Capacity Analysis of Correlated MIMO Channels

Leif Hanlen, *Member, IEEE*, and Alex Grant, *Senior Member, IEEE*

***Abstract*—This paper gives expressions for the capacity of ergodic multiple-input multiple-output channels with finite dimensions, in which the channel gains have a correlated complex normal distribution and receivers experience independent Gaussian noise. The particular correlated normal distribution considered corresponds to flat Rayleigh fading with arbitrary transmit and receive correlation. Knowledge of the correlation matrices is assumed at both the transmitter and receiver, while the receiver, but not the transmitter, has complete knowledge of the channel realization. The optimal input density is characterized via a necessary and sufficient condition for optimality, along with an iterative algorithm for its numerical computation. The resulting capacity is expressed in terms of hypergeometric functions of matrix argument, which depend on the channel correlation matrices only through their eigenvalues. Some closed-form expressions are also given in the case of single-sided correlation. Some consideration is given to high- and low-power asymptotics. Easily computable asymptotic expressions are also given for receive-side only correlation in the case that the number of transmitters is large. In that case, the capacity can be divided into two components: one arising from the dominant eigenvalues of the receiver-end correlation matrix, and the other from the remaining spherically distributed eigenvalues. Some numerical results are also presented.**

***Index Terms*—Correlated multiple-input multiple-output (MIMO) channels, MIMO channel capacity, random matrix theory.**

I. Introduction

THE work in [2]–[4] has shown that under the assumption of an i.i.d. transfer matrix, the capacity of a multiple-input multiple-output (MIMO) channel grows linearly in proportion to the minimum of the number of inputs and outputs. This is the often cited *linear growth* in capacity of MIMO channels. The channel assumptions model flat Rayleigh fading in a point-to-point wireless system and have been verified under the condition of widely separated (uncorrelated) antenna elements, and dense multipath scattering [5].

The application of MIMO transmission to spatially diverse systems does not necessarily lead to the large growth predicted in [2] and [3]. Several authors have recently investigated correlation of the MIMO channel due to various physical mechanisms and shown a reduction in the rate of growth in channel capacity. The work in [6] and [7] showed that for widely separated array elements, sparse scattering results in a correlation of the channel and lower mutual information for equal-power white transmission. These results, however, require parameterizations of the channel in terms of physical placement of scatterers and are difficult to apply in an abstract setting.

Recently, Chuah *et al.* [8] have used the *Stieltjes transform* (see, for example, [9] and [10]) to obtain asymptotic results for the capacity of a correlated channel. The results obtained provide insight into the asymptotic properties of MIMO channels. Unfortunately, the Stieltjes transform does not lend itself well to numerical solutions. Authors in [7] and [8] have suggested that the growth of the MIMO channel will remain linear for a correlated channel, although the proportionality constant for the linear growth may change.

This paper provides an analysis of correlated MIMO channels using random matrix theory [11]. Using hypergeometric functions of matrix argument, we find the capacity for transmit-end and/or receive-end correlation, where the correlation matrices are known to both the transmitter and receiver and the channel realization is known to the receiver but not the transmitter. In the special case of single ended correlation (i.e., transmit correlation, or receive correlation, but not both) closed-form results are obtained. High- and low-power asymptotes are considered, along with the corresponding effect on the optimal input density. The asymptotic channel capacity (as the number of transmitters becomes large, with receive-end only correlation) is also derived.

Related work,[1] taking a number of different approaches is found in [12]–[18] (see also [19] for channel measurement results). In contrast, this paper concentrates on the direct generalization of [2]. In the final stages of the September 2004 revision of this paper, we became aware of related independent work [20], [21], also using hypergeometric functions to obtain capacity results. In that work, the authors find the characteristic function of capacity for receive-side correlation only. Asymptotics are not considered in [20] and [21].

This paper is arranged as follows. Section III presents the main results, namely a fixed-point equation for the optimal input density and expressions for capacity. High- and low-power asymptotics are considered in Section IV. Section V gives an asymptotic channel capacity as the number of transmitters becomes large with fixed numbers of receivers and receive-side only correlation. This capacity may be expressed in terms of an i.i.d. component and a correlated component. This is used to compare the correlated channel with the well-known i.i.d. channel. Using a trace-rule to compare different correlated matrices, numerical results are presented in Section VI. All proofs are contained in the Appendix, which also contains an overview of some properties of hypergeometric functions.

Manuscript received December 30, 2002; revised January 09, 2003; accepted March 14, 2012. Date of publication September 07, 2012; date of current version October 16, 2012. A. Grant was supported in part by the Australian Government under ARC Grant DP0209658. This paper was presented at the 2003 International Symposium on Information Theory.

L. Hanlen was with the School of Electrical Engineering and Computer Science, University of Newcastle, Callaghan, N.S.W. 2308, Australia. He is now with the National ICT Australia, Canberra, A.C.T. 2601 Australia, and also with the Australian National University, Acton, A.C.T. 0200, Australia.

A. Grant is with the Institute for Telecommunications Research, University of South Australia, Adelaide, S.A. 5000, Australia.

Communicated by E. Telatar, Associate Editor for Shannon Theory, and by A. Lozano, Associate Editor for Communications.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2012.2216393

[1]This section refers only to related work up to the 2004 revision of this paper. Section II details further developments, 2004–2011.

Section II, a new section, describes relevant developments that have been published since the 2004 revision of this paper, up to the present, 2011. The remainder of the paper is identical to the 2004 revision, expect for minor typographical corrections and some minor technical fixes, the latter noted in the footnotes.

II. Subsequent Developments

Since the original submission and subsequent revision of this paper in 2004, several of our results have been independently rediscovered and in some cases generalized by other authors. This section aims to cover those more recent developments of particular relevance to this paper. The contemporary state of the art (*circa* 2004) is summarized in [22] and [23]. In particular, work in [22, Sec. 3.3.3] gives a summary of the separable correlation case. An excellent recent reference is the book by Couillet and Debbah [24], which has an up-to-date treatment of correlated channels (including reporting on the results of this paper as announced in [1]).

The optimal covariance for a zero mean Kronecker random matrix is well known to diagonalize the transmit covariance—a result already established in [25] and used here and elsewhere [26]–[29]. Theorem 1 that gives necessary and sufficient conditions for optimality of an input covariance matrix was rediscovered (using similar arguments stemming from the Karush–Kuhn–Tucker conditions) in [30, eq. (5)]. In that paper, the authors also usefully reexpress the result in terms of the minimum mean squared error [30, eqs. (10) and (11)]. For numerical computation, [30, Algorithm 1] is an implementation of our fixed-point (6), again expressed in terms of MMSE. A similar fixed-point equation to that given in this paper was also provided in [31]; however, a transformation into "virtual channels" was required in order to use the equation. Another related iteration can be found in [24, p. 314]. In another direction, there has been a significant amount of work considering optimal input covariances, and optimized precoding for restricted inputs [30], [32], [33].

A key observation of this paper is that correlations may be beneficial to MIMO capacity under certain combinations of correlation and power limits, a phenomenon which cannot be seen under isotropic transmission assumptions. Similar conclusions that correlated channels could have higher capacity under optimal signaling were provided by [34] using a virtual channel argument. Feedback of the statistics of the channel to the transmitter was shown to be beneficial as a medium access protocol [35], and Jafar and Goldsmith [36] drew similar conclusions. Kang and Alouini [37] also showed that correlation of MIMO channels might not lead to capacity reduction under appropriate covariance feedback.

A variation on random channels that generalized the Kronecker model and used the algorithm of [30] to find the input covariance was given in [33]. Extensions of this work to precoding with parallel channels [32] and nondiagonal channels [38] under certain input constraints have been developed.

Many authors have developed similar input covariance optimization algorithms for specific special cases of channel. In [39], an algorithm for determining the optimal input covariance based on asymptotically large antenna numbers was described and Dupuy and Loubaton in [40] extended this to include frequency selectivity. Marques and Abrantes [41] found an exact closed-form expression for the receive-side correlated channel using hypergeometric functions and developed a program to evaluate a formula similar to (8). Li *et al.* [42] offered a method that was related to Karush–Khun–Tucker conditions for the covariance matrix and provided a solution for the input covariance for a two-input multiple-output system under the constraint that either the matrix was deterministic (i.e., fixed channel known to both transmitter and receiver), or Jensen's inequality could be used, effectively corroborating [43]. For deterministic channels, Vandenberghe *et al.* [44] provide the optimization solution for matrix channel capacity with total power constraint, per-antenna power constraint, and with crosstalk.

Works such as [45]–[47] considered alternate power constraints, where the transmit covariance is premultiplied by a known deterministic matrix.

Other researchers have developed bounds [48]–[50] for Ricean models and for outage capacity of a isotropic source [51], [52]. Vu and Paulraj [53] considered so-called dynamic channel state information (CSI). This provides a bridge from full CSI at the transmitter to statistical models only (as used in this paper). The extension of this is toward uncertain statistical knowledge given by [54] which then bridges to channels that are completely unknown [55].

Some results that effectively extend work in this paper may be seen in [56] where the authors considered the nonuniqueness of the optimal covariance, and provided necessary and sufficient conditions for unique input covariance. The work of multiuser components of [56] was developed to include interference cases [57]. Gao *et al.* [58] were unable to form analytic solutions to capacity but developed closed-form (tight) upper bounds for capacity for nonseparable correlation, going beyond work in this paper.

The fixed-point equation of this paper has been shown to hold in the case of *arbitrary* random matrix channels [59] where the transmitter possesses the channel statistics. In this case, the algorithm is simply repeated application of (6), followed by a trace constraint. The factorization of the covariance matrix into upper triangular matrices used in [60] is unnecessary. Gohary *et al.* [61] applied the algorithm in [59] and compared this to the water filling on the covariance matrix directly [43].

There is still no clear approach for nonzero mean channels when the channel mean and channel correlation(s) are not jointly diagonalizable. For general channels, the optimal transmit direction is not a simple combination of the eigenvectors of the mean and/or transmit correlation [62, Myth 2], [59, Fig. 2].

Existing closed-form results for optimal covariance are for specific numbers of elements, specific correlations [28], or numeric algorithms [30], [63]. The differential of mutual of information [30] and other secondary functions (such as moments) do yet not have closed-form solutions for general input covariance. The expectation given in (5) still does not appear to have a closed-form solution. This represents a significant barrier to advances in optimal input covariance: the objective function itself (let alone the optimal solution) has not yet been expressed in closed form. One approach which side-steps the lack of a closed form for the expectation (5) is to use the iteration in [30], where the fixed point is expressed in terms of MMSE, for which a closed-form expression has been given in the case of Rayleigh fading and separable correlation [64, Th. 1].

The numerical evaluation of the hypergeometric function remains an impediment for evaluation of capacity formulae. Work in [20], [21], and this paper applied particular channel simplifications to provide closed-form solutions that could be evaluated. Advances by Koev and Edelman [65] may yet remove this barrier for general input covariance. High- and low-signal-to-noise-ratio (SNR) asymptotic results have since been considered by several authors. Our high-SNR asymptotes (11) and (12) are special cases of [66, Proposition 4], which generalizes the result in several useful ways including for double-ended correlation and for Ricean channels. Low-SNR asymptotes have since been extensively treated in [67], which considers approximations derived from second-order expansions rather than our first-order approximations in Section IV. While work in [67] provides generalizations for Ricean channels and polarization diversity; the results were all constrained to isotropic input signaling. We have also subsequently given second-order asymptotes and second-order conditions for optimality for the input covariance at low SNR in [68]. These conditions follow from the fact that it is possible to simultaneously diagonalize both terms in the second-order Taylor expansion by the eigenvectors of the transmit correlation. More general results for SNR intercept, slope, and first-order optimality were given in [69]. At low SNR, it is of interest to know the conditions under which beamforming is optimal. These conditions are provided in [26].

Large dimension asymptotes for correlated channels have been subsequently considered in [28], [33], and [34], achieving generalization to double-ended correlation, and dropping our requirement for $t = r$.

