since the parameter ζ has formally the same meaning to (32) as the system load β, we call ζ *(radio) channel load*, in what follows.

Since it is required in Section V-A, we calculate the asymptotic eigenvalue distribution of the matrix $\boldsymbol{C}'_\ell \triangleq \boldsymbol{H}_\ell\boldsymbol{H}_\ell^{\mathrm{H}}$ as well. Replacing S_ℓ by R in (17), this is a straightforward exercise leading to

$$s^2 P_\ell G^3_{\boldsymbol{C}'}(s) + sP_\ell(\rho+\beta-2)G^2_{\boldsymbol{C}'}(s) + (s\rho+P_\ell(\rho-1)(\beta-1))G_{\boldsymbol{C}'}(s) - \rho = 0. \tag{33}$$

Previously, the asymptotic eigenvalue distribution was characterized in terms of its Stieltjes transform. In the following, the Stieltjes transform is used to derive the asymptotic eigenvalue distribution.

In order to invert the Stieltjes transform implicitly, we make use of the Stieltjes inversion formula (see [8, eq. (3.1.5)][4])

$$\mathrm{dF}(x) = \mathrm{f}(x)\,\mathrm{d}x = \frac{1}{\pi}\lim_{\epsilon\to 0}\Im G(-x-\mathrm{j}\epsilon)\,\mathrm{d}x. \tag{34}$$

As shown in Appendix B-C, this leads to the following construction for the probability density.

Theorem 2: Under Assumptions 4–6, the asymptotic probability density of the eigenvalues of $\boldsymbol{C}_\ell$ is composed of a point mass at zero that is only nonvanishing for $\max\{\zeta,\beta\} > 1$ and a continuous part with compact support. The continuous part is given by $c(x)/\pi$ where $c(x)$ is the unique positive solution to

$$c^2(x) = 3g^2(x) - 2\,\frac{\beta+\zeta-2\beta\zeta}{\beta\zeta x}\,g(x) + \frac{(\beta-1)(\zeta-1)-x}{\beta\zeta x^2} \tag{35}$$

for $x_1 < x < x_2$ and 0 elsewhere. Here the function $g(x)$ represents the (up to three) real solutions of

$$8\beta\zeta x^2 g^3(x) - 8x(\beta+\zeta-2\beta\zeta)g^2(x) + 2\left((\beta-1)(\zeta-1)-x+\frac{(\beta+\zeta-2\beta\zeta)^2}{\beta\zeta}\right)g(x) + \frac{(x-(\beta-1)(\zeta-1))(\beta+\zeta-2\beta\zeta)}{\beta\zeta x} + 1 = 0 \tag{36}$$

[4]The difference of the sign in the reference is due to the different definition of the Stieltjes transform.

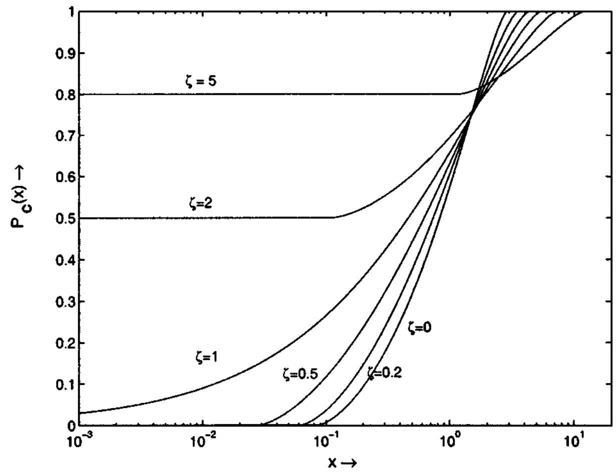

Fig. 1. Asymptotic eigenvalue distribution for system load $\beta = 0.5$ and various channel loads.

and x_1, x_2 are the two largest nonnegative solutions of the equation

$$\begin{aligned}4\beta\zeta x^3 &- \left(10\beta^2\zeta + 10\beta\zeta + 10\beta\zeta^2 - \beta^2 - \beta^2\zeta^2 - \zeta^2\right) x^2 \\ &+ 2\left(4\beta^3\zeta + 4\beta\zeta + 4\beta\zeta^3 - 2\beta^2\zeta - 2\beta^2\zeta^2 - 2\beta\zeta^2 \right. \\ &\left. - \beta^3\zeta^2 - \beta^2\zeta^3 - \beta^3 - \zeta^3 - \beta^2 - \zeta^2\right) x \\ &+ (\beta-1)^2(\zeta-1)^2(\beta-\zeta)^2 = 0. \end{aligned} \tag{37}$$

The influence of the two loads on the asymptotic eigenvalue distribution is illustrated in Fig. 1. It can be observed that the spectral radius increases with the channel load. While the maximum eigenvalue increases monotonically with the two loads, cf. Fig. 2, the smallest positive eigenvalue, becomes minimal if

$$\zeta \le \beta = 1$$
$$\beta \le \zeta = 1$$

or

$$\beta = \zeta \ge 1.$$

In the latter cases, the smallest positive eigenvalue approaches zero arbitrary closely.

B. Multiple Delays

Section III-A has shown that the asymptotic eigenmodes of the channel conditioned on a fixed path delay are determined by only two parameters: the richness ρ and the system load β. The latter is a system parameter rather than a property of the channel. Therefore, the *physical* channel conditions are canonically determined by the richness ρ.

The richness may certainly depend on the path delay. Instead of the scatterer–count delay profile S_ℓ, we can use the richness delay profile ρ_ℓ to characterize the channel. The latter characterization offers the advantage that it does not depend on the *absolute*, but on the *relative* number of scattering objects, and thus scales automatically with the size of the array.

