Multiple-Antennas and Isotropically-Random Unitary Inputs: The Received Signal Density in Closed-Form

BABAK HASSIBI* THOMAS L. MARZETTA**

*Department of Electrical Engineering
California Institute of Technology, Pasadena, CA 91125

**Mathematical Sciences Research Center
Bell Laboratories, Lucent Technologies
600 Mountain Avenue, Murray Hill, NJ 07974

June 19, 2001

Abstract

An important open problem in multiple antenna communications theory is to compute the capacity of a wireless link subject to flat Rayleigh block-fading, with no channel-state information available either to the transmitter or to the receiver. The isotropically-random (i.r.) unitary matrix–having orthonormal columns, and a probability density that is invariant to pre-multiplication by an independent unitary matrix–plays a central role in the calculation of capacity and in some special cases happens to be capacity-achieving. In this paper we take an important step towards computing this capacity by obtaining, in *closed-form*, the probability density of the received signal when transmitting i.r. unitary matrices. The technique is based on analytically computing the expectation of an exponential quadratic function of an i.r. unitary matrix and makes use of a Fourier integral representation of the constituent Dirac delta functions in the underlying density. Our formula for the received signal density enables us to evaluate the mutual information for any case of interest, something that could previously only be done for single transmit and receive antennas. Numerical results show that at high SNR the mutual information is maximized for $M = \min(N, T/2)$ transmit antennas, where N is the number of receive antennas and T is the length of the coherence interval, whereas at low SNR the mutual information is maximized by allocating all transmit power to a single antenna.

Index Terms—wireless communications, multiple antennas, unitary space-time modulation, isotropically-random unitary matrix

*Formerly with Mathematical Sciences Research Center, Bell Laboratories, Lucent Technologies.

1 Introduction

We consider a single-user multiple-antenna wireless link, subject to flat Rayleigh block-fading, with no channel-state information (CSI) available either to the transmitter or to the receiver. The assumption of no CSI - as opposed to CSI available at the receiver, for example - converts a straightforward problem in Shannon theory [1, 2, 3] into a rather difficult, albeit more realistic, one [4, 5, 6, 7]. The capacity-attaining signals are non-Gaussian, the maximum-likelihood receiver is noncoherent, and there is a hard upper limit on the number of usable transmit antennas.

The isotropically-random unitary matrix - having orthonormal columns, and a probability density that is invariant to pre-multiplication by an independent unitary matrix - plays a central role in the theory of our multiple-antenna link [5]. The capacity-attaining input signal is the product of an isotropically random unitary matrix, and an independent nonnegative real diagonal matrix. In certain limiting regimes [5, 6, 7] the diagonal matrix is constant, and the message is carried entirely by the unitary matrix: a type of modulation called unitary space-time (USTM) [6]. A number of practical considerations make unitary space-time modulation attractive for general usage.

It follows that calculations of capacity or mutual information invariably require taking expectations with respect to an isotropically random unitary matrix. This was done for single transmit and receive antennas [5], but has not been done for more complicated problems. The contribution of this paper is a technique for taking this expectation analytically for certain cases. In particular, when the input signal is isotropically random unitary we obtain a closed-form expression, in the form of a determinant, for the probability density of the received signal. When combined with a simple Monte-Carlo integration, it is possible to compute mutual information for any case of interest.

Section 2 reviews the signal model, and known results. Section 3 contains a derivation of the probability density of the received signal. Our technique uses a Fourier representation for the Dirac delta functions that comprise the isotropic unitary density. An integration with respect to the received signal, which we cannot perform analytically, yields mutual information (Section 4). We perform this integration via Monte-Carlo in Section 5 to compute mutual information. Appendix A reviews the isotropically random unitary density, and derives its Fourier-based version.[1] Appendix B reviews a technique for converting a particular M-dimensional integral into a determinant of an $M \times M$ Hankel matrix, whose entries are one-dimensional integrals. Appendix C provides a simple derivation of the Jacobian determinant for a coordinate transformation from a Hermitian matrix to its eigenvectors and eigenvalues. Finaly Appendix D discusses a certain complex exponential quadratic integral.