III. Main Results

Consider a point-to-point communication link with t inputs and r outputs. Throughout the paper, $m = \min\{r, t\}$ and $n = \max\{r, t\}$. At each symbol interval, the received complex vector $y \in \mathbb{C}^r$ depends on the transmitted vector $x \in \mathbb{C}^t$ according to

$$y = Hx + w. \tag{1}$$

Element y_j is the jth matched-filter output, while x_i is the ith transmitted symbol. The transmitter has an average power limit $\mathrm{E}\{x^* x\} \leq \gamma$, where the superscript * denotes conjugate transpose and $\mathrm{E}\{\cdot\}$ denotes expectation. The matrix $H \in \mathbb{C}^{r \times t}$ has entries H_{ji}, which are the complex gains between input i and output j. The vector $w \in \mathbb{C}^r$ contains i.i.d. circularly symmetric Gaussian noise samples [70, p. 134], $\mathrm{E}\{ww^*\} = \eta^2 I_r$, a scaled $r \times r$ identity matrix. Without loss of generality, assume $\eta^2 = 1$. More general choices of noise covariance are easily accommodated by appropriate modification of γ and H.

According to the notation of [71], the random[2] matrix H is chosen from a complex Gaussian ensemble $H \sim N_{r,t}(0, R \otimes T)$ where the *receive covariance* R is $r \times r$ and the *transmit covariance* T is $t \times t$ (both R and T are assumed to be symmetric, nonsingular).

[2]Assume H is chosen independently each symbol interval, although H need only change often enough to consider the channel ergodic.

Subject to this choice of H rewrite (1) as

$$y = R^{1/2} \tilde{H} T^{1/2} x + w \tag{2}$$

where $\tilde{H}$ has i.i.d. circularly symmetric complex Gaussian entries, with independent real and imaginary parts, each with variance $1/2$ [2] and $R = R^{1/2} R^{1/2}$ and $T = T^{1/2} T^{1/2}$ (such "square roots" exist since the covariance matrices are positive definite). Accordingly, the scope of this paper is restricted to correlation structures of the form $\mathrm{E}\{\mathrm{vec}(H)\mathrm{vec}(H)^*\} = R \otimes T$, where $\mathrm{vec}(H)$ stacks the columns of H into a single column vector and $\otimes$ is the Kronecker product. The well-known case $R = I_r$ and $T = I_t$ is considered in [2]. It is assumed throughout that R and T are known at the transmitter and receiver and that the channel realization H is known at the receiver but not at the transmitter.

This choice of correlation structure for H is the so-called *Kronecker model* [72], used for analytical purposes by many authors (see, for example, [8] and [73]–[77]). The accuracy of this model has received recent scrutiny [19], [78], [79] and there may be cases of practical interest which require more general structures [80]. Despite potential deficiencies of the model, it provides a tractable approach to the analysis of correlated MIMO channels and it is for that reason adopted here.

For exact knowledge of H at the receiver, a zero-mean Gaussian distribution on the input vector x is optimal [2], [81], although this distribution is not necessarily i.i.d. [25], [82]. The capacity-achieving input distribution is, therefore, completely specified its covariance $Q = \mathrm{E}\{xx^*\}$, which must be optimized for the determination of capacity.

The capacity of the ergodic correlated MIMO channel (1) with channel matrix $H = R^{1/2} \tilde{H} T^{1/2}$ with correlation matrices T and R and power constraint γ is then given by

$$C(T, R, \gamma) = \max_{\mathrm{tr}\, Q \leq \gamma} \mathrm{E}_{\tilde{H}} \left\{ \log\det \left(I_r + R\tilde{H} T^{1/2} Q T^{1/2} \tilde{H}^* \right) \right\} \tag{3}$$

and the optimizing transmit covariance will be denoted $\hat{Q}$. In the case $T = \tau I$ and $R = \rho I$, the optimal input covariance is the well-known $\hat{Q} = I_t \gamma / t$ equal power "white" transmit scenario [2]. The following lemma shows that nonwhite receive covariance does not affect the form of the optimal input distribution.

Lemma 1: Given the channel (2), with $T = I_t$, power limit γ and both R and T known at the transmitter, the capacity-achieving input distribution is equal power, independent white Gaussian signals $\mathrm{E}\{xx^*\} = \hat{Q} = \gamma I_t / t$ and is thus independent of R.

For the remainder of the paper, it will be assumed that T and R are diagonal. That this results in no loss of generality is demonstrated by the following Lemma.

Lemma 2: Let $T = U_T \Sigma_T U_T^*$ and $R = U_R \Sigma_R U_R^*$ be the respective eigenvalue decompositions of T and R (i.e., U_T, U_R, unitary and Σ_T, Σ_R diagonal). Then, $C(T, R, \gamma) = C(\Sigma_T, \Sigma_R, \gamma)$.

The main results of the paper are Theorem 1 which gives an implicit characterization of the optimal input density as a function of T and R, Theorem 2 which expresses capacity in terms of a hypergeometric function of matrix argument, and Theorem 3 which gives closed-form expressions for capacity in the case of single-ended correlation.

In the case that the transmitter has complete knowledge of the channel realization H, the capacity-achieving strategy is transmission of independent Gaussian symbols on the right singular vectors of H. The optimal power allocation is obtained by water-pouring on the corresponding singular values. The following theorem gives the form of the optimal input covariance when the transmitter only has knowledge of T and R but not H.

Theorem 1 (Optimal Covariance): Consider the channel (2), with $R = \operatorname{diag}\{\rho_1, \dots, \rho_r\}$ and $T = \operatorname{diag}\{\tau_1, \dots, \tau_t\}$, where both R and T are known at the transmitter, and power limit γ. A necessary and sufficient condition for the optimality of an input covariance $Q = \operatorname{diag}\{q_1, \dots, q_t\}$ is

$$\mathrm{E}\left\{\left((I+SQ)^{-1}S\right)_{kk}\right\} = \mu, \qquad q_k > 0 \tag{4}$$

$$\mathrm{E}\left\{\left((I+SQ)^{-1}S\right)_{kk}\right\} < \mu, \qquad q_k = 0 \tag{5}$$

for $k = 1, 2, \dots, t$ and some constant μ. The expectation is with respect to the random matrix $S = H^*H$, $H \sim N_{r,t}(0, R \otimes T)$.

In the case $Q > 0$, condition (4) may be rewritten as a fixed-point equation[3]

$$Q = \nu\, \mathrm{E}\left\{\left(Q^{-1}+S\right)^{-1}S\right\} \tag{6}$$

which suggests the following iterative procedure for numerically finding the optimal Q. Starting from an initial diagonal $Q^{(0)} > 0$, compute

$$q_k^{(i+1)} = \nu^{(i+1)}\left[\mathrm{E}\left\{\left((Q^{(i)})^{-1}+S\right)^{-1}S\right\}\right]_{kk}$$

selecting $\nu^{(i)}$ at each step to keep $\operatorname{tr} Q^{(i)} = \gamma$. Although there is no existing closed-form solution for $\mathrm{E}\left\{\left(Q^{-1}+S\right)^{-1}S\right\}$, it may be accurately estimated using Monte Carlo techniques.

Conditions (4), (5) may be compared with the corresponding condition for parallel Gaussian channels. Suppose $y = Sx + n$ where S is a deterministic diagonal matrix known to both the transmitter and receiver; then, the condition for optimality of the input covariance is

$$\left((I+SQ)^{-1}S\right)_{kk} = \mu, \qquad q_k > 0$$

$$\left((I+SQ)^{-1}S\right)_{kk} < \mu, \qquad q_k = 0.$$

Thus, Theorem 1 can be recognized as a direct generalization of the classical water-pouring result for parallel channels.

The following theorem, proved in the Appendix, expresses the capacity in terms of hypergeometric functions of matrix argument.

[3]Here, we have assumed that $Q > 0$. A little more care must be taken if some of the q_k are zero.

Theorem 2: The capacity of the ergodic correlated MIMO channel (1) with $H \sim N_{r,t}(0, R \otimes T)$ and power constraint $\mathrm{E}\{x^*x\} \le \gamma$ is given by

$$\begin{aligned} C(R,T,\gamma) = {} & \frac{\pi^{m(m-1)}}{\tilde{\Gamma}_m(n)\tilde{\Gamma}_m(m)\det(R)^r\det(T\hat{Q})^t} \\ & \cdot \int_\Lambda {}_0F_0\left(-R^{-1}, \Lambda, (T\hat{Q})^{-1}\right) \\ & \cdot \prod_{i=1}^m \lambda_i^{(n-m)} \prod_{i<j}^m (\lambda_i - \lambda_j)^2 \sum_{i=1}^m \log(1+\lambda_i)\, d\Lambda \end{aligned} \tag{7}$$

where $\Lambda = \operatorname{diag}\{\lambda_1, \dots, \lambda_m\}$, $\hat{Q}$ is selected optimal according to Theorem 1, $\tilde{\Gamma}_m(a) = \pi^{m(m-1)/2}\prod_{i=1}^m \Gamma(a-i+1)$ is the complex multivariate gamma function and ${}_0F_0(\cdot,\cdot,\cdot)$ is a hypergeometric function of three matrix arguments.

This theorem is a natural generalization of the result in [2], as ${}_0F_0(\cdot,\cdot,\cdot)$ is a generalization of etr. Setting $T = I_t$ and $R = I_r$ directly recovers Telatar's result since ${}_0F_0(I, \Lambda, I) = {}_0F_0(\Lambda) = \operatorname{etr}(\Lambda)$. In the case of single-ended correlation, i.e., $T = I_t$ and $R \ne I_r$ or $T \ne I_t$ and $R = I_r$, the resulting hypergeometric functions in Theorem 2 can be computed in closed form using a determinant formula [83]. The resulting capacity for both cases may be written in terms of a function $C(A, \gamma)$, defined for diagonal $A = \operatorname{diag}\{a_i\}$ as

$$\begin{aligned} C(A,\gamma) & \triangleq \frac{1}{m\det(\mathbf{V}(A^{-1}))} \\ & \cdot \sum_{j=1}^m \sum_\pi (-1)^{\operatorname{per}(\pi)} \tilde{C}\left(n-m+\pi(j), 1, a_j\gamma\right) \prod_{i=1}^m (a_i)^{\pi(i)-m} \end{aligned} \tag{8}$$

$$\begin{aligned} & = \frac{1}{m\det(\mathbf{V}(A^{-1}))} \\ & \cdot \sum_{k=1}^m \det\left\{a_j^{i-m} \cdot \tilde{C}\left(n-m+k, 1, a_j\gamma\right)^{\delta_{ik}}\right\} \end{aligned} \tag{9}$$

where the sum $\sum_\pi$ is over all $m!$ permutations of $(1, 2, \dots, m)$, $\operatorname{per}(\pi)$ is 0 for even and 1 for odd permutations, $\mathbf{V}(X)$ denotes a Vandermonde matrix of the eigenvalues of X and $\delta_{ik} = 1$ if $i = k$ and is zero otherwise. Finally, $\tilde{C}(r, 1, \gamma) = C(I_r, 1, \gamma)$ is the capacity for r receivers, 1 transmitter and power γ, defined in (69) and given in [2].

Theorem 3: For the channel (2)

$$\begin{aligned} C(R, I_t, \gamma) &= C(R, \gamma) \\ C(I_r, T, \gamma) &= C(\hat{Q}T, \gamma) \end{aligned} \tag{10}$$

where for the second line $\hat{Q}$ is selected according to Theorem 1.

It is interesting to note that using (10) or (9), the capacity for a single-ended-correlated t-input r-output channel is expressed completely as weighted sums of single-input multiple-output channels for both transmit-side and receive-side correlation.

IV. Asymptotic Power Results

The results of the previous section show how to numerically compute the optimal power distributions across the singular vectors of the transmit covariance matrix. In this section, the low- and high-SNR extremes are considered, yielding simple closed-form results for capacity and corresponding optimal power distributions.