The benefit of rich scattering from a fundamental perspective is in providing enough linearly independent dimensions in

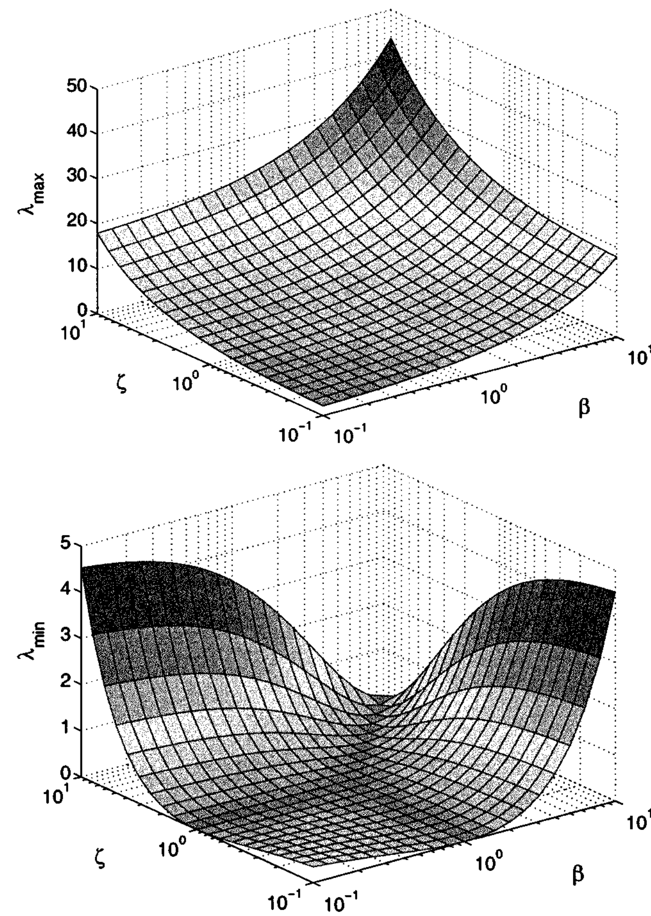

Fig. 2. Largest and smallest (nonzero) eigenvalue of $\boldsymbol{C}$ as a function of the two loads β and ζ in the asymptotic limit.

the signal space to span at least one dimension per transmit antenna. In the case of different delays, it is not immediately clear whether the delayed replicas of the signal indeed span new dimensions.

Mathematically, the problem can be formulated as follows. Stack the symbol vectors $\boldsymbol{x}[k]$ transmitted at subsequent time instances k one below each other in a single vector of transmitted data. Without loss of conceptual scope, the transmission in space and time can be written as a single matrix equation (38), as shown at the top of the following page. Note that $\boldsymbol{\mathcal{H}} \in \{0;1\}^{\infty\times\infty} \otimes \mathbb{C}^{R\times T}$ while $\boldsymbol{H}_\ell \in \mathbb{C}^{R\times T}$. Nevertheless, we prove the following theorem in Appendix B-D.

Theorem 3: Define $\boldsymbol{\mathcal{H}}$ as in (38), let the entries of $\boldsymbol{H}_\ell \in \mathbb{C}^{R\times T}$ be rotationally invariant random variables, and define

$$\boldsymbol{\Sigma} \triangleq \sum_{\ell=1}^{L} \boldsymbol{H}_\ell. \tag{39}$$

Then the eigenvalue distributions of $\boldsymbol{\mathcal{H}}^{\mathrm{H}}\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\Sigma}^{\mathrm{H}}\boldsymbol{\Sigma}$ averaged over the ensemble $\boldsymbol{H}_\ell$, $1 \le \ell \le L$, are identical.

Moreover, if the eigenvalue distributions of $\boldsymbol{H}_\ell^{\mathrm{H}}\boldsymbol{H}_\ell$, $1 \le \ell \le L$, converge in a certain sense to a nonrandom limit as $R, T \to \infty$ and T/R remains fixed, the eigenvalue distributions of $\boldsymbol{\mathcal{H}}^{\mathrm{H}}\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\Sigma}^{\mathrm{H}}\boldsymbol{\Sigma}$ converge in the same sense to identical asymptotic limits.

$$\begin{bmatrix} \vdots \\ \boldsymbol{y}[0] \\ \boldsymbol{y}[1] \\ \boldsymbol{y}[2] \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \\ \cdots & \mathbf{0} & \boldsymbol{H}_L & \cdots & \boldsymbol{H}_2 & \boldsymbol{H}_1 & \mathbf{0} & \cdots & & \\ & \cdots & \mathbf{0} & \boldsymbol{H}_L & \cdots & \boldsymbol{H}_2 & \boldsymbol{H}_1 & \mathbf{0} & \cdots & \\ & & \cdots & \mathbf{0} & \boldsymbol{H}_L & \cdots & \boldsymbol{H}_2 & \boldsymbol{H}_1 & \mathbf{0} & \cdots \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}}_{\triangleq \boldsymbol{\mathcal{H}}} \begin{bmatrix} \vdots \\ \boldsymbol{x}[0] \\ \boldsymbol{x}[1] \\ \boldsymbol{x}[2] \\ \vdots \end{bmatrix}. \tag{38}$$

Note that Theorem 3 does not require any of Assumptions 1–6 to hold. The importance of Theorem 3 is that it shows that for any matrix-valued channel with memory $\boldsymbol{H}_\ell$, $1 \le \ell \le L$, there is a virtual memoryless channel $\boldsymbol{\Sigma}$ with the same eigenvalue distribution.

The summation over the matrix-valued impulse response in (39) can be written with (9) as an inner product of composite matrices

$$\sum_{\ell=1}^{L} \boldsymbol{H}_\ell = \underbrace{\left[\boldsymbol{\Phi}_1^{\rm H} \cdots \boldsymbol{\Phi}_L^{\rm H}\right]}_{\triangleq \boldsymbol{\Phi}^{\rm H}} \underbrace{\operatorname{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_L)}_{\triangleq \boldsymbol{A}} \underbrace{\left[\boldsymbol{\Theta}_1^{\rm H} \cdots \boldsymbol{\Theta}_L^{\rm H}\right]^{\rm H}}_{\triangleq \boldsymbol{\Theta}}. \tag{40}$$

Note that (40) describing the channel with multiple delays takes the same form as (9) which is conditioned on a particular path delay. The sum in (40) just adds up all propagation paths regardless of their delays and stacks them in the composite matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$. The dimensions of the composite matrices are $R \times S_{\max}$ and $T \times S_{\max}$, respectively, with

$$S_{\max} = \sum_{\ell=1}^{L} S_\ell. \tag{41}$$

This is a very nice property. It shows that the equivalent memoryless channel Σ can be also modeled as an antenna-array channel. To be more precise, it shows the following result.

Theorem 4: If the conditions required for Theorem 3 are satisfied and the antenna-array channel satisfies Assumptions 1–3, the equivalent memoryless channel $\boldsymbol{\Sigma}$ also satisfies Assumptions 1–3.

Note that the validity of Theorem 4 crucially depends on Assumption 1 which prohibits multifold scattering.

A channel with varying delays is found equivalent to a channel with identical delays if the richness per delay is replaced by the *total richness*

$$\rho \triangleq \sum_{\ell=1}^{L} \rho_\ell \tag{42}$$

while the system load remains unaffected. This means that for asymptotically large antenna arrays *the eigenvalue distribution of the space–time channel matrix does not change if delay times of particular paths vary*. Moreover, there is no need to distinguish between the distributions of path attenuations conditioned on different delays. Therefore, no joint richness- and power-delay profile is necessary to define the asymptotic eigenvalue distribution of the space–time channel matrix $\boldsymbol{\mathcal{H}}$. It is sufficiently characterized by

- the system load β,
- the total richness ρ,
- and the distribution of the attenuations A_κ, $1 \le \kappa \le S_{\max}$ as $S_{\max} \to \infty$.