2 Signal Model; Mutual Information

2.1 Signal model

Consider a single-user block-fading multiple antenna link with M transmit and N receive antennas described by a propagation matrix that is constant during coherence intervals of length T symbols, after which it jumps to a new independent value for T more symbols, and so on [5]. This is a reasonable model for systems employing some form of time-division-multiplexing, or frequency-hopping. During any coherence interval, a $T \times M$ complex matrix S is transmitted, and a $T \times N$ complex matrix X, is received

$$X = \sqrt{\frac{\rho}{M}} SH + W, \tag{1}$$

where H is a $M \times N$ propagation matrix, and W is a $T \times N$ additive noise matrix. The elements of H and W are independent, zero-mean, circularly symmetric, complex Gaussian with unit variance, whose distribution is denoted $\mathrm{CN}(0, 1)$. The values of H and W are unknown to both the transmitter and the receiver. At each

[1]Throughout the paper, all probability densities are with respect to Lebesgue measure, rather than Haar measure.

time t the transmitted signal has, on average, unit-variance entries,

$$\mathrm{E}\left\{\sum_{m=1}^{M}|s_{tm}|^2\right\} = M. \tag{2}$$

Since S, H, and W are independent, the normalization $\sqrt{\frac{\rho}{M}}$ in (1) implies that ρ is equal to the SNR at each receive antenna, independently of M.

2.2 Mutual information

The independence of H and W over different coherence intervals implies that Shannon coding, applied over many coherence intervals, yields reliable transmission at any rate less than the mutual information,

$$I(X;S) = \mathrm{E}\left\{\log\frac{p(X \mid S)}{p(X)}\right\}, \tag{3}$$

where the conditional density $p(X \mid S)$ is

$$p(X \mid S) = \frac{1}{\pi^{TN}} \cdot \frac{\exp\left[-\mathrm{tr}X^\dagger\left(I_T + \frac{\rho}{M}SS^\dagger\right)^{-1}X\right]}{\det\left(I_T + \frac{\rho}{M}SS^\dagger\right)^N}, \tag{4}$$

where the superscript $\dagger$ denotes conjugate transpose. The maximization of the mutual information over the distribution of S, subject to the power constraint (2), yields the capacity. There is no point in making the number of transmit antennas greater than the symbol duration of the coherence interval [5], so we assume that $M \leq T$. The capacity attaining signal is the product of two independent random matrices [5] ,

$$S = \sqrt{T}\Phi D, \tag{5}$$

where Φ is a $T \times M$ isotropically distributed unitary matrix, and D is a $M \times M$ nonnegative real diagonal matrix. The matrix Φ has orthonormal columns, $\Phi^\dagger \Phi = I_M$, and the joint probability density of its elements is unchanged when Φ is multiplied by any $T \times T$ unitary matrix Ψ that is independent of Φ, i.e., $p(\Phi) = p(\Psi\Phi)$. See [5], or Appendix A, for a derivation of this density. The power constraint (2), combined with (5), implies that

$$\mathrm{E}\left\{\mathrm{tr} D^2\right\} = M. \tag{6}$$

Except for the case $T = M = N = 1$ [4], and for certain limiting cases, little is known concerning the capacity-attaining joint density of the diagonal elements of D.

For the two limiting cases: 1) $T >> M$ [5], or 2) $M \le \min(T/2, N)$ and $\rho >> 1$ [7], capacity is attained when $D = I_M$, and the optimal transmitted signal is proportional to an isotropically random unitary matrix. The transmission of orthonormal signals is called unitary space-time modulation (USTM) [6], and its general use is motivated by several practical considerations [5, 8]. For either of the limiting cases above, mutual information can be evaluated indirectly through asymptotic formulas for capacity [5, 2, 7]. In general, however, the optimizing joint distribution on D is not known. In fact, even the mutual information obtained by $D = I_M$ in the general case was not known and is the main focus of this paper.

2.3 Difficulties in computing mutual information

The combination of (3) and (5) yields a formidable multiple integration,

$$\begin{aligned} I(X;S) &= \int dX \int d\Phi \int dD \, p(X \mid \Phi, D) \cdot p(\Phi) \cdot p(D) \\ &\quad \cdot \; \log \left\{ \frac{p(X \mid \Phi, D)}{\int d\acute{\Phi} \int d\acute{D} \, p(X \mid \acute{\Phi}, \acute{D}) \cdot p(\acute{\Phi}) \cdot p(\acute{D})} \right\}. \end{aligned} \tag{7}$$

The term that arises from the log-numerator reduces to an expectation with respect to D. Most of the trouble results from the log-denominator.