For high SNR

$$\lim_{\gamma\to\infty} C(R,T,\gamma) = \lim_{\gamma\to\infty} \mathrm{E}\left\{\log\det\left(\gamma H\hat{Q}H\right)\right\}$$

where now $\operatorname{tr}\hat{Q} = 1$. This may be simplified in the following two cases, using [71, Prob.3.15]. First, for $T = I_t$ and $t \geq r$, $\hat{Q} = I/t$ and[4]

$$C(R,I,\gamma) = m\log\left(\frac{\gamma}{t}\right) + \log\det(R) + \sum_{i=1}^{m}\psi(i) + o(1) \quad (11)$$

where $\psi(i)$ is the psi function. This shows that at high SNR $C(R,I,\gamma)$ depends on R only through its determinant. On the other hand, for $R = I_r$, $T = \operatorname{diag}\{\tau_1,\ldots,\tau_t\}$ and $r \geq t$,

$$C(I,T,\gamma) = m\log(\gamma) + \log\det\left(T\hat{Q}\right) + \sum_{i=1}^{m}\psi(i) + o(1) \quad (12)$$

and clearly (12) is maximized for $\hat{Q} = \alpha I$ which is independent of T.

Some care is needed with this result, which requires that the product $\gamma H\hat{Q}H \gg I$. This requires that γ is large and T is nonsingular. For the case where T is almost singular, a significantly larger value of γ is required before the asymptotic result holds. In practical terms, what this means is that as long as the variation in the τ_k is small compared to γ, the gain to be had by optimizing the power allocation is minimal.

For low SNR (once again setting $\operatorname{tr}\hat{Q} = 1$)

$$C(R,T,\gamma) = \gamma\mathrm{E}\{\operatorname{tr}(S)\} + o(\gamma) \quad (13)$$

with $S = XX^*$ and $X \sim N_{r,t}\left(0, R\otimes T\hat{Q}\right)$. This is a consequence of $\log(1+x) \approx x$ for small x. Now, since trace and expectation commute [2]

$$\mathrm{E}\{\operatorname{tr}(S\} = \operatorname{tr}(T\hat{Q})\operatorname{tr}(R) \quad (14)$$

according to [71, Th. 7.3.5(i)]. For the case $T = I$

$$C(R,I,\gamma) \approx \gamma\operatorname{tr}(R) \quad (15)$$

and the optimal input strategy is not unique: all inputs with diagonal covariance Q satisfying $\operatorname{tr}(Q) = \gamma$ achieve capacity. For $T \neq I$, $\operatorname{tr}(T\hat{Q})$ is clearly maximized when $\hat{Q} = \operatorname{diag}\{1,0,\ldots,0\}$ so that $\operatorname{tr}(T\hat{Q}) = \tau_1$ and

$$C(T \neq I, R, \gamma) \approx \gamma\tau_1\operatorname{tr}(R). \quad (16)$$

This corresponds to the well-known beamforming results of [25], [82], and [84]. Note that (13) is valid for small power levels and/or singular T. As opposed to the high-SNR dependence on $\det R$, the low-SNR dependence is on $\operatorname{tr} R$.

[4]Equations (11)–(13) have been modified from the 2004 revision to replace a sloppy $\lim_{\gamma\to\infty}$ on the left-hand side with an order expansion. Other than this, the expressions are identical to the 2004 version. Thanks to A. Lozano for pointing out this correction.

V. Asymptotic Behavior Large Dimension

Motivated by a down-link scenario in which a base station transmits to a small mobile terminal, this section concentrates on the behavior of capacity as $m = r$ is held fixed and $n = t \to \infty$ with $T = I_t$. Let $\Omega = nHH^*$. The following lemma results from [11, Corollary 9.5.7], converting for complex R and Ω.

Lemma 3 [11]: Without loss of generality, assume that R is diagonal. Suppose that for some value of $0 < k < m$, R has eigenvalues $\rho_1, \rho_2, \ldots, \rho_m$ satisfying

$$\rho_1 > \cdots > \rho_k > \rho_{k+1} = \cdots = \rho_m = \rho > 0 \quad (17)$$

which incorporates R having distinct eigenvalues as a special case. Define

$$z_i = \left(\frac{n}{2}\right)^{1/2}\left(\frac{\hat{\lambda}_i - \rho_i}{\rho_i}\right), \quad i = 1,\ldots,m. \quad (18)$$

Then, the limiting joint density function of $z_1,\ldots,z_m$ for $n \to \infty$ is given by

$$p_z(z_1,\ldots,z_m) = p_z(z_1,\ldots,z_k)\cdot p_z(z_{k+1},\ldots,z_m) \quad (19)$$

$$p_z(z_1,\ldots,z_k) = \prod_{i=1}^{k}\phi(z_i) \quad (20)$$

$$p_z(z_{k+1},\ldots,z_m) = \frac{\pi^{q(q-1)}}{2^q\Gamma_q(q)}\exp\left(-\sum_{j=k+1}^{m} z_j^2\right)\prod_{\substack{k+1\\ i<j}}^{m}(z_i - z_j)^2 \quad (21)$$

where $q = m - k$, and $\phi(\cdot)$ is the standard zero-mean, Gaussian distribution with variance $1/2$ and z_i is independent of z_j for $i \neq j$ and $i,\ j \leq k$.

For the case $k = 0$, i.e., $R = \rho\cdot I_m$, (21) reduces to the standard Wishart distribution for i.i.d. matrices. A similar result [85] as above holds even in the case of $\rho_1 \geq \cdots \geq \rho_m > 0$ where R contains several (different) repeated eigenvalues.

With the changes of variable (18) (7) may be expanded into two parts, C_d for the first k *distinct* eigenvalues of Σ_m and C_{eq} for the remaining $m - k$ *equal* eigenvalues of Σ_m

$$C(R,T,\gamma) = C_d + C_{\mathrm{eq}} \quad (22)$$

$$C_d = K_{k,n}\int\prod_{i=1}^{k}\phi(z_i) \cdot\sum_{i=1}^{k}\log\left(1 + \frac{\gamma}{t}n\rho_i\left[z_i\left(\frac{n}{2}\right)^{-1/2} + 1\right]\right)dz_i \quad (23)$$

$$C_{\mathrm{eq}} = K_{m-k,n}\int_{-\infty}^{\infty} p_z(z_{k+1},\ldots,z_m) \cdot\sum_{i=k+1}^{m}\log\left(1 + \frac{\gamma}{t}n\rho\left[z_i\left(\frac{n}{2}\right)^{-1/2} + 1\right]\right)dz_{k+i} \quad (24)$$

where $K_{k,n}$ and $K_{m-k,n}$ are normalizing constants. The terms (23) and (24) may be simplified as follows.

Reversing the order of sum and integral in (23), write (23) as a summation of iterated integrals

$$C_d = K_{k,n} \sum_{j=1}^{k} \int \cdots \int \prod_{i=1}^{k} \frac{\exp\left(-z_i^2\right)}{\sqrt{\pi}} \cdot \log\left(1 + \frac{\gamma}{t} n\rho_i \left[z_j \left(\frac{n}{2}\right)^{-1/2} + 1\right]\right) \prod dz_i. \tag{25}$$

In each iterated integral, only one particular term $z_{i=j}$ is inside the $\log(\cdot)$, all the others $z_{j\neq i}$ may be integrated out as they are independent. As the limits for all the integrals in (23) are identical, this results in the simplified form:

$$C_d = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{k} \int \exp\left(-z^2\right) \cdot \log\left(1 + \frac{\gamma}{t} n\rho_i \left[z \left(\frac{n}{2}\right)^{-1/2} + 1\right]\right) dz. \tag{26}$$

The asymptotic limit is completed using the following substitution:

$$\mu = z\left(\frac{n}{2}\right)^{-1/2} + 1, \qquad dz = \left(\frac{n}{2}\right)^{1/2} d\mu. \tag{27}$$

As z has normal distribution, so too does μ. However, $\mu \backsim N(1, \frac{1}{2n})$ and consequently, in the limit $n \to \infty$ the distribution of μ becomes $\delta(\mu - 1)$ where $\delta(\cdot)$ is the Dirac delta function. Therefore, in the limit of large nC_d may be approximated by

$$C_d \to \sum_{i=1}^{k} \log\left(1 + \frac{\gamma}{t} n\rho_i\right). \tag{28}$$

The integral of (24) is independent of ordering of z_i and the unordered density function $p_z(z_{k+1})$ may be used to simplify the calculations. Following [2], the analysis for C_{eq} is very similar. First, remove the sum of (24) by noting that the terms are identically distributed

$$C_{\text{eq}} = (m-k) \int_{-\infty}^{\infty} p_z(z_{k+1}) \cdot \log\left(1 + \frac{\gamma}{t} n\rho \left[z_{k+1} \left(\frac{n}{2}\right)^{-1/2} + 1\right]\right) dz_{k+1}. \tag{29}$$

In order to calculate $p_z(z_{k+1})$, integrate the density function $p_z(z_{k+1}, \ldots, z_m)$ over the terms $z_{k+2} \ldots z_m$

$$p_z(z_{k+1}) = \int \cdots \int p_z(z_{k+2}, \ldots, z_m) \prod_{i=k+2}^{m} dz_i. \tag{30}$$

Note that (19) contains a similar Jacobian term $\prod_{i<j}(z_i - z_j)^2$ of [2] which suggests the use of the Vandermonde determinant identity. Define the Vandermonde matrix $V(z_{k+1}, \ldots, z_m)$ over the eigenvalues $z_{k+1}, \ldots, z_m$ as

$$V(z_{k+1}, \ldots, z_m) = \begin{bmatrix} 1 & \cdots & 1 \\ z_{k+1} & \cdots & z_m \\ \vdots & & \vdots \\ z_{k+1}^{m-k-1} & \cdots & z_m^{m-k-1} \end{bmatrix} \tag{31}$$

and write (30) as

$$p_z(z_{k+1}, \ldots, z_m) = (m-k)! \hat{K}_{m-k,n} \cdot \det\left[V(z_{k+1}, \ldots, z_m)\right]^2 \prod_{i=k+1}^{m} e^{-z^2}. \tag{32}$$

Applying a set of (orthonormal) functions $\varphi_1(\cdot), \ldots, \varphi_{m-k}(\cdot)$, to the rows of the Vandermonde matrix. Then, $V(z_{k+1}, \ldots, z_m)$ may be rewritten as $\bar{V}(z_{k+1}, \ldots, z_m)$ with

$$\bar{V}(z_{k+1}, \ldots, z_m) = \begin{bmatrix} \varphi_1(z_{k+1}) & \cdots & \psi_1(z_m) \\ \vdots & & \vdots \\ \varphi_{m-k}(z_{k+1}^{m-k-1}) & \cdots & \varphi_{m-k}(z_m^{m-k-1}) \end{bmatrix}.$$

The orthonormality condition on $\varphi_i(\cdot)$ is given by

$$\langle \varphi_a(z), \varphi_b(z) \rangle = \int_{-\infty}^{\infty} \varphi_a(z)\varphi_b(z) e^{-z^2}\, dz = \delta_{a,b} \tag{33}$$

where $\delta_{a,b}$ is the Kronecker delta. The orthonormal functions that satisfy (33) are Hermite polynomials [86] giving

$$\varphi_j(z) = \left(\frac{1}{2^j j! \sqrt{\pi}}\right)^{1/2} H_j(z), \quad j = 0, \ldots, m-k-1. \tag{34}$$

Now $\det\left[\bar{V}\right]$ can be written in terms of the determinant identity for row operations as [2], [87]

$$\det\left[\bar{V}\right] = \sum_{\alpha} (-1)^{\text{per}(\alpha)} \prod_{j} \varphi_{\alpha(j)}(z_{k+j}). \tag{35}$$

Substituting (35) into (32) and integrating over $z_{k+2} \ldots z_m$ gives

$$p_z(z_{k+1}) = \frac{1}{m-k} \sum_{j=1}^{m-k} \varphi_j(z_{k+1})^2 e^{-z_{k+1}^2}. \tag{36}$$

From (34), C_{eq} can be written as

$$C_{\text{eq}} \approx \int_{-\infty}^{\infty} \log\left(1 + \frac{\gamma}{t} n\rho \left[z\left(\frac{n}{2}\right)^{-1/2} + 1\right]\right) \cdot \sum_{j=1}^{m-k} \frac{1}{2^j j! \sqrt{\pi}} \left[H_j(z)\right]^2 e^{-z^2}\, dz \tag{37}$$

where $H_j(\cdot)$ is the jth Hermite polynomial [88]. Hence, combining (28) and (37) results in the following limit theorem.