If all attenuations are identical, the results found in Section III-A by conditioning on a particular delay apply to the general space–time channel with multipath propagation in both space and time.

Remark 1: The considerations in this subsection should not be misinterpreted as concluding that, as increased bandwidth refines time resolution, it also increases total richness via adding more terms to the sum in (42). Note that a path which differs only in delay time, but neither in angle of arrival nor in angle of departure, cannot increase the richness: It does add a new row to each of the space–time matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$, but these two rows are collinear to those two rows representing the path which differs in delay time only. Such collinear rows are a violation of Assumption 4. They cannot be neglected via considerations such as those outlined in Appendix A-B.

IV. Performance Measures

The results of the previous section can be used to calculate performance measures such as SINRs and channel capacity for transmission over the antenna-array channel

$$\boldsymbol{y}[k] = \sum_{\ell=1}^{L} \boldsymbol{H}_\ell \boldsymbol{x}[k-\ell-1] + \boldsymbol{n}[k] \tag{43}$$

in the presence of AWGN $\boldsymbol{n}[k]$ provided that Assumptions 1–5 hold. For convenience, the equivalent memoryless channel model

$$\tilde{\boldsymbol{y}}[k] = \boldsymbol{\Sigma}\tilde{\boldsymbol{x}}[k] + \tilde{\boldsymbol{n}}[k] \tag{44}$$

established in Section III-B is used for further considerations. In order to allow for analytic expressions in closed form, the attenuations of all paths are assumed identical, i.e.,

$$A_\kappa = \sqrt{P/S_{\max}}, \qquad 1 \le \kappa \le S_{\max}.$$

Knowledge about the channel is assumed to be available at the receiver site, but unavailable to the transmitters. This prohibits the transmitter to use the eigenmodes of the channel for orthogonal signaling. Therefore, the receiver suffers from crosstalk between the signals represented by the components of $\tilde{\boldsymbol{x}}[k]$.

Methods for mitigation of crosstalk on a channel described by (44) are comprehensively discussed in [23] in the context of multiuser detection for CDMA. Many of those do not rely on the particular structure of CDMA signals and can also be applied

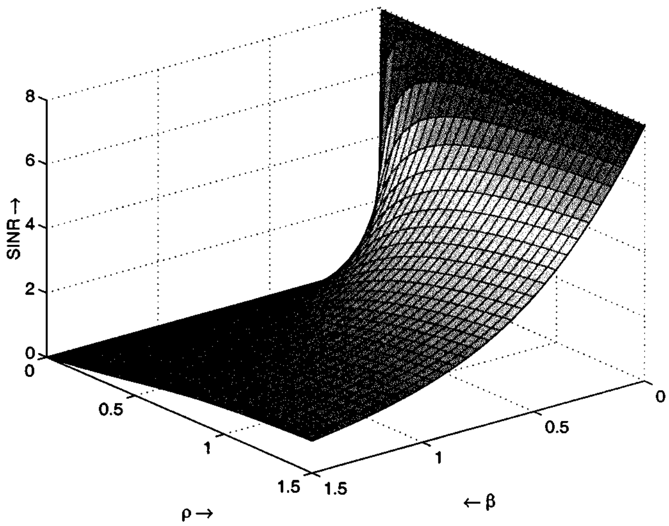

Fig. 3. SINR as function of system load β and richness ρ for $10\log_{10}(P/\sigma^2) = 9$ dB.

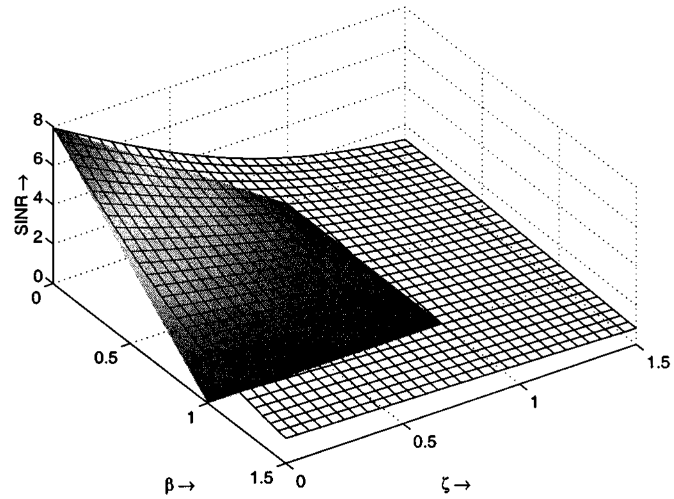

Fig. 4. SINRs of nulling (shaded) and optimum beam forming for $10\log_{10}(P/\sigma^2) = 9$ dB.

to antenna arrays. Two of them, the linear MMSE detector and the decorrelator, are examined in greater detail in the following subsections.

A. Optimum Linear Detection

The Stieltjes transform found in (31) and (32), can be easily used to express the SINR with optimum linear detection, also known as Wiener filtering, optimum linear beamforming, and linear MMSE detection. For this purpose, it is sufficient to make use of standard results in mean-square estimation, i.e.,

$$\mathrm{SINR} = \frac{1}{\mathrm{MMSE}} - 1 \tag{45}$$

and [23, eq. (6.49)]

$$\mathrm{MMSE} = \sigma^2 G_{\boldsymbol{C}}(\sigma^2). \tag{46}$$

with σ^2 denoting the variance of the AWGN. Plugging (46) into (45) obviously gives (2). Plugging (2) into (32) gives the SINR

$$\frac{\sigma^2}{P_\ell}\mathrm{SINR}^3 + \left(2\frac{\sigma^2}{P_\ell} - (\beta-1)(\zeta-1)\right)\mathrm{SINR}^2 + \left(\frac{\sigma^2}{P_\ell} + \beta + \zeta - 2\right)\mathrm{SINR} - 1 = 0 \tag{47}$$

in terms of the two loads.

The SINR is illustrated in Fig. 3. It can be observed that rich scattering is crucial for high SINR. The importance of rich scattering has been stressed in the literature [25], [26] using intuitive arguments. In this work, we demonstrate its importance analytically using an abstract channel model. We also find that rich scattering is more crucial the higher the system load, see Fig. 3.