The outer integration of (7) with respect to Φ may be eliminated as follows. For any $T \times T$ unitary matrix Ψ, it follows from (4) that $p(X \mid S) = p(\Psi^\dagger X \mid \Psi^\dagger S)$—see also [5]. Now use this property for $\Psi = \begin{bmatrix} \Phi & \Phi_\perp \end{bmatrix}$, where $\Phi_\perp$ is the $T \times (T-M)$ orthogonal complement to Φ, and perform the following change of variable in (7): $\Psi^\dagger X \to X$. This yields

$$
\begin{aligned}
I(X;S) &= \int dX \int d\Phi \int dD \, p(\Psi^\dagger X \mid \Phi_0, D) \cdot p(\Phi) \cdot p(D) \\
&\quad \cdot \log \left\{ \frac{p(X \mid \Phi, D)}{d\acute{\Phi} \int d\acute{D} \, p(X \mid \acute{\Phi}, \acute{D}) \cdot p(\acute{\Phi}) \cdot p(\acute{D})} \right\} \\
&= \int dX \int d\Phi \int dD \, p(X \mid \Phi_0, D) \cdot p(\Phi) \cdot p(D) \\
&\quad \cdot \log \left\{ \frac{p(\Psi X \mid \Phi, D)}{d\acute{\Phi} \int d\acute{D} \, p(\Psi X \mid \acute{\Phi}, \acute{D}) \cdot p(\acute{\Phi}) \cdot p(\acute{D})} \right\} \\
&= \int dX \int d\Phi \int dD \, p(X \mid \Phi_0, D) \cdot p(\Phi) \cdot p(D) \\
&\quad \cdot \log \left\{ \frac{p(X \mid \Phi_0, D)}{d\acute{\Phi} \int d\acute{D} \, p(X \mid \Psi^\dagger \acute{\Phi}, \acute{D}) \cdot p(\acute{\Phi}) \cdot p(\acute{D})} \right\},
\end{aligned}
$$

where

$$
\Phi_0 \triangleq \Psi^\dagger \Phi = \begin{bmatrix} I_M \\ 0 \end{bmatrix}. \tag{8}
$$

Now make the change of variable $\Psi^\dagger \acute{\Phi} \to \acute{\Phi}$, and use the fact that $\acute{\Phi}$ is isotropically distributed to obtain

$$
I(X;S) = \int dX \int dD \, p(X \mid \Phi_0, D) \cdot p(D) \cdot \log \left\{ \frac{p(X \mid \Phi_0, D)}{d\acute{\Phi} \int d\acute{D} \, p(X \mid \acute{\Phi}, \acute{D}) \cdot p(\acute{\Phi}) \cdot p(\acute{D})} \right\}. \tag{9}
$$

Note that in the above expression we have easily integrated out Φ since the only dependence of the integrand on Φ was through $p(\Phi)$.

The log-numerator term in (9) is

$$
\begin{aligned}
\mathrm{E}\left\{\log p(X \mid \Phi_0, D)\right\} &= -TN\log(\pi e) - N\mathrm{E}\left\{\log\det\left(I_T + \rho\beta\Phi_0 D^2\Phi_0^\dagger\right)\right\} \\
&= -TN\log(\pi e) - N\mathrm{E}\left\{\log\det\left(I_M + \rho\beta D^2\Phi_0^\dagger\Phi_0\right)\right\} \\
&= -TN\log(\pi e) - N\mathrm{E}\left\{\log\det\left(I_M + \rho\beta D^2\right)\right\}, \qquad (10)
\end{aligned}
$$

where we define $\beta \triangleq T/M$.

One further simplification of (9) occurs, because the logarithm depends only on the eigenvalues of $XX^\dagger$, as we discuss in the next section.