Theorem 4 (Asymptotic Correlated MIMO Capacity): Consider a correlated MIMO channel (1) such that the correlation matrix R has eigenvalues which satisfy (17), and $T = I$. Then, the asymptotic capacity C_∞ of the channel, as $n \to \infty$ with $\gamma t/(\mathrm{tr}\, T^{-1}) \to \gamma$ and finite m is given by

$$C_\infty \backsim \sum_{i=1}^{k} \log\left(1 + \frac{\gamma}{t} n\rho_i\right) + \int_{-\infty}^{\infty} \log\left(1 + \frac{\gamma}{t} n\rho \left[z\left(\frac{n}{2}\right)^{-\frac{1}{2}} + 1\right]\right) \cdot \sum_{j=1}^{m-k} \frac{1}{2^j j! \sqrt{\pi}} \left[H_j(z)\right]^2 e^{-z^2}\, dz \tag{38}$$

where $H_j(\cdot)$ is the jth Hermite polynomial, and the approximations are accurate to $O(n^{-1})$. The notation $A \backsim B$ denotes $A/B \to 1$ for $n \to \infty$, [11, p. 391].

It is also possible to accommodate MIMO channels with reduced rank correlation matrices R having eigenvalues satisfying

$$\rho_1 > \cdots > \rho_k > \rho_{k+1} = \cdots = \rho_p > \rho_{p+1} = \rho_m = 0 \tag{39}$$

that is, the smallest $m - p$ eigenvalues are identically zero. In this case, apply the transform [71]

$$H = U_p X \tag{40}$$

where U_p is $m \times p$ and $U_p U_p^* = I_p$ and replace m with p in the Theorem 4. This generalization permits arbitrary correlation matrices R which may not be strictly full rank.

It is interesting to compare the result of Theorem 4 with the well-known i.i.d. case [2]. In the following, constants arising from the integrals will be suppressed. Using the change of variable (27), the integral of Theorem 4 has the form

$$\mathcal{I} = c\int_{-\infty}^{\infty} \log(1 + \gamma\rho\mu) \sum_{j=1}^{r-k} H_j\left[n(\mu-1)\right]^2 \delta(\mu-1)\, d\mu \tag{41}$$

for constant c. A similar argument may be applied to the logarithmic term in (41) as was used in (28). The Hermite polynomial, however, requires special care. While the variance of μ shrinks to zero, the argument of the Hermite polynomial will have constant variance, independent of n. As such, the summation term in (41) must be integrated over the range $(-\infty, \infty)$ and the above result simplifies to give

$$C_\infty \to \sum_{i=1}^{k} \log(1 + \gamma\rho_i) + (r - k)\log(1 + \gamma\rho_m) \tag{42}$$

with $\rho_m \leq 1$ and $\hat{\rho} = 1$ in [2]. Fig. 1 plots the value of C_{eq} in for various $r - k$, and $\gamma = 1$. It can be seen that the integral of Theorem 4 is closely approximated by a linear function.

The relation (42) may also be compared with the result for $t \gg r$ from [2]

$$C_{\mathrm{iid}} \to r\log(1 + \gamma\hat{\rho}). \tag{43}$$

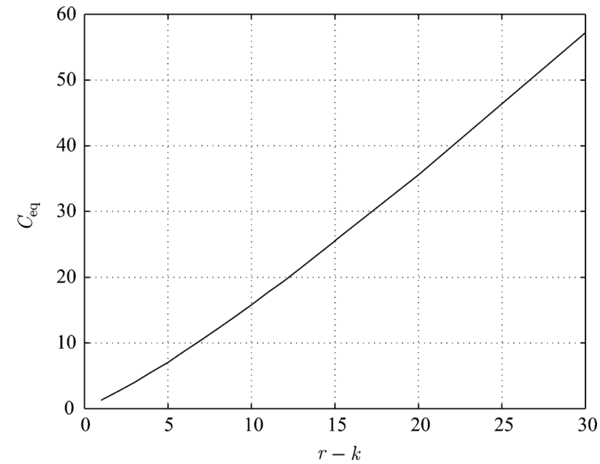

Fig. 1. C_{eq} as a function of $r - k$.

For $k = 0$, (42) and (43) are identical. The parameter k may be interpreted as the *degree* of correlation of the channel matrix H. Note that the asymptotic channel capacity C_∞ grows linearly with increasing r although the rate of growth is a function of k.

In order to compare different correlation matrices R with (43), adopt a constraint on the trace of R. The elements of an arbitrary R are confined to satisfy

$$\mathrm{tr}\, R = \sum_{i}^{m} \rho_i = m. \tag{44}$$

This is equivalent to requiring that each correlated channel transfers the same *total power* from transmitters to receivers, given the same power input. In the case of equal eigenvalues, $\rho_i = \rho$ for some value ρ, we note that (44) gives $\rho_i = \rho = 1$ corresponding to the i.i.d. channel.

Note that the number of distinct eigenvalues k is insufficient to fully determine the "severity" of the correlation of the channel. The range of values that the eigenvalues take must also be considered. Therefore, consider different correlation matrices in terms of a parameter ϵ, where ϵ is defined as

$$\epsilon = \frac{\rho_m}{\bar{\rho}} = \rho_m \tag{45}$$

and $\bar{\rho}$ is the mean eigenvalue of Σ_m. The second equality arises from the normalization of (44) giving $\bar{\rho} = 1$. The correlation matrix R may be considered in terms of the combination of k and ϵ.

The eigenvalues of H give the energy transfer through the MIMO channel. As shown above, these eigenvalues are dominated by the eigenvalues of R. For fixed total energy transfer (44), the parameters k and ϵ may be considered as distributions of energy across the effective parallel channels between transmitters and receivers.

The value of k determines the concentration of channels—for small $k > 0$, the total fading channel is well approximated by a singular line-of-sight channel. For larger values of k, the energy of the MIMO channel is distributed over a wider range of directions. The value of ϵ gives the variation of the eigenvalues of R. For small ϵ, the smaller (equal) eigenvalues will diminish. The worst case is where $\epsilon \to 0$ which corresponds to a reduced

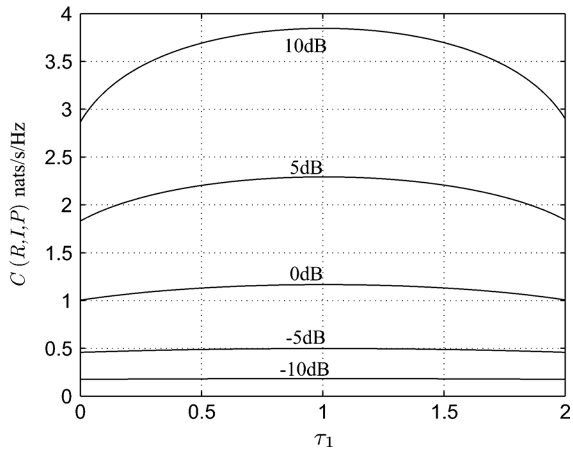

Fig. 2. Capacity $C(R, I, P)$ for $r = t = 2$ and $R = \mathrm{diag}\{\tau_1, \tau_2\}$, such that $\mathrm{tr}(R) = r = 2$.

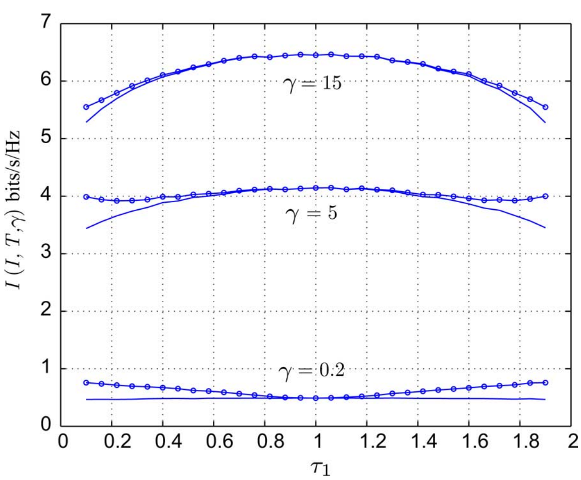

Fig. 3. Mutual Information $I(I, T, \gamma)$ for $r = t = 2$ and $T = \mathrm{diag}\{\tau_1, \tau_2\}$, such that $\mathrm{tr}(T) = t = 2$ for various SNR levels. The optimal input distribution $\hat{Q}$ (circled) is compared with $Q = I$, shown solid.

rank (singular) channel. The optimal value for ϵ is unity, which results in an i.i.d. channel.

VI. Numerical Results

Fig. 2 plots the capacity of a $t = r = 2$ channel, with $T = I_t$ and $R = \mathrm{diag}\{\rho_1, \rho_2\}$ such that $\rho_1 + \rho_2 = 2$. For low SNR, the correlation R has little impact on the capacity. Similarly, for high SNR, the capacity is not significantly impacted for moderate values of R—the plot becomes flat for a wide range of ρ_1.

Fig. 3 shows the capacity of a $t = r = 2$ channel, with $R = I$ and $T = \mathrm{diag}\{\tau_1, \tau_2\}$ such that $\tau_1 + \tau_2 = 2$. The capacity resulting from optimal $\hat{Q}$ given by (6) (shown solid) is compared to the mutual information given by equal power $Q = I \cdot P/t$ for various values of γ. It can be seen for low values of γ, and/or large ratios τ_1/τ_2, that the optimal covariance outperforms equal power allocation, although the benefit reduces as γ increases. In particular, for low SNR, the optimal covariance reduces to beamforming. This corresponds to well-known results on water filling for deterministic channels: for high-SNR optimal power allocation offers little benefit over equal power distributions, while at low SNR, optimal power allocations significantly outperform equal power distributions.

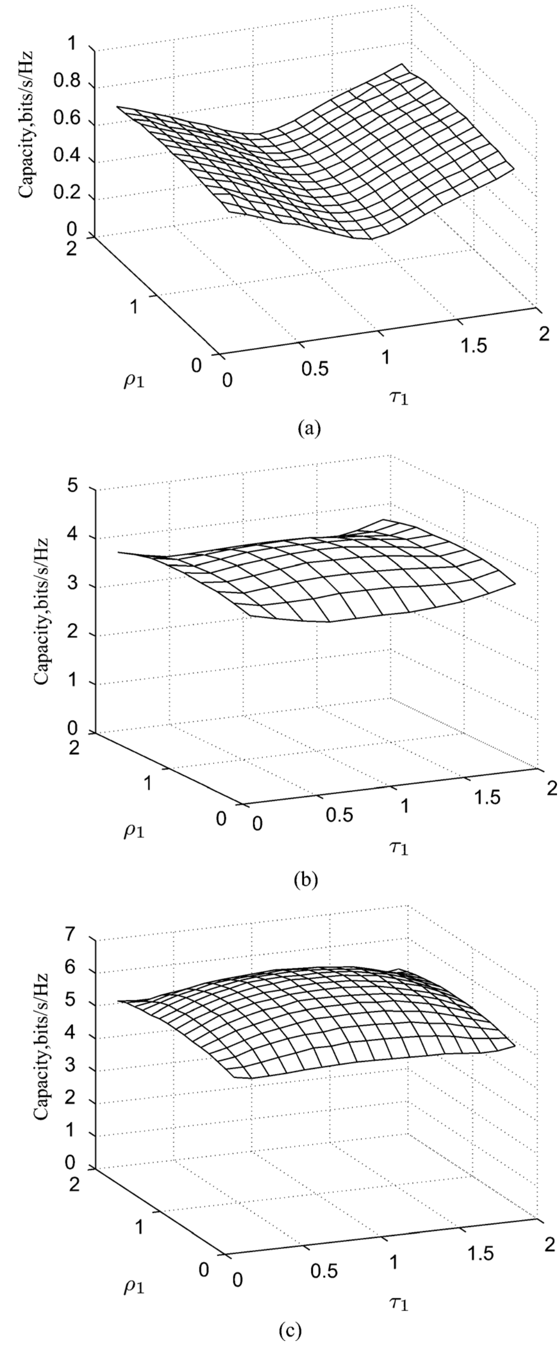

Fig. 4. Capacity, for $T = \mathrm{diag}\{\tau_1, \tau_2\}$, with $\mathrm{tr}(T) = t = 2$ and $R = \mathrm{diag}\{\rho_1, \rho_2\}$, with $\mathrm{tr}(R) = r = 2$ and γ given above (a) $\gamma = 0.2$, (b) $\gamma = 5$, (c) $\gamma = 15$.