The special case of independent entries in the matrix $\boldsymbol{H}$ is included in (31) and appears when the relative number of scattering objects ρ goes to infinity. As expected, in that case $G_{\boldsymbol{C}}(s)$ in (31) becomes equivalent to $G_{\boldsymbol{\Theta}\boldsymbol{\Theta}^{\mathrm{H}}/S}(s)$ in (14) when ζ is replaced by β.

B. Nulling

Let the receiver process the received signal in such a way that the crosstalk is completely nulled out. The drawback of this method is the inevitable enhancement of the AWGN. In matrix notation, this type of processing is identified as channel inversion. In the CDMA literature [23] it is known as decorrelation. It is illustrative to calculate the SINR at the output of such a receiver front-end and compare it to the optimum linear beamformer (Wiener filter) analyzed previously.

In the asymptotic limit, the SINR is given as

$$\mathrm{SINR}' = \frac{P}{\sigma^2 \lim\limits_{T\to\infty} \frac{1}{T}\,\mathrm{tr}\left(\boldsymbol{C}^{-1}\right)} \tag{48}$$

$$= \frac{P}{\lim\limits_{T\to\infty} \frac{\sigma^2}{T} \sum\limits_{i=1}^{T} \frac{1}{\lambda_i}} = \frac{1}{\sigma^2 G_{\boldsymbol{C}}(0)} \tag{49}$$

since P is the power of the signal of interest and $\sigma^2\boldsymbol{C}^{-1}$ is the covariance of the filtered noise. Setting $s = 0$ in (32) immediately gives the asymptotic SINR as

$$\mathrm{SINR}' = \frac{P_\ell}{\sigma^2}(1-\beta)(1-\zeta). \tag{50}$$

In order to ensure full rank of $\boldsymbol{C}$, both system and channel load are required to be smaller than one. This means that the richness has to be larger than the system load.

Fig. 4 compares the SINRs of the nulling strategy to the optimum beam former. The degradation of the nulling strategy remains small for $\max\{\beta, \zeta\} \ll 1$.

C. Channel Capacity

In general, calculating the information-theoretic capacity of a channel requires an optimization over the statistics of the transmitted signals, unless forming those statistics is considered as part of the channel. In the present case, we assume that the transmitter array is not aware of the channel at all and cannot adapt the statistics of its signals to the channel.

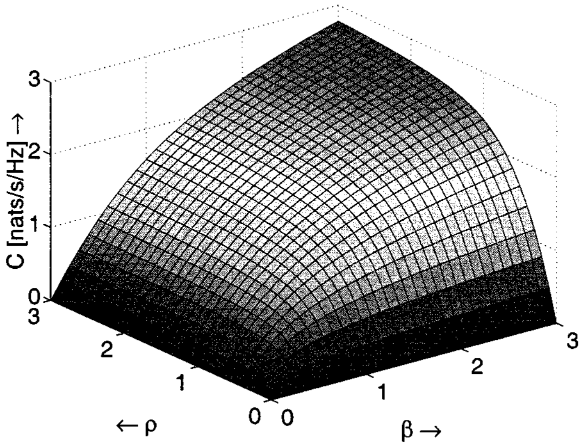

Fig. 5. Channel capacity of the antenna-array channel per unit number of receive antennas with $10\log_{10}(P/\sigma^2) = 9$ dB.

Under this assumption, the channel capacity per receive antenna of in presence of complex AWGN is given by [27]

$$C(\sigma^2) = \frac{1}{R}\sum_{i=1}^{T}\log\left(1+\frac{\lambda_i}{\sigma^2}\right). \quad (51)$$

For asymptotically large arrays, this means

$$C(s) = \beta\int\log\left(1+\frac{x}{s}\right)\mathrm{dF}_{\boldsymbol{C}}(x) \quad (52)$$

with the formal noise variance s. Differentiation with respect to s yields

$$\frac{\partial C}{\partial s} = \beta G_{\boldsymbol{C}}(s) - \frac{\beta}{s}. \quad (53)$$

This can be plugged into (31) giving the following differential equation for channel capacity:

$$\left(s\frac{\partial C}{\partial s}\right)^3 + (\beta+\rho+1)\left(s\frac{\partial C}{\partial s}\right)^2 + \left(\frac{s\rho}{P_\ell}+\rho+\beta\rho+\beta\right)\left(s\frac{\partial C}{\partial s}\right) + \rho\beta = 0 \quad (54)$$

with the boundary condition $\lim_{s\to\infty} C(s) = 0$.

Numerical solutions to this differential equation are easily obtained by solving (54) via Cardano's formula and integrating numerically. An example is depicted in Fig. 5. For rich scattering, capacity grows almost linearly with the system load until unit system load is reached. For higher system loads, no free dimensions are left in the signal space and capacity grows only logarithmically with increasing transmitted signal power. For poor scattering, the channel capacity of the antenna array degrades severely.

V. Multiple Users

Previously, we considered only single-user communication links. With respect to applications in cellular communications, it is also important to consider the more general setting where several terminals are signaling via their respective antenna arrays to a common transceiver station that also employs an array antenna.

There are two possible directions of signal flow: one common transmitter with several receivers, and several transmitters with one common receiver. These two cases correspond to different types of multiuser channels, i.e., the broadcast and the multiple-access channel. For both types, multiple delays are treated with an equivalent memoryless channel.

Let there be K users. Let $\boldsymbol{\Sigma}_k = \boldsymbol{\Phi}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\Theta}_k$ denote the equivalent memoryless channel of user k. There are several important cases.

1) The users see the same set of scattering objects. For these users, the individual channel matrices $\boldsymbol{\Sigma}_k$ are strongly dependent, since the matrix factors that describe propagation between scattering objects and the common transceiver station—$\boldsymbol{\Theta}_k$ and $\boldsymbol{\Phi}_k$ for forward and reverse links, respectively—are identical. The properties of such a setting depend on the statistical dependencies of the other matrix factors—those that describe propagation from scattering objects to the users; these are $\boldsymbol{\Phi}_k$ and $\boldsymbol{\Theta}_k$ for the forward and reverse links, respectively. If they are independent, the users can simply be considered as a single virtual user with the number of antenna elements among the real users added up.
2) The users do not share any scattering objects. Then, the channels of all users are statistically independent. This case will be discussed in detail within this section.
3) The users share some, but not all of the scattering objects. This general case is difficult and exceeds the scope of this paper.

Assumption 7: The channels of all users are statistically independent.