All of the above simplifications, combined with numerical integration, were used to evaluate capacity and mutual information for cases where $M = N = 1$ and $T > 1$ [5]. Classical numerical integration of the quadrature type is helpless when dealing with more than a few integration variables, so the same technique cannot be applied to bigger problems. Another possibility - not investigated in this paper - is Monte-Carlo integration. However this could only be used in conjunction with some clever importance sampling, for the reason that $p(X \mid \Phi, D)$ is sharply peaked with respect to Φ. In the next section we show how to calculate analytically the inner integral with respect to the isotropically distributed unitary matrix, for the case where $D = I_M$.

3 Computing $p(X)$

Using (4) and (5), we may write

$$
p(X|S) = \frac{1}{\pi^{NT}} \cdot \frac{\exp\left[-\mathrm{tr}X^\dagger\left(I_T + \rho\beta\Phi D^2\Phi^\dagger\right)^{-1}X\right]}{\det\left(I_T + \rho\beta\Phi D^2\Phi^\dagger\right)^N}
$$

$$= \frac{1}{\pi^{NT}} \cdot \frac{\exp\left\{-\mathrm{tr}\left[X^\dagger X - X^\dagger\Phi\left(\Phi^\dagger\Phi + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger X\right]\right\}}{\det\left(I_T + \rho\beta D^2\Phi^\dagger\Phi\right)^N}$$

$$= \frac{1}{\pi^{NT}} \cdot \frac{\exp\left(-\mathrm{tr}X^\dagger X\right)}{\det\left(I_M + \rho\beta D^2\right)^N} \cdot \exp\left[\mathrm{tr}X^\dagger\Phi\left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger X\right]. \tag{11}$$

This implies that

$$p(X|D) = \frac{1}{\pi^{NT}} \cdot \frac{\exp\left(-\mathrm{tr}X^\dagger X\right)}{\det\left(I_M + \rho\beta D^2\right)^N} \cdot \mathrm{E}_{|D}\exp\left[\mathrm{tr}X^\dagger\Phi\left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger X\right], \tag{12}$$

where the expectation is taken over the isotropically-distributed unitary matrix Φ. Consider now the eigenvalue decomposition of $XX^\dagger$,

$$XX^\dagger = U \underbrace{\begin{bmatrix} \Sigma & 0_{K\times(T-K)} \\ 0_{(T-K)\times K} & 0_{(T-K)\times(T-K)} \end{bmatrix}}_{\triangleq A} U^\dagger,$$

where U is $T \times T$ and unitary, Σ is $K \times K$ real positive diagonal, and $K \triangleq \min(T, N)$. Then

$$\begin{aligned}
\mathrm{E}_{|D}\exp\left[\mathrm{tr}X^\dagger\Phi\left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger X\right] &= \mathrm{E}_{|D}\exp\left[\mathrm{tr}XX^\dagger\Phi\left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger\right] \\
&= \mathrm{E}_{|D}\exp\left[\mathrm{tr}UAU^\dagger\Phi\left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger\right] \\
&= \mathrm{E}_{|D}\exp\left[\mathrm{tr}AU^\dagger\Phi\left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger U\right] \\
&= \mathrm{E}_{|D}\exp\left[\mathrm{tr}A\Phi\left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger\right],
\end{aligned}$$

where to obtain the last equality, we have used the fact that, since Φ is isotropically-distributed, $U^\dagger\Phi$ is also unitary and isotropically-distributed. Defining $B \triangleq \left(I_M + \frac{1}{\rho\beta}D^{-2}\right)^{-1}$, we note that to compute $p(X|D)$

we need to compute $\mathrm{E}\left\{\exp\left(\mathrm{tr}A\Phi B\Phi^\dagger\right)\right\}$, where the expectation is taken over the $T\times M$ isotropically-distributed unitary matrix Φ. For reasons to be made clear shortly, it will be convenient to transform the diagonal matrix $A\geq 0$ into a negative-definite matrix via the following transformation: Choose the scalar $a>0$ such that $aI_T > A$, then

$$\begin{aligned}
\mathrm{E}e^{\mathrm{tr}A\Phi B\Phi^\dagger} &= e^{\mathrm{tr}aB}\mathrm{E}e^{\mathrm{tr}A\Phi B\Phi^\dagger}e^{-\mathrm{tr}aB} \\
&= e^{\mathrm{tr}aB}\mathrm