Fig. 4 shows the capacity of the $t = r = 2$ as a function of $R = \mathrm{diag}\{\rho_1, \rho_2\}$ and $T = \mathrm{diag}\{\tau_1, \tau_2\}$ such that $\mathrm{tr}(R) = \mathrm{tr}(T) = 2$. For $\rho_1 = \tau_1 = 1$, we have the well-known i.i.d. case studied in [2]. Note that for variations in τ_1, the capacity may increase (at low SNR) for more correlated channels, as beamforming on a correlated channel may outperform equal power on the i.i.d. channel, while for high SNR, we must have quite severe correlation before beamforming offers an advantage over equal power. There is no capacity benefit achieved by having greater receive-side correlation.

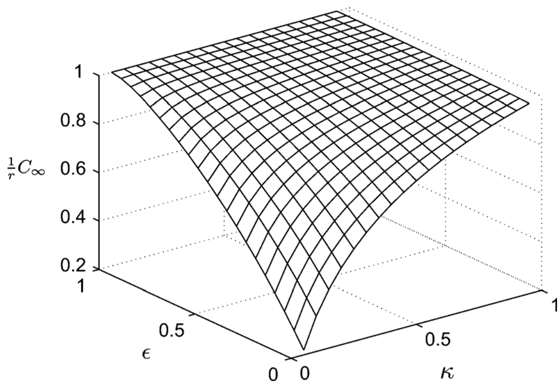

Fig. 5. Capacity $\frac{C_\infty}{r}$ for κ and ϵ, with $T = I$.

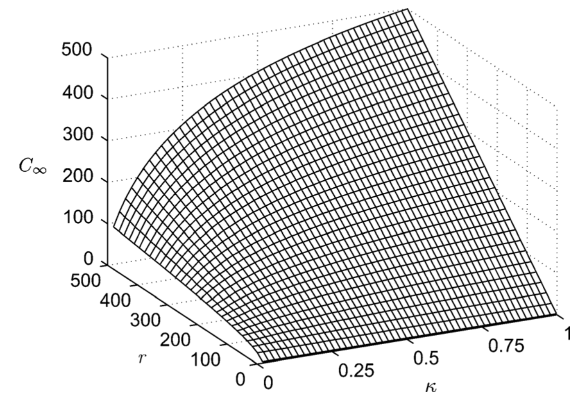

Fig. 6. Capacity C_∞ for κ and r, given $\epsilon = 0.07$ and $T = I$.

Now consider the case that $t \gg r$. For simplicity, $\gamma = 1$. Fig. 5 shows the effective growth rate of the channel capacity as a function of $\kappa = k/r$ and ϵ. As shown above, the growth is linear, although the proportionality constant is a function of the correlation of the channel. For this reason, we define the normalized capacity

$$C_{\text{norm}}(\Sigma_m) = \frac{C_\infty}{m} = \frac{C_\infty}{r} \tag{46}$$

C_{norm} gives the *rate of growth* of the capacity of the MIMO channel for increasing r. For the i.i.d. channel, $C_\infty/r = 1$ resulting in the largest rate of growth in channel capacity for increasing numbers of receivers. This corresponds to the i.i.d. channel being entropy maximizing [2], and hence, an i.i.d. channel exhibits the greatest increase in capacity for a given increase in numbers of receivers.

Eigenvalues were generated as k random samples with the last $r - k$ values set to unity. The entire eigenvalue set was then normalized in accordance with (44). Fig. 5 shows the result of several Monte Carlo simulations.

The value of κ gives the ratio of the number of distinct eigenvalues, to the number of (smaller) equal eigenvalues, independent of r. For $\kappa \to 0$ (no distinct eigenvalues), the capacity of the correlated channel approaches the capacity of the i.i.d. channel and $C_\infty/r \to 1$. For other values of κ and ϵ, the correlation becomes more severe giving a reduction in capacity growth rate. Note that C_∞/r diminishes most rapidly as κ varies form 0 to small, nonzero, values. This is seen in both Figs. 5 and 6. In Fig. 5, note that the plot does not extend completely to $\kappa = 0$ for $\epsilon > 0$, although it is possible to plot arbitrarily small values of κ, resulting in a step from $C_\infty/r = 1$. This is because for $\kappa = 0$, there is no variation in eigenvalues, and so nonzero values of ϵ are meaningless.

Fig. 6 shows C_∞ with respect to r for a particular $\epsilon = 0.07$ and various κ. As can be seen, for any value of κ, the growth is "linear," although the rate of growth is a function of κ. For highly correlated Σ_m, the growth rate becomes negligible.

The most severe correlation can be seen for small, nonzero, values of κ with large ϵ. This is due to (44). For large ϵ and small κ, the (large) variance of the eigenvalues is given by large values of a small number of eigenvalues—the remainder being very small to satisfy (44). This approximates a channel H which is singular, with only a small number of significant eigenvalues—the rest being approximately zero.

As κ becomes closer to unity the effect of a large variance is distributed over many eigenvalues, and consequently, the effective correlation of the channel is reduced.

Intuitively, this can be seen from a beamforming perspective. For a fixed variation in eigenvalues ϵ, small $\kappa > 0$ values correspond to a small number of dominant signal paths—such as occurs in Ricean fading. In the extreme case, the channel has only one significant eigenvector or beam direction, the rest being negligible. In this case, the total variation in eigenvalues is accommodated by one direction.

A transmitter using equal power white Gaussian signals wastes the majority of its power: only a tiny component of the signals uses the particular beam direction, and the capacity is correspondingly reduced. As κ increases, the number of dominant signal paths increases. For a particular ϵ value, this corresponds to a reduction in the difference between the largest eigenvalue and the smallest eigenvalue, so that a larger number of signal directions have similar power gains. Note that Fig. 6 does not extend to $\kappa = 0$ where all the eigenvalues of Σ_m are equal, as there can be *no variation* in eigenvalue magnitudes, as such the regions $\epsilon > 1$ and $\kappa = 0$ are mutually exclusive.

VII. Conclusion

The main results of this paper are expressions for the capacity and corresponding optimal input density of a correlated ergodic MIMO channel, where the channel is assumed to have a (Kronecker) correlated normal distribution. These expressions are valid under the assumption that the noise experienced by the receiver is independent (although not necessarily i.i.d.).

The optimal input density is Gaussian with a covariance matrix which satisfies a certain fixed-point equation, being a function of the receive and transmit correlation matrices. This equation is a natural generalization of the well-known water filling solution (although the resulting optimal power allocation is not water filling). Using this fixed-point equation, an iterative method for numerical computation of the optimal covariance has been suggested. Numerical capacity results obtained using this algorithm were compared with the mutual information given by the equal power strategy. In the case of receive-only correlation, the optimal input strategy is equal power independent Gaussian signals.

In the case of receive- or transmit-only (single-ended) correlation, a closed-form expression for the capacity has been given, which is a function of the correlation matrix eigenvalues. It was shown that the capacity may be written in terms of a summation of single-input multiple-output channel capacities. High- and low-SNR asymptotic expressions were also obtained in the general case, along with the corresponding asymptotically optimal input distributions.

An asymptotic closed-form expression has been given for the case when the number of transmit antennas is increased, while holding the number of receive antennas fixed and keeping $T = I$. This asymptotic limiting capacity C_∞ has two components: the first component is due to the dominant (distinct) eigenvalues of the receive correlation matrix, and the second component is due to the remaining (equal) eigenvalues. The equal eigenvalue component has a spherical distribution, corresponding to the well-known i.i.d. channel case.

It was shown that receive-end correlated channels exhibit a linear growth in (effective) channel capacity with respect to the rank of the correlation matrix. The proportionality of this growth is a function of the variance of the eigenvalues of the correlation matrix. In the case where the correlation matrix is full rank, and the number of transmitters t is large, this corresponds to a linear growth with respect to number of receivers.

Where the variance of the eigenvalues of the receive covariance matrix R are small, it was shown that the capacity of correlated MIMO channel becomes close to the i.i.d. case. It was also shown that there are some regions (where the ratio of correlation matrix eigenvalues is large) for which the rate of linear growth in capacity is extremely small. The i.i.d. channel is optimal in terms of linear growth, giving the greatest increase in channel capacity, for each additional receiver.

Finally, it was noted that for high- and low-SNR levels, the correlation R of the channel does not play a significant role in the capacity: only the most extreme (near-singular) correlation matrices result in a significant loss in capacity.

Appendix A
Proofs

Proof of Lemma 1: For unitary U, the distribution of $\tilde{H}U$ is the same as $\tilde{H}$. Fixing $T = I_t$, apply the result preceding [2, Th. 1] which ensures that diagonal Q is sufficient. Since $\log\det(I + AQA^*)$ is strictly convex on positive definite matrices [87], [89] use [2] to note that

$$\begin{aligned}&\mathrm{E}_{\tilde{H}}\left\{\log\det\left(I + R\tilde{H}Q(R\tilde{H})^*\right)\right\}\\&\quad\le \frac{1}{t!}\sum_{\gamma i}\mathrm{E}_{\tilde{H}}\left\{\log\det\left(I + R\tilde{H}\gamma i Q(R\tilde{H}\gamma i)^*\right)\right\}\end{aligned}\tag{47}$$

where the sum is over all permutation matrices γi. The inequality of (47) is an equality for $Q = \alpha I$, and thus $Q = PI_t/t$ is capacity achieving.[5] ■

[5]We thank I. E. Telatar for this observation.

Proof of Lemma 2: The proof results from a series of mutual information preserving transformations and the isotropic property of the normal distribution. Let $\sim$ denote "has same distribution as." Then, for $H \sim N_{r,t}(0, I)$ and singular value decompositions $T^{1/2} = U_T\Sigma_T^{1/2}V_T^*$, $R^{1/2} = U_R\Sigma_R^{1/2}V_R^*$

$$\begin{aligned}y &= U_R\Sigma_R^{1/2}V_R^*HU_T\Sigma_T^{1/2}V_T^*x + w\\&\sim U_R\Sigma_R^{1/2}H\Sigma_T^{1/2}V_T^*x + w\\&\sim \Sigma_R^{1/2}H\Sigma_T^{1/2}V_T^*x + w.\end{aligned}$$

Now the capacity-achieving input distribution is Gaussian with a certain covariance $\hat{Q} = \mathrm{E}\{xx^*\}$. The unitary matrix V_T may be absorbed into the transmit covariance since multiplication by a unitary matrix will not affect the trace constraint. ■

Proof of Theorem 1: From [2] and [81], the optimal input distribution for a Gaussian fading channel, with exact knowledge of the channel realization H at the receiver, is Gaussian. From [25], [82], and [84], it is already known that the eigenvectors of the optimizing input covariance $\hat{Q}$ diagonalize T. This means that the optimal transmit strategy is to transmit independent symbols in the direction of the eigenvectors of T. For this reason, and with recourse to Lemma 2, it suffices to consider only diagonal T and R, in which case $\hat{Q}$ is also diagonal.

The goal is to find the optimal covariance $\hat{Q} = \operatorname{diag}(q_1, \ldots, q_t)$ subject to $\sum_{i=1}^t q_i \le \gamma$ and $q_i \ge 0,\ i = 1, 2, \ldots, t$, which maximizes $\mathrm{E}_H\{I(x; y \mid H)\}$ for the channel (2) with $T = \operatorname{diag}(\tau_1, \ldots, \tau_t)$, and $R = \operatorname{diag}(\rho_1, \ldots, \rho_r)$. It will be assumed that $\tau_1 \ge \cdots \tau_t > 0$.