A. Common Receiver

Let user number k transmit via T_k antennas to a common receiver with R antenna elements. Denoting each user's signal as $\boldsymbol{x}_k$, the communication link is described as

$$\boldsymbol{y} = \underbrace{[\boldsymbol{\Sigma}_1\ \boldsymbol{\Sigma}_2\ \cdots\ \boldsymbol{\Sigma}_K]}_{\triangleq \boldsymbol{H}}\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \vdots\\ \boldsymbol{x}_K\end{bmatrix}. \quad (55)$$

For a complete description of the channel's capacity region, all possible subsets of users have to be considered separately [27]. This is a straightforward exercise once capacity is found for the set of all users. We do not pay attention to this detail here and look only at the eigenvalue distribution of the channel's covariance matrix

$$\boldsymbol{C}' = \boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} = \sum_{k=1}^{K}\boldsymbol{\Phi}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\Theta}_k\boldsymbol{\Theta}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{\Phi}_k \quad (56)$$

comprising all users' subchannels.

Under Assumptions 4–7, the terms summed in (56) form a family of free random variables due to [8, Theorem 4.3.5] (see Appendix B-A for details). A mathematical tool to calculate the eigenvalue statistics of a sum from the eigenvalue statistics of free components is the *additive free convolution*. Additive free convolution is easily calculated via the R-*transform* which is defined in terms of the Stieltjes transform as

$$R(w) \triangleq \frac{1}{w} - G^{-1}(w) \tag{57}$$

where $G^{-1}(\cdot)$ denotes the inverse function of $G(\cdot)$ with respect to composition [8], [22]. It can be calculated from (33) by straightforward algebra and reads

$$R_{\mathbf{\Sigma}_k \mathbf{\Sigma}_k^{\mathrm{H}}}(w) = \frac{\beta_k}{2\zeta_k w}\left(\zeta_k + 1 + \frac{1}{P_k w} - \sqrt{\frac{1}{P_k^2 w^2} + \frac{2(\zeta_k+1)}{P_k w} + (\zeta_k - 1)^2}\right) \tag{58}$$

with

$$\beta_k \triangleq \frac{T_k}{R} \quad \text{and} \quad \zeta_k \triangleq \frac{T_k}{S_k}. \tag{59}$$

The R-transform is the counterpart in free probability theory to the log-moment generating function in conventional probability theory. This means the R-transform of a sum of free terms is the sum of the individual R-transforms [8], [22].

Thus, the R-transform in the multiuser case is given in terms of the R-transforms in the single-user case as

$$R_{\mathbf{C}'}(w) = \sum_{k=1}^{K} R_{\mathbf{\Sigma}_k \mathbf{\Sigma}_k^{\mathrm{H}}}(w). \tag{60}$$

Note from (58) that the two loads β and ζ are not additive, in general.

B. Common Transmitter

Let there be a common transmitter with a T-element antenna array signaling to the K users, each of whom employs an antenna array with R_k elements. Again, the channel is characterized in terms of the eigenvalue distribution of its covariance matrix. Note, however, that the considered channel is a broadcast channel [27] and it is not clear whether the eigenvalue distribution yields a canonical description of the channel's capability to support reliable transmission of information.

Denoting each user's signal as $\mathbf{y}_k$, the communication link is described as

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_K \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{\Sigma}_1 \\ \mathbf{\Sigma}_2 \\ \vdots \\ \mathbf{\Sigma}_K \end{bmatrix}}_{\triangleq \mathbf{H}} \mathbf{x}. \tag{61}$$

The channel's covariance matrix

$$\mathbf{C} = \mathbf{H}^{\mathrm{H}}\mathbf{H} = \sum_{k=1}^{K} \mathbf{\Theta}_k^{\mathrm{H}} \mathbf{A}_k^{\mathrm{H}} \mathbf{\Phi}_k \mathbf{\Phi}_k^{\mathrm{H}} \mathbf{A}_k \mathbf{\Theta}_k \tag{62}$$

differs from its counterpart in (56) by an exchange of the matrices $\mathbf{\Phi}_k$ and $\mathbf{\Theta}_k$ only. This is as if here the receivers and transmitters were switched in comparison to Section V-A.

In the equal power case, the R-transform is given with (57) and (32) by

$$R_{\mathbf{\Sigma}_k^{\mathrm{H}} \mathbf{\Sigma}_k}(w) = \frac{1}{2\beta_k \zeta_k w}\left(\beta_k + \zeta_k + \frac{1}{P_k w} - \sqrt{\frac{1}{P_k^2 w^2} + \frac{2(\beta_k+\zeta_k)}{P_k w} + (\beta_k - \zeta_k)^2}\right) \tag{63}$$

where

$$\beta_k \triangleq \frac{T}{R_k} \quad \text{and} \quad \zeta_k \triangleq \frac{T}{S_k}. \tag{64}$$

As in Section V-A, the R-transform in the multiuser case is given in terms of the R-transforms in the single-user case as

$$R_{\mathbf{C}}(w) = \sum_{k=1}^{K} R_{\mathbf{\Sigma}_k^{\mathrm{H}} \mathbf{\Sigma}_k}(w). \tag{65}$$

Similar to the common receiver case in Section V-A, there is no additivity for general parameters β_k and ζ_k.

VI. Summary and Outlook

We have introduced a random matrix model for communication via antenna arrays. The relative number of scattering objects (richness) in the channel was found to be a key parameter to understand the behavior of the channel's eigenvalues and, therefore, its capacity. The problem has been found analytically tractable, and the asymptotic eigenvalue distribution was calculated. This distribution can be used for the design of receivers that rely on iterative analysis tools, e.g., [28]–[30]. Additionally, spatio–temporal multipath propagation could be described by an equivalent memoryless channel model.

Within the assumptions of the model and with respect to the performance measures addressed in the paper, only spatial multipath propagation was found to improve the channel conditions for reliable communication. Propagation paths which can be separated by time delay only were not found helpful.

Several questions remain open: Can the model be extended in order to include multifold scattering? How many antenna elements are needed for the asymptotic limit to be a good approximation? Can the model be confirmed by measured data? Concerning the last question, promising results were recently reported in [31].

Appendix A
Discussion of Assumptions

A. On Assumptions 1–3

Assumption 1 reflects two points:

- Propagation via line of sight does not contribute to the boost in capacity due to the use of antenna arrays at both ends.
- Signals that are bounced off several times on their way from the transmitter to the receiver lose much of their energy. Therefore, they are negligible unless they are the dominant means of propagation. Multifold scattering exceeds the scope of this paper. We refer the interested reader to [32], [19].

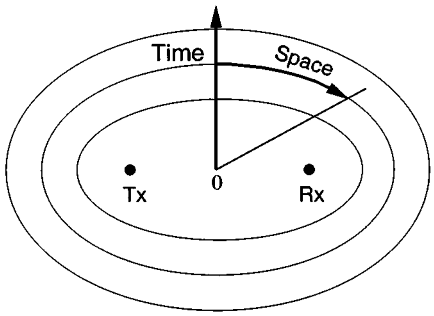

Fig. 6. Ellipsoid model of propagation in time and one angular spatial dimension.