The objective function we wish to maximize is

$$f(Q) = \mathrm{E}_H\{\log\det(I + HQH^*)\}\tag{48}$$
$$= \mathrm{E}_S\{\log\det(I + SQ)\}\tag{49}$$

where $S = H^*H$. Equation (49) is known to be concave on positive definite matrices Q, [2], [87], [89]. The conditions for optimality are [90, p. 87]

$$\begin{aligned}\frac{\partial f(Q)}{\partial q_k} &= \mu \quad q_k > 0\\ \frac{\partial f(Q)}{\partial q_k} &< \mu \quad q_k = 0.\end{aligned}$$

The partial differentiation may be taken inside the expectation, and using $\partial\log\det(X)/\partial X = X^{-1}$ and the chain rule of calculus results in

$$\frac{\partial f(Q)}{\partial q_k} = \mathrm{E}\left\{\frac{\partial\log\det(I + SQ)}{\partial q_k}\right\}\tag{50}$$
$$= \mathrm{E}\left\{\operatorname{tr}\left[(I + SQ)^{-1}\frac{\partial(I + SQ)}{\partial Q}\frac{\partial Q}{\partial q_k}\right]\right\}\tag{51}$$
$$= \mathrm{E}\left\{\operatorname{tr}\left[(I + SQ)^{-1}SE_{kk}\right]\right\}\tag{52}$$
$$= \mathrm{E}\left\{\left((I + SQ)^{-1}S\right)_{kk}\right\}\tag{53}$$

where E_{ij} is an all-zero matrix apart from entry i, j, which is 1. ■

Proof of Theorem 2: The basic quantity of interest is the mutual information $I \triangleq I(x;y)$ of the channel (1) with "white" Gaussian input distribution, $\mathrm{E}\{xx^*\} = \gamma I_t/t$. The capacity is found by then substituting $T \mapsto \hat{Q}T$ with $\hat{Q}$ given by Theorem 1.

From [11, p. 79], the density of $H \sim N_{r,t}(0, R \otimes T)$ is

$$p(H) = (2\pi)^{-rt} \det(R)^{-r} \det(T)^{-t} \cdot \mathrm{etr}\left(-R^{-1}HT^{-1}H^*\right) \tag{54}$$

where $\mathrm{etr}(S)$ denotes $\exp\{\mathrm{tr}(S)\}$ (see, for example, [11]). Then, the mutual information of interest is

$$I = \int_H \log\det\left(I_r + HQH^*\right) p(H)\, dH. \tag{55}$$

Two transformations will be applied to obtain the result. First, the (singular value decomposition) transformation $H \to UDV$ with U, V unitary, and D diagonal with the $0 < q \le m$ nonzero singular values of H, in descending order, and second $DD^* \to \Lambda$. The Jacobians are [71], [91]

$$H \to UDV: \quad dH = \det(D)^{2(n+m-2q+1)} \prod_{i<j}^{q} \left(D_{ii}^2 - D_{jj}^2\right)^2 \, dD\, dU\, dV \tag{56}$$

$$DD^* \to \Lambda : dD = \left[2^m \prod_i D_{ii}^2 dD\right]^{-1} \tag{57}$$

where dU, dV are invariant Haar measures on the Stiefel manifold. The result of [91] has been modified to account for complex variables, transcribing the result of [92, Th. 3.2].

Applying $H \to UDV$ gives

$$\begin{aligned} I = (2\pi)^{-rt} \det(R)^{-r} \det(T)^{-t} &\int_D \log\det\left(I + \frac{\gamma}{t}DD^*\right) \\ &\cdot 2^m \det(D)^{2(n+m-2q+1)} \prod_{i<j}^{q} (D_{ii}^2 - D_{jj}^2)^2 \\ &\cdot \int_U \int_V \mathrm{etr}\left(-R^{-1}UDVT^{-1}V^*D^*U^*\right) \, dD\, dU\, dV. \end{aligned} \tag{58}$$

The integrals with respect to U and V in (58) are over all unitary matrices. Repeated application of (73) gives

$$\begin{aligned} I = \frac{\pi^{2n}}{\tilde{\Gamma}_m(n)\tilde{\Gamma}_m(m)} &\det(R)^{-r} \det(T)^{-t} \\ &\cdot \int_D \log\det\left(I + \frac{\gamma}{t}DD^*\right) \det(D)^{2(n+m-2q+1)} \\ &\cdot \prod_{i<j}^{q} \left(D_{ii}^2 - D_{jj}^2\right)^2 {}_0F_0\left(-R^{-1}, DD^*, T^{-1}\right) \, dD. \end{aligned} \tag{59}$$

Now applying the second transformation, $DD^* \to \Lambda$ where the matrix Λ is diagonal with entries λ_i (ith largest eigenvalue of HH^*) and $\lambda_i = |D_i|^2$

$$\begin{aligned} I = \frac{\pi^{2n}}{\tilde{\Gamma}_m(n)\tilde{\Gamma}_m(m)} &\det(R)^{-r} \det(T)^{-t} \\ &\cdot \int_\Lambda \log\det\left(I + \frac{\gamma}{t}\Lambda\right) \det(\Lambda)^{(n+m-2q)} \prod_{i<j}^{q} (\lambda_i - \lambda_j)^2 \\ &\cdot {}_0F_0\left(-R^{-1}, \Lambda, T^{-1}\right) d\Lambda. \end{aligned}$$

Finally, noting $m = q$ gives the desired result. ■

Alternate Proof: The previous proof takes care of possible problems due to $t < r$ through use of the singular value decomposition. For the case $t \ge r$, the following calculation applies. The quantity of interest is

$$\mathrm{E}\left\{\log\det\left(I + X\hat{Q}X^*\right)\right\} \tag{60}$$

where from Theorem 1, $\hat{Q} = \mathrm{diag}\{q_1, \ldots, q_t\}$, and $X \sim N_{r,t}(0, R \otimes T)$. The variable $S = X\hat{Q}X$ has density [71, eq. (7.2.5)]

$$\begin{aligned} p(S) = \frac{1}{2^{tr}\Gamma_p(t)} &\det(\hat{Q}T)^{-t} \det(R)^{-r} \det(S)^{t-r} \\ &\cdot {}_0F_0\left(T^{-1}\hat{Q}^{-1}, -R^{-1}S\right). \end{aligned} \tag{61}$$

Using the change of variables $S \to U\Lambda U^*$ (eigenvalue decomposition) with Jacobian [92, p. 33] $J(S \to U, \Lambda) = \prod_{i<j}(\lambda_i - \lambda_j)^2$ and integrating over U gives the above result. Generalizing (61) for $t < r$ requires similar effort as the previous proof. ■

Proof of Theorem 3: If either $T = I_t$ or $R = I_r$ then $I(x;y) = \mathrm{E}\{\log\det(I + S)\}$ where, according to the notation of [71, eq. (7.2.5)] $S \sim Q_{m,n}(A, I_t, I_r)$, where $A = \mathrm{diag}\{a_i\}$. In the case of $T = I_t$ (where $\hat{Q} = I$ is optimal), $A = R$. For $R = I_r$, then $A = T\hat{Q}$. The remainder of the proof will be carried through for $A = R$. Change of variables may be used to obtain the $A = T\hat{Q}$ result.

The strategy of the proof is to obtain the density of an unordered eigenvalue of S. Now from (62).

Let $\Delta = \mathrm{diag}\{\delta_1, \ldots, \delta_n\}$ with $\delta_1 > \cdots > \delta_n > 0$, and $\Gamma = \mathrm{diag}\{\gamma_1, \ldots, \gamma_n\}$ with $\gamma_1 > \cdots > \gamma_n > 0$. From [83, eq. (4.6)]

$$\begin{aligned} {}_0F_0(\Delta, \Gamma) = \det &\begin{bmatrix} {}_0\mathcal{F}_0(\delta_1\gamma_1) & \cdots & {}_0\mathcal{F}_0(\delta_n\gamma_1) \\ \vdots & & \vdots \\ {}_0\mathcal{F}_0(\delta_1\gamma_n) & \cdots & {}_0\mathcal{F}_0(\delta_n\gamma_n) \end{bmatrix} \\ &\cdot \frac{\tilde{\Gamma}_m(m)\pi^{-m(m-1)/2}}{\prod_{i<j}(\delta_i - \delta_j)(\gamma_i - \gamma_j)} \end{aligned} \tag{62}$$

where ${}_0\mathcal{F}_0(\cdot)$ is the scalar hypergeometric function. Note that ${}_0\mathcal{F}_0(x) = \exp(x)$ for scalar x. By repeated application of l'Hôpital's rule, the limits $\lim_{\epsilon\to 0}\delta_i = \delta_j+\epsilon$ and $\lim_{\epsilon\to 0}\delta_k = \epsilon$ may be taken to generalize (62) to the case $\delta_1 \geq \cdots \geq \delta_n \geq 0$, and $\gamma_1 \geq \cdots \geq \gamma_n \geq 0$. This allows padding either of the matrices Δ, Γ with zero eigenvalues to ensure equal dimension, and reduces ${}_0F_0(\Delta, \Gamma)$ to ${}_0F_0(\alpha\Gamma) = \mathrm{etr}(\alpha\Gamma)$ in the case that $\Delta = \alpha I$. See, for example, [93].

From (62), we may write

$$ {}_0F_0\left(R^{-1}, \Lambda\right) = \frac{\det\left[{}_0\mathcal{F}_0\left(\frac{\lambda_j}{\rho_i}\right)\right]\tilde{\Gamma}_m(m)\pi^{-m(m-1)/2}}{\prod_{i<j}\left(\frac{1}{\rho_i}-\frac{1}{\rho_j}\right)(\lambda_i-\lambda_j)}. \quad (63)$$

Combining (61) and (63), the density for the ordered eigenvalues is

$$\begin{aligned} p_o(\lambda_1,\ldots,\lambda_m) &= K_{m,n,R}\det\left[{}_0\mathcal{F}_0\left(\frac{\lambda_i}{\rho_j}\right)\right]\cdot\prod_i \lambda_i^{m-n}\prod_{i<j}(\lambda_i-\lambda_j)\\ &= K_{m,n,R}\sum_\alpha(-1)^{\mathrm{per}(\alpha)}\prod_i \exp\left(-\frac{\lambda_i}{\rho_{\alpha(i)}}\right)\\ &\qquad\cdot\prod_i\lambda_i^{m-n}\prod_{i<j}(\lambda_i-\lambda_j) \end{aligned} \quad (64)$$

where the constant $K_{m,n,R}$ is

$$K_{m,n,R} = \left(\det(R)^r\prod_i(i-1)!\prod_{i<j}\left(\frac{1}{\rho_i}-\frac{1}{\rho_j}\right)\right)^{-1} \cdot\frac{\pi^{m(m-1)}}{\tilde{\Gamma}_m(n)\tilde{\Gamma}_m(m)}.$$

Now the expression (64) is invariant to permutations of the eigenvalues λ_i, and thus the *unordered* density $p_u(\lambda_1,\ldots,\lambda_m) = p_o(\lambda_1,\ldots,\lambda_m)/m!$.