Assumption 1 also suggests to characterize the location of each scattering object in elliptical coordinates (see Fig. 6) with the transmitter and receiver locations being the foci of the ellipsoid. There might be two or only one spatial dimension depending on whether the antenna arrays are able to resolve both azimuth and elevation or not. Though any multipath propagation takes place in space, those paths that correspond to distinguishable delays are called, for convenience, multipath over time, whereas paths with indistinguishable delays but distinguishable locations of the scattering objects are called multipath in space. Whether two delays are distinguishable in time is determined by the bandwidth of the transmitted signal. Whether scattering objects are distinguishable in space is a more difficult question: For uniform linear arrays, the separability of wave-emitting sources depends on the number of antenna elements, the wavelength, and the angles of arrival and departure. For general array geometries, the question is a topic of recent research, see, e.g., [33], [34]. Note that the number of scattering objects depends implicitly on the number of antenna elements and the array geometry via the question of separability.

Assumptions 2 and 3 just assume quantization of the environment within the elliptical coordinates of Fig. 6.

B. On Assumption 4

Assumption 3 ensures that paths with different delays ℓ correspond to different scattering objects. Assuming the locations of the scattering objects to be independent random variables, Assumption 4 follows for any set of matrix entries in which the delay indexes of all elements differ. For matrix entries with identical delay index, however, Assumption 4 is the most critical assumption of the proposed model. In order to see this, condition without loss of generality on delay $\ell = 1$. Assuming the locations of all scattering objects and of all antenna elements to be three-dimensional random vectors, there are at most $3(S_1 + R + T)$ degrees of freedom available to generate the $S_1(R+T)$ entries of $\mathbf{\Theta}_1$ and $\mathbf{\Phi}_1$. Obviously, it is easy to find a set $\{S_1; R; T\}$ such that the number of matrix entries exceeds the number of degrees of freedom. Therefore, the matrix entries cannot, in general, be statistically independent.

Though statistical independence is, in general, impossible, numerical evidence shows that the assumption of statistical independence is a good approximation. In the following, some intuition is given to illustrate this surprising phenomenon. For simplicity, consider uniform linear arrays with adjacent elements spaced by distance d. Note that this is a pessimistic assumption, since it determines the positions of all but one of the elements within each of the antenna arrays, since it further reduces the number of degrees of freedom that are available compared to an array with a random geometry. In the far field, the relative carrier phase from antenna μ to the scattering object κ at delay $\ell = 1$ is given by [35, eq. (17.15)]

$$\vartheta_{\kappa,\mu,1} = -\frac{2\pi(\mu-1)d}{\lambda}\sin(\alpha_{\kappa,1}) \tag{66}$$

with $\alpha_{\kappa,1}$ denoting the angle under which scattering object κ is seen from the first element of the array and λ denoting the wavelength. Thus, the entries of the matrix $\mathbf{\Theta}_1$ satisfy the recursion

$$e^{j\vartheta_{\kappa,\mu+1,1}} = e^{j(\vartheta_{\kappa,\mu,1} - 2\pi d \sin(\alpha_{\kappa,1})/\lambda)}. \tag{67}$$

Due to (67), $\mathbf{\Theta}_1^{\mathrm{H}}$ is a Vandermonde matrix. The only random parameters within $\mathbf{\Theta}_1$ are the angles $\alpha_{\kappa,1}$. The mathematical form of (67), however, is similar to the *linear congruential pseudorandom number generator* [36] defined by the recursion

$$X_{n+1} = (aX_n + c) \bmod m, \qquad n \geq 0. \tag{68}$$

This kind of random number generator is frequently used for fast generation of random integers. Its components are the initial value X_0, the multiplier a, the increment c, and the modulus m. For uniform linear arrays, the initial value is given by the absolute carrier phase that is not included in (67). The multiplier is chosen as $a = 1$, the increment is $-2\pi d \sin(\alpha_{\kappa,1})/\lambda$, and the modulus 2π is taken by the exponential functions with imaginary argument that are applied on the relative carrier phases. In order to obtain good pseudorandom numbers, the modulus should be relatively prime to the increment [36]. This prohibits fast repetition of the sequence. In case of uniform linear arrays, the sequence is guaranteed to be nonperiodic with probability one, since both the modulus and the increment are transcendent numbers without common multiple. Provided that $|\vartheta_{\kappa,\mu,1}|$ is large enough that the modulo reduction actually takes place for most antenna indexes μ—a condition that is obviously fulfilled for a large number of antenna elements or wide element spacing—the choice of $a = 1$ is not critical. In that case, the matrix $\mathbf{\Theta}_1$ has as much randomness as if each of its rows were generated by independent linear congruential pseudorandom generators with different increments. The same considerations apply to the matrix $\mathbf{\Phi}_1$ as well.

C. On Assumption 5

A matrix is said to be unitarily invariant if its distribution does not change when it is simultaneously multiplied by an arbitrary unitary matrix $\boldsymbol{U}$ from the left and by $\boldsymbol{U}^{\mathrm{H}}$ from the right. For instance, if $\boldsymbol{X}$ is a Gaussian random matrix with independent and identically distributed random entries, $\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}$ is unitarily invariant. It is well known [6], [7] that under Assumption 4, the asymptotic singular value distributions of the matrices $\mathbf{\Theta}_\ell$ and $\mathbf{\Phi}_\ell$ are not affected by Assumption 5. Moreover, it is conjectured in [37] that unitary invariance is a property of a much larger class of random matrices than the Gaussian independent and identically distributed ensemble. However, to the author's knowledge, a proof has not yet been found.

APPENDIX B
PROOFS AND DERIVATIONS

A. Proof of Theorem 1

We have to show that $\mathrm{tr}(\boldsymbol{C}_\ell^n)/S_\ell$ converges almost surely for all nonnegative integers n as $S_\ell \to \infty$. With

$$\mathrm{tr}(\boldsymbol{XY}) = \mathrm{tr}(\boldsymbol{YX}) \tag{69}$$

it is easy to see from (10) and (9) that $\boldsymbol{C}_\ell^n$ can be expressed as a noncommutative polynomial in the $S_\ell \times S_\ell$ matrices $\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\mathrm{H}}$, $\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\mathrm{H}}$, $\boldsymbol{A}_\ell$, and $\boldsymbol{A}_\ell^{\mathrm{H}}$.

First, consider the case that $S_\ell \leq \min\{R; T\}$. Under Assumption 4, the self-adjoint random matrices $\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\mathrm{H}}$ and $\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\mathrm{H}}$ converge in distribution almost surely to a compactly supported probability measure as shown in [38], [21]. Moreover, they are unitarily invariant due to Assumption 5. Thus, the matrices $\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\mathrm{H}}$, $\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\mathrm{H}}$, $\boldsymbol{A}_\ell$, and $\boldsymbol{A}_\ell^{\mathrm{H}}$ fulfill the conditions of [8, Theorem 4.3.5] which says that they form a family that is asymptotically free almost everywhere as $S_\ell \to \infty$.

If $S_\ell > \min\{R; T\}$, the limiting probability measure of either $\boldsymbol{\Theta}_\ell\boldsymbol{\Theta}_\ell^{\mathrm{H}}$ or $\boldsymbol{\Phi}_\ell\boldsymbol{\Phi}_\ell^{\mathrm{H}}$, or the two of them is not really compactly supported, but is composed of a continuous density with compact support and a mass point at zero that lies out of the compact support. However, the mass point at zero is not critical. It is deterministic without need for asymptotic limits and [8, Theorem 4.3.5] can be easily extended, as its proof [8, pp. 155–156] shows. Therefore, asymptotic freeness almost everywhere holds for $S_\ell > \min\{R; T\}$, as well.

Asymptotic freeness almost everywhere, however, implies almost sure convergence of all normalized moments of the respective eigenvalue distribution (see [8, pp. 146–147]). The proof is complete.

B. Derivation of (17)

In general, the eigenvalue densities $\mathrm{f}_{\boldsymbol{C}}(x)$ and $\mathrm{f}_{\tilde{\boldsymbol{C}}}(x)$ are composed of a continuous density and a point mass at zero. Since the nonzero eigenvalues of $\boldsymbol{C}_\ell$ and $\tilde{\boldsymbol{C}}_\ell$ are identical, the continuous parts of the respective densities are identical up to a multiplicative factor that ensures $\int \mathrm{dF}(x) = 1$.

Since for any nonnegative random variable X

$$\Pr(X = 0) - 1 = -\Pr(X > 0)$$

we have the stronger result

$$\mathrm{f}_{\boldsymbol{C}}(x) - \delta(x) = a(\mathrm{f}_{\tilde{\boldsymbol{C}}}(x) - \delta(x)) \tag{70}$$

with a to be determined.

The probability that a randomly chosen eigenvalue of the $T \times T$ matrix $\boldsymbol{C}_\ell$ and the $S_\ell \times S_\ell$ matrix $\tilde{\boldsymbol{C}}_\ell$ is nonzero is obviously $E\,\mathrm{rank}(\boldsymbol{C}_\ell)/T$ and $E\,\mathrm{rank}(\tilde{\boldsymbol{C}}_\ell)/S_\ell$, respectively. Since both matrices have identical rank, we find

$$a = \frac{S_\ell}{T} \tag{71}$$

which gives (16) with (70).

C. Proof of Theorem 2

Let

$$g(x) \triangleq \lim_{\epsilon\to 0} \Re G_{\boldsymbol{C}}(-s - \mathrm{j}\epsilon) \tag{72}$$

$$c(x) \triangleq \lim_{\epsilon\to 0} \Im G_{\boldsymbol{C}}(-s - \mathrm{j}\epsilon). \tag{73}$$

For $c(x) \neq 0$, the real and imaginary parts of (32) give the following system of equations:

$$\begin{aligned}&\beta\zeta x^2 g^3(x) - 3\beta\zeta x^2 g(x) c^2(x)\\ &\quad - x(\beta + \zeta - 2\beta\zeta)(g^2(x) - c^2(x))\\ &\quad + ((\beta - 1)(\zeta - 1) - x)g(x) - 1 = 0\end{aligned} \tag{74}$$

$$\begin{aligned}&3\beta\zeta x^2 g^2(x) - \beta\zeta x^2 c^2(x) - 2x(\beta + \zeta - 2\beta\zeta)g(x)\\ &\quad - x + (\beta - 1)(\zeta - 1) = 0.\end{aligned} \tag{75}$$

Solving (75) for $c^2(x)$ gives (35). Solving (74) for $c^2(x)$ and setting its solution equal to the right-hand side of (35) gives (36).

In order to get the supporting interval of $c(x)$, we note that $c(x_1) = c(x_2) = 0$, since $c(x)$ is continuous. Plugging (37) into (36) and (35) verifies, after some tedious but straightforward algebra, that x_1 and x_2 are zeros to $c(x)$.

D. Proof of Theorem 3

The space–time channel matrix $\boldsymbol{\mathcal{H}} \in \{0; 1\}^{\infty\times\infty} \otimes \mathbb{C}^{R\times T}$ is circulant in the space of $\infty \times \infty$ matrices on $\mathbb{C}^{R\times T}$. Therefore, it can be decomposed into

$$\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{T}}_R \boldsymbol{\mathcal{L}} \boldsymbol{\mathcal{T}}_T^{\mathrm{H}} \tag{76}$$

where $\boldsymbol{\mathcal{T}}_i \in \mathbb{C}^{\infty\times\infty} \otimes \{\mathbf{I}_i\}$ are $\infty \times \infty$ Fourier matrices on the set of $i \times i$ identity matrices $\mathbf{I}_i$ and $\boldsymbol{\mathcal{L}} \in \{\mathbf{I}_\infty\} \otimes \mathbb{C}^{R\times T}$ is an $\infty\times\infty$ diagonal matrix on the set of $R\times T$ matrices. The Fourier matrices are unitary with respect to both the matrix entries and the underlying complex scalar entries. Thus, the singular values of the space–time channel matrix $\boldsymbol{\mathcal{H}}$ in $\{0; 1\}^{\infty\times\infty} \otimes \mathbb{C}^{R\times T}$ are identical to those of $\boldsymbol{\mathcal{L}}$ in $\{\mathbf{I}_\infty\} \otimes \mathbb{C}^{R\times T}$.

Since the space–time channel matrix is circulant, the nonzero matrix-valued entries of $\boldsymbol{\mathcal{L}}$ are given as the matrix-valued Fourier transform of the first matrix-valued row of $\boldsymbol{\mathcal{H}}$ [39]. Therefore, the ith matrix-valued diagonal element of $\boldsymbol{\mathcal{L}}$ can be written as

$$\boldsymbol{\mathcal{L}}_{ii} = \sum_{\ell=1}^{L} \boldsymbol{H}_\ell \mathrm{e}^{\mathrm{j}\alpha_{i,\,\ell}} \tag{77}$$

with some $\alpha_{i,\,\ell}$ being determined by the Fourier transform. The complex phase factors $\exp(\mathrm{j}\alpha_{i,\,\ell})$ do not change the statistics of the random matrices $\boldsymbol{H}_\ell$, as the entries of $\boldsymbol{H}_\ell$ are rotationally invariant. Therefore, $\alpha_{i,\,\ell}$ is assumed to be zero without changing the expected singular value distribution of $\boldsymbol{\mathcal{L}}_{ii}$.