Note also that the expression $\prod_{i<j}(\lambda_i-\lambda_j)$ in (64) may be written as a Vandermonde determinant

$$\prod_{i<j}(\lambda_i-\lambda_j) = -\sum_\beta(-1)^{\mathrm{per}(\beta)}\prod_i\lambda_i^{\beta(i)-1}$$

which means

$$\begin{aligned} &p_u(\lambda_1,\ldots,\lambda_m) =\\ &\frac{K_{m,n,R}}{m!}\sum_{\alpha,\beta}(-1)^{\mathrm{per}(\alpha)+\mathrm{per}(\beta)+1}\prod_i e^{\lambda_i/\rho_{\alpha(i)}}\lambda_i^{n-m+\beta(i)-1} \end{aligned} \quad (65)$$

for permutations α and β. From [86, 3.381.4], for $\Re(a) > 0$ and $\Re(k) > -1$

$$\int_0^\infty e^{-x/a}x^k\,dx = a^{k+1}\Gamma(k+1).$$

Making the substitution $k = n-m+\beta(i)-1$ and $a = \rho_{\alpha(i)}$, the $\lambda_2,\ldots,\lambda_m$ may be integrated out to give the probability density of a randomly chosen eigenvalue λ

$$\begin{aligned} &p_1(\lambda)\\ &= \frac{K_{m,n,R}}{m!}\sum_{\alpha,\beta}(-1)^{\mathrm{per}(\alpha)+\mathrm{per}(\beta)+1}e^{-\lambda/\rho_{\alpha(1)}}\lambda^{n-m+\beta(1)-1}\\ &\quad\cdot\prod_{i=2}^m\left(\rho_{\alpha(i)}\right)^{n-m+\beta(i)}\Gamma(n-m+\beta(i))\\ &\quad= \frac{K_{m,n,R}}{m!}\sum_{\alpha,\beta}(-1)^{\mathrm{per}(\alpha)+\mathrm{per}(\beta)+1}\\ &\quad\cdot\frac{\prod_{i=1}^m\left(\rho_{\alpha(i)}\right)^{n-m+\beta(i)}\Gamma(n-m+\beta(i))}{\Gamma(n-m+\beta(1))\left(\rho_{\alpha(1)}\right)^{n-m+\beta(1)-1}}\\ &\qquad\cdot e^{-\lambda/\rho_{\alpha(1)}}\lambda^{n-m+\beta(1)-1}. \end{aligned}$$

For permutations α, β, and π, the identities are used to obtain the final results:

$$\begin{aligned} &\sum_{\alpha,\beta}(-1)^{\mathrm{per}(\alpha)+\mathrm{per}(\beta)}f(\alpha_1,\beta_1)\prod_{i=1}^m g(\alpha(i),\beta(i))\\ &= (m-1)!\sum_j^m\sum_\pi(-1)^{\mathrm{per}(\pi)}f(j,\pi(j))\prod_{i=1}^m g(i,\pi(i)) \end{aligned} \quad (66)$$

$$= (m-1)!\sum_{k=1}^m\det\left\{g(i,j)\left[f(k,j)\right]^{\delta_{ik}}\right\} \quad (67)$$

where (66) arises by noting $\alpha(i) = \pi(\beta(i))$, and δ_{ik} denotes the Kronecker delta in (67).

Substituting

$$\begin{aligned} f(\alpha(1),\beta(1)) &= \frac{e^{-\lambda/\rho_{\alpha(1)}}\lambda^{n-m+\beta(1)-1}}{\Gamma(n-m+\beta(1))\left(\rho_{\alpha(1)}\right)^{-\beta(1)}}\\ &\cdot g(\alpha(i),\beta(i)) = \left(\rho_{\alpha(i)}\right)^{n-m+\beta(i)} \end{aligned}$$

results in

$$\begin{aligned} p_1(\lambda) &= \frac{K_{m,n,R}\tilde{\Gamma}_m(n-1)}{m\pi^{m(m-1)/2}}\sum_{j=1}^m\sum_\pi(-1)^{\mathrm{per}(\pi)}\\ &\cdot\frac{e^{-\lambda/\rho_j}\lambda^{n-m+\pi(j)-1}}{\Gamma(n-m+\pi(j))(\rho_j)^{n-m+\pi(j)}}\prod_{i=1}^m(\rho_i)^{n-m+\pi(i)}. \end{aligned} \quad (68)$$

The result of (68) gives the density for a randomly chosen eigenvalue λ which can be used to calculate the expectation of $\log\det(\cdot)$.

The expectation of interest is $\int_0^\infty p_1(\lambda)\log(1+\gamma\lambda)$ with $p_1(\lambda)$ given by (68). This integral has the same form as [2, Example 3], i.e., the capacity of a Gaussian random channel with $t = 1$ transmitters and r receivers. Denote the integral as

$$\tilde{C}(r,1,\gamma) = \frac{1}{\Gamma(r)}\int_0^\infty \log(1+\gamma\lambda)\lambda^{r-1}e^{-\lambda}. \quad (69)$$

A change of variables gives

$$\int_0^\infty e^{-\lambda/\rho_j}\lambda^{n-m+\pi(j)-1}\log(1+\gamma\lambda) = \Gamma(n-m+\pi(j))\rho_j^{n-m+\pi(j)}\tilde{C}(n-m+\pi(j),1,\rho_j\gamma). \tag{70}$$

Substituting (70) into (68) gives the desired result. ■

APPENDIX B
PROPERTIES OF THE HYPERGEOMETRIC FUNCTION

This section briefly outlines some properties of the hypergeometric function and zonal polynomials of matrix arguments. Many of these properties may be found in [11], [71], [83], and [94]–[97]

Let S be a symmetric $p\times p$ matrix, and V_k be the subspace of polynomials of degree k in $p(p+1)/2$ distinct elements of S

$${}_0F_0\,(S) = \operatorname{etr}(S) = \sum_k^\infty \sum_\kappa \frac{C_\kappa(S)}{k!}. \tag{71}$$

The zonal polynomial $C_\kappa(S)$ is the component of $(\operatorname{tr} S)^k$ in the subspace V_k, and $\sum_\kappa$ denotes summation over all partitions κ.

The hypergeometric function of n matrix arguments ${}_0F_0\,(S_1,\ldots,S_n)$, where each S_i is a symmetric $p_i\times p_i$ matrix, is given by

$${}_0F_0\,(S_1,\ldots,S_n) = \sum_k^\infty \sum_\kappa \left(\frac{C_\kappa(S_1)}{k!}\prod_{i=2}^n \frac{C_\kappa(S_i)(2\pi)^{q-p_i}}{C_\kappa(I_q)}\right) \tag{72}$$

where $q=\max(p_1,\ldots,p_n)$.

The order of matrix arguments is unimportant. If any $S_i = I$, the hypergeometric function reduces to an equivalent function with $n-1$ arguments, i.e., ${}_0F_0\,(S_1,\ldots,S_{n-1},I) = {}_0F_0\,(S_1,\ldots,S_{n-1})$. The ${}_0F_0\,(\cdots)$ function is generated by simple extension of the hypergeometric function of two matrix arguments [71, definition 1.6.2]. It can be found by integrating a hypergeometric function of one matrix argument over multiple Stiefel manifolds:

$$\int_{Q_1\cdots Q_n}\cdots\int {}_0F_0\,(S_1Q_1\cdots S_n\cdots Q_1{}^*)\prod_i^{n-1}[dQ_i] \tag{73}$$

where $[dQ_i]$ are normalized invariant Haar measures on the orthogonal group $O(p_i,p_{i+1})$, [71]. This extends the normal two-matrix argument hypergeometric function, by noting that for nonequal dimension S_i

$$\operatorname{tr}(SQ_1TQ_1') = \operatorname{tr}\left([\,S:0\,]\begin{bmatrix}Q_1\\ Q_2\end{bmatrix}T\,[\,Q_1:Q_2\,]\right) \tag{74}$$

with $[Q_1:Q_2]\in O(q)$, so we may pad each symmetric matrix with zeros. Repeated integration over the (larger) $O(q)$ manifold may be performed, noting that the "additional" terms (due to Q_2) result in a volume discrepancy of $(2\pi)^{q-p_i}$ for each integral.

The normalized Haar measure is related to the nonnormalized Haar measure by

$$dQ = \frac{\pi^n}{\tilde{\Gamma}_m(n)}[dQ] \qquad \int_Q dQ = (2\pi)^m\frac{\pi^n}{\tilde{\Gamma}_m(n)}. \tag{75}$$

The hypergeometric function of matrix argument ${}_0F_0\,(X)$ is an extension of the scalar hypergeometric function ${}_0\mathcal{F}_0\,(x) = \exp(x)$.

ACKNOWLEDGMENT

The authors wish to thank I. E. Telatar for several valuable, and helpful comments which improved the scope and clarity of the paper. We also thank A. Lozano for facilitating the publication of the manuscript and assisting with the section on subsequent developments. The authors would also like to thank the anonymous reviewers for their helpful comments.

REFERENCES

[1] L. W. Hanlen and A. J. Grant, "Capacity analysis of correlated MIMO channels," in *IEEE Int. Symp. Inf. Theory*, Yokohama, Japan, Jun. 29–Jul. 4 2003, p. 263.
[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.
[4] A. J. Grant, "Rayleigh fading multiple-antenna channels," *EURASIP J. App. Sig. Proc., Special Issue Space-Time Coding, Part I*, no. 3, pp. 316–329, Mar. 2002.
[5] P. W. Wolniansky, G. D. Golden, G. J. Foschini, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals, Syst., Electron.*, Sep. 29–Oct. 2 1998, pp. 295–300.
[6] D. Gesbert, H. Bölcskei, D. Gore, and A. Paulraj, "MIMO wireless channels: Capacity and performance prediction," in *Proc. IEEE Global Commun. Conf.*, 2000, vol. 2, pp. 1083–1088.
[7] L. W. Hanlen and M. Fu, "Multiple antenna wireless communication systems: Limits to capacity growth," in *Proc. IEEE Wireless Commun. Net. Conf.*, Orlando, FL, Mar. 17–21, 2002, pp. 172–176.
[8] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.
[9] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of a class of large dimensional random matrices," *J. Mult. Anal.*, vol. 54, pp. 175–192, Aug. 1995.
[10] D. Tse, "Multiuser receivers, random matrices and free probability," in *Proc. 37th Annu. Allerton Conf. Commun., Control Comput.*, Monticello, IL, Sep. 1999, pp. 1055–1064.
[11] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, ser. Wiley Series in Probability and Mathematical Statistics. New York: Wiley, 1982.
[12] 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, no. 3, pp. 502–513, Mar. 2000.
[13] P. Rapajic, "Information capacity of a multipath mobile communication channel with large number of receiving antennas," in *Proc. IEEE Inf. Theory Workshop*, Cairns, Qld., Australia, Sep. 2–7, 2001, pp. 104–106.
[14] N. Chiurtu, B. Rimoldi, and E. Telatar, "Dense multiple antenna systems," in *Proc. IEEE Inf. Theory Workshop*, Cairns, Qld., Australia, Sep. 2–7, 2001, pp. 108–109.
[15] A. Scaglione, "Statistical analysis of the capacity of MIMO frequency selective Rayleigh fading channels with arbitrary number of inputs and outputs," in *Proc. IEEE Int. Symp. Inf. Theory*, Lausanne, Switzerland, Jun. 30–Jul. 5 2002, p. 278.
[16] M. T. Ivrlac and J. A. Nossek, "On the impact of correlated fading for MIMO-systems," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2002, pp. 655–658.
[17] J. W. Wallace and M. A. Jensen, "The capacity of MIMO wireless systems with mutual coupling," in *Proc. IEEE Veh. Technol. Conf.*, Vancouver, BC, Canada, 2002, pp. 696–700.

[18] J. W. Wallace and M. A. Jensen, "Intrinsic capacity of the MIMO wireless channel," in *Proc. IEEE Veh. Technol. Conf.*, Vancouver, BC, Canada, 2002, pp. 701–705.

[19] J. W. Wallace and M. A. Jensen, "Modeling the indoor MIMO wireless channel," *IEEE Trans. Antennas Propag.*, vol. 50, no. 5, pp. 591–599, May 2002.

[20] M. Chiani, M. Z. Win, and A. Zanella, "The distribution of eigenvalues of a Wishart matrix with correlation and application to MIMO capacity," in *Proc. IEEE Global Commun. Conf.*, Dec. 1–5, 2003, vol. 4, pp. 1802–1805.

[21] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[22] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications, Foundations and Trends in Communications and Information Theory*. Boston, MA: Now Publishers, 2004.

[23] E. Biglieri and G. Taricco, *Transmission and Reception with Multiple Antennas: Theoretical Foundations, Foundations and Trends in Communications and Information Theory*. Boston, MA: Now Publishers, 2004.

[24] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[25] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Int. Conf. Commun.*, Jun. 11–14, 2001, vol. 7, pp. 2266–2271.

[26] S. A. Jafar and A. J. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Inf. Theory*, vol. 3, no. 4, pp. 1165–1175, Jul. 2004.

[27] M. Guillaud, M. Debbah, and A. L. Moustakas, "A maximum entropy characterization of spatially correlated MIMO wireless channels," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2007, pp. 1040–1045.

[28] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: Moment-generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.

[29] S. H. Simon and A. L. Moustakas, "Eigenvalue density of correlated complex random Wishart matrices," *Phys E. Rev.*, vol. 69, pp. 065101-1–065101-4, Jun. 2004.

[30] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multiple-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 662–671, Mar. 2006.

[31] V. Raghavan and A. M. Sayeed, "Sublinear capacity scaling laws for sparse MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 345–364, Jan. 2011.

[32] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, Jul. 2006.

[33] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[34] V. V. Veeravalli, Y. Liang, and A. M. Sayeed, "Correlated MIMO wireless channels: Capacity, optimal signaling, and asymptotics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2058–2072, Jun. 2005.

[35] H. Wang, P. Wang, L. Ping, and X. Lin, "Beneficial impact of antenna correlation at mobile units in MIMO MACs with covariance feedback," *IEE Elect. Lett.*, vol. 45, no. 20, pp. 1039–1040, 2009.

[36] S. A. Jafar and A. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 990–997, May 2005.

[37] M. Kang and M. S. Alouini, "Capacity of correlated MIMO Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 143–155, Jan. 2006.

[38] S. K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "Precoding by pairing subchannels to increase MIMO capacity with discrete input alphabets," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4156–4169, Jul. 2011.

[39] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim, "On the capacity achieving covariance matrix for Rician MIMO channels: An asymptotic approach," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1048–1069, Mar. 2010.

[40] F. Dupuy and P. Loubaton, "On the capacity achieving covariance matrix for frequency selective MIMO channels using the asymptotic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5737–5753, Sep. 2011.

[41] P. M. Marques and S. A. Abrantes, "On the derivation of the exact, closed-form capacity formulas for receiver-sided correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1139–1161, Mar. 2008.

[42] X. Li, S. Jin, X. Gao, and K.-K. Wong, "Near-optimal power allocation for MIMO channels with mean or covariance feedback," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 289–300, Jan. 2010.

[43] L. W. Hanlen, "Water-filling with channel correlation," in *Proc. 4th Aust. Commun. Theory Workshop*, Melbourne, Vic., Australia, Feb. 4–5, 2003, pp. 19–23.

[44] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 19, no. 2, pp. 299–533, 1998.

[45] M. L. Morris, M. A. Jensen, and J. W. Wallace, "Superdirectivity in MIMO systems," *IEEE Trans. Antennas Propag.*, vol. 53, no. 9, pp. 2850–2857, Sep. 2005.

[46] J. W. Wallace and M. A. Jensen, "Electromagnetic considerations for communicating on correlated MIMO channels with covariance information," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 543–551, Feb. 2008.

[47] M. A. Jensen and J. W. Wallace, "Capacity of the continuous-space electromagnetic channel," *IEEE Trans. Antennas Propag.*, vol. 56, no. 2, pp. 524–531, Feb. 2008.

[48] S. Jin, M. R. McKay, K.-K. Wong, and X. Li, "Low-SNR capacity of multiple-antenna systems with statistical channel-state information," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2874–2884, Jul. 2010.

[49] M. R. McKay and I. B. Collings, "Improved general lower bound for spatially-correlated Rician MIMO capacity," *IEEE Commun. Lett.*, vol. 10, no. 3, pp. 162–164, Mar. 2006.

[50] M. R. McKay and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3121–3145, Sep. 2005.

[51] C. Zhong, S. Jin, K.-K. Wong, and M. R. McKay, "Ergodic mutual information analysis for multi-keyhole MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1754–1763, Jun. 2011.

[52] G. Levin and S. Loyka, "On the outage capacity distribution of correlated keyhole MIMO channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3232–3245, Jul. 2008.

[53] M. Vu and A. Paulraj, "On the capacity of MIMO wireless channels with dynamic CSIT," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1269–1283, Sep. 2007.

[54] C. D. Charalambous, S. Z. Denic, and C. Constantinou, "Capacity of the class of MIMO channels with incomplete CDI: Properties of mutual information for a class of channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3725–3734, Aug. 2009.

[55] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Uniform power allocation in MIMO channels: A game-theoretic approach," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1707–1727, Jul. 2003.

[56] W. Rhee and G. Taricco, "On the ergodic capacity-achieving covariance matrix of certain classes of MIMO channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3810–3817, Aug. 2006.

[57] G. Taricco and E. Riegler, "On the ergodic capacity of correlated Rician fading MIMO channels with interference," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4123–4137, Jul. 2011.

[58] X. Gao, B. Jiang, X. Li, A. B. Gershman, and M. R. McKay, "Statistical eigenmode transmission over jointly correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3735–3750, Aug. 2009.

[59] L. W. Hanlen and A. Grant, "Optimal transmit covariance for MIMO channels with statistical transmitter side information," [Online]. Available: http://users.cecs.anu.edu.au/leifh/flow/pubs/IT2006-granhan.pdf 2006

[60] L. W. Hanlen and A. J. Grant, "Optimal transmit covariance for MIMO channels with statistical transmitter side information," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, S.A., Australia, Sep. 4–9, 2005, pp. 1818–1822.

[61] R. H. Gohary, W. Mesbah, and T. N. Davidson, "Rate-optimal MIMO transmission with mean and covariance feedback at low SNRS," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3802–3807, Sep. 2009.

[62] A. Lozano, A. M. Tulino, and S. Verdú, *Space-Time Wireless Systems: From Array Processing to MIMO Communications, Chapter Multiantenna Capacity: Myths and Realities*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[63] P. Wijesinghe, U. Gunawardana, and R. Liyanapathirana, "An efficient algorithm for capacity and outage probability estimation in MIMO channels," *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 644–646, Jun. 2011.

[64] J. S. Kwak, J. G. Andrews, and A. Lozano, "MIMO capacity in correlated interference-limited channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, 2007, pp. 106–110.
[65] P. Koev and A. Edelman, "The efficient evaluation of the hypergeometric function of a matrix argument," *Math. Comput.*, vol. 75, no. 256, pp. 833–846, 2006.
[66] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
[67] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.
[68] L. W. Hanlen and A. J. Grant, "On capacity of ergodic multiple-input multiple-output channels," in *Proc. 6th Aust. Commun. Theory Workshop*, Brisbane, Qld., Australia, Feb. 2–4, 2005, pp. 121–124.
[69] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1310–1343, Jun. 2002.
[70] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
[71] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions,Vol. 104 of Monographs and Surveys in Pure and Applied Mathematics*. Boca Raton, FL: Chapman & Hall/CRC, 2000.
[72] J. P. Kermoal, L. Schumacher, K. I. Pederson, P. E. Morgensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.
[73] D. Gesbert, H. Bölcskei, D. Gore, and A. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.
[74] K. Yu and B. Ottersten, "Models for MIMO propagation channels: A review," *Wireless Commun. Mobile Comput.*, vol. 2, pp. 653–666, Nov. 2002.
[75] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
[76] P. Kyritsi and D. C. Cox, "Correlation properties of MIMO radio channels for indoor scenarios," in *Proc. 35th Asimolar Conf. Signals, Syst. Comput.*, Nov. 4–7, 2001, vol. 2, pp. 994–998.
[77] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading environments," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 550–560, Apr. 2002.
[78] H. Ozcelik, M. Herdin, W. J. Weichselberger, and E. Bonek, "Deficiencies of 'Kronecker' MIMO radio channel model," *Electron. Lett.*, vol. 39, no. 16, pp. 1209–1210, Aug. 2003.
[79] T. S. Pollock, "Correlation modelling in MIMO systems: When can we Kronecker?," in *Proc. Aust. Commun. Theory Workshop*, Newcastle, N.S.W., Australia, Feb. 4–6, 2004, pp. 149–153.
[80] Y.-H. Kim and A. Lapidoth, "On the log determinant of noncentral Wishart matrices," in *Proc. IEEE Int. Symp. Inf. Theory*, Yokohama, Japan, Jun. 29–Jul. 4 2003, p. 54.
[81] T. Ericson, "A Gaussian channel with slow fading," *IEEE Trans. Inf. Theory*, vol. IT-16, no. 3, pp. 353–355, May 1970.
[82] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 406–417, Apr. 2003.
[83] K. I. Gross and D. S. P. Richards, "Total positivity, spherical series, and hypergeometric functions of matrix argument," *J. Approx. Theory*, no. 59, pp. 224–246, 1989.
[84] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.
[85] T. W. Anderson, "Asymptotic theory for principal component analysis," *Ann. Math. Statist.*, vol. 34, pp. 122–148, Jul. 1963.
[86] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed. New York: Academic, 2000.
[87] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
[88] G. Szegö, *Orthogonal Polynomials*. Providence, RI: American Math. Soc., 1939.
[89] A. W. Marshall and O. Ingram, *Inequalities: Theory of Majorization and its Applications*, ser. (Mathematics in Science and Engineering, vol. 143). New York: Academic, 1979.
[90] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[91] H. Le and D. G. Kendall, "The Riemannian structure of Euclidean shape spaces: A novel environment for statistics," *Ann. Math. Statist.*, vol. 21, no. 3, pp. 1225–1271, Sep. 1993.
[92] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, 1989.
[93] R. Alexander, C. A. Tracy, and H. Widom, "Random words, Toeplitz determinants and integrable systems. I," in *Random Matrices and their Applications*. New York: Cambridge Univ. Press, 2001, pp. 245–258.
[94] D. S. P. Richards, "Total positivity properties of generalized hypergeometric functions of matrix argument," *J. Statist. Phys.*, vol. 116, pp. 225–231, 2004.
[95] A. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 36, pp. 475–501, 1964.
[96] Y. Chikuse and A. W. Davis, "Some properties of invariant polynomials with matrix arguments and their applications in econometrics," *Ann. Inst. Stat. Math*, vol. 38, pp. 109–122, 1986.
[97] A. W. Davis, "On the construction of a class of invariant polynomials in several matrices, extending the zonal polynomials," *Ann. Inst. Stat. Math*, vol. 33, pp. 297–313, 1981.

Leif Hanlen (S'00–M'03) received the B.E. and B.Comp.Sci. degrees in 2000 and Ph.D. degree in 2003 all from the University of Newcastle Australia. Since June 2003, he has been with National ICT Australia where he is currently Director of the Health Business Team. Dr. Hanlen holds a joint appointment with the Australian National University, in the Research School of Engineering. He is a member of Australian Standards for Telehealth, IT014-012 and the Australian Information Industry Association eHealth taskforce.

Alex Grant (S'93–M'96–SM'03) received the B.E. and Ph.D. degrees from the University of South Australia, in 1993 and 1996, respectively. In 1997, he was a Postdoctoral Research Fellow at the Laboratory for Signal and Information Processing, ETH Zurich. Since March 1998, he has been with the Institute for Telecommunication Research, University of South Australia, where he is now Director and Research Professor of Information Theory. Prof. Grant served as Chairman for the IEEE SA/ACT/Vic Sections Information Theory Society Joint Chapter (2000–2003). He served as Technical Co-Chair for the 2001 (Cairns) and 2009 (Taormina, Italy) IEEE Information Theory Workshop and as general Co-Chair for the 2005 IEEE International Symposium on Information Theory, Adelaide, Australia. He is a Member of the Board of Governors of the IEEE Information Theory Society, and has served as Associate Editor for the IEEE Transactions on Information Theory and for the IEEE Transactions on Wireless Communications. Prof. Grant is co-founder of Cohda Wireless, a leading developer of Safe Vehicle and Connected Vehicle design solutions and services for public safety, outdoor, and automotive wireless-based systems.