The matrix $\boldsymbol{\mathcal{L}}$ is block-diagonal. Therefore, its expected singular value distribution is the mean of the expected singular value distributions of the block matrices $\boldsymbol{\mathcal{L}}_{ii}$. The expected singular value distributions of all the block matrices $\boldsymbol{\mathcal{L}}_{ii}$ are identical, as $\alpha_{i,\,\ell}$ may be assumed zero without changing the expected singular value distribution. Therefore, the expected singular value distributions of

$$\boldsymbol{\mathcal{H}} \quad \text{and} \quad \boldsymbol{\Sigma} = \sum_{\ell=1}^{L} \boldsymbol{H}_\ell \tag{78}$$

in $\{0; 1\}^{\infty\times\infty} \otimes \mathbb{C}^{R\times T}$ and $\mathbb{C}^{R\times T}$, respectively, are identical. Since the nonzero eigenvalues of any matrix $\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}$ are the squared nonzero singular values of $\boldsymbol{X}$, equivalence holds also for the respective eigenvalue distributions.

If the singular value distributions of $\boldsymbol{H}_\ell$ converge asymptotically to a nonrandom limits, for all ℓ, all of the above considerations for expected singular value distributions hold respectively for asymptotic singular value distributions.

Acknowledgment

The author would like to thank M. Toeltsch, C. Mecklenbräuker, A. Molisch, E. Bonek, G. Caire, H. Hofstetter, E. Schofield, and the anonymous reviewers for helpful suggestions.

References

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

[2] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.

[3] G. G. Raleigh and J. M. Cioffi, "Spatio–temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[5] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.

[6] V. L. Girko, *Theory of Random Determinants*. Dordrecht, The Netherlands: Kluwer Academic, 1990.

[7] M. L. Mehta, *Random Matrices*, 2nd ed. Boston, MA: Academic, 1991.

[8] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*. Providence, RI: Amer. Math. Soc., 2000.

[9] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2059–2078, Sept. 2000.

[10] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, Nov./Dec. 1999.

[11] D. N. C. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.

[12] J. Bach Andersen, "Antenna arrays in mobile communications: Gain, diversity, and channel capacity," *IEEE Antennas Propagat. Mag.*, vol. 42, pp. 12–16, Apr. 2000.

[13] S. V. Hanly and D. N. C. Tse, "Resource pooling and effective bandwidth in CDMA networks with multiuser receivers and spatial diversity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1328–1351, May 2001.

[14] Gigabit Wireless Inc., "Experimental results for MIMO technology in an indoor environment," Tech. Rep., San Jose, CA, Mar. 1999.

[15] J. P. Kermoal, P. E. Mogensen, S. H. Jensen, J. B. Andersen, F. Frederiksen, T. B. Sørensen, and K. I. Pedersen, "Experimental investigation of multipath richness for multi-element transmit and receive antenna arrays," in *Proc. IEEE Vehicular Technology Conf. (VTC)*, vol. 3, Tokyo, Japan, May 2000, pp. 2004–2008.

[16] E. Bonek and M. Steinbauer, "Double-directional channel measurements," in *Proc. 11th IEE Int. Conf. Antennas and Propagation*, Manchester, U.K., Apr. 2001, pp. 226–230.

[17] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 2, San Francisco, CA, Nov./Dec. 2002, pp. 1083–1088.

[18] H. Bateman, *Table of Integral Transforms*. New York: McGraw-Hill, 1954, vol. 2.

[19] R. R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2086–2091, July 2002.

[20] ——, "Multiuser receivers for randomly spread signals: Fundamental limits with and without decision-feedback," *IEEE Trans. Inform. Theory*, vol. 47, pp. 268–283, Jan. 2001.

[21] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *J. Multivariate Anal.*, vol. 54, pp. 175–192, 1995.

[22] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables*. Providence, RI: Amer. Math. Soc., 1992.

[23] S. Verdú, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.

[24] P. F. Driessen and G. J. Foschini, "On the capacity formula for multiple input–multiple output wireless channels: A geometric interpretation," *IEEE Trans. Commun.*, vol. 47, pp. 173–176, Feb. 1999.

[25] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[26] J. Bach Andersen, "Multiple antennas—The promise of high spectral efficiency," *Telektronikk*, vol. 97, no. 1, pp. 40–48, 2001.

[27] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[28] R. R. Müller and S. Verdú, "Design and analysis of low-complexity interference mitigation on vector channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1429–1441, Aug. 2001.

[29] G. Caire and R. R. Müller, "The optimal received power distribution for IC-based iterative multiuser joint decoders," in *Proc. 39th Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Oct. 2001.

[30] R. R. Müller and G. Caire, "Efficient implementation of iterative multiuser decoding," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Lausanne, Switzerland, June 2002, p. 446.

[31] R. R. Müller and H. Hofstetter, "Confirmation of random matrix model for the antenna array channel by indoor measurements," in *Proc. IEEE Antennas Propagation Society Int. Symp.*, Boston, MA, July 2001, pp. 472–475.

[32] R. R. Müller, "On the effect of multi-fold scattering onto the capacity of BLAST-like systems," in *Proc. Int. Conf. Telecommunications (ICT)*, Bucharest, Romania, June 2001.

[33] A. Manikas, A. Sleiman, and I. Dacos, "Manifold studies of nonlinear antenna array geometries," *IEEE Trans. Signal Processing*, vol. 49, pp. 497–506, Mar. 2001.

[34] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, "Detection-estimation of more uncorrelated Gaussian sources than sensors in nonuniform linear antenna arrays—Part I: Fully augmentable arrays," *IEEE Trans. Signal Processing*, vol. 49, pp. 959–971, May 2001.

[35] S. R. Saunders, *Antennas and Propagation for Wireless Communication Systems*. Chichester, U.K.: Wiley, 1999.

[36] D. E. Knuth, *The Art of Computer Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1981, vol. 2.

[37] J. W. Silverstein, "Eigenvalues and eigenvectors of large dimensional sample covariance matrices," in *Random Matrices and Their Applications*, J. E. Cohen, H. Kesten, and C. M. Newman, Eds. Providence, RI: Amer. Math. Soc., 1986, pp. 153–159.

[38] ——, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *J. Multivariate Anal.*, vol. 55, pp. 331–339, 1995.

[39] P. J. Davis, *Circulant Matrices*. New York: Wiley, 1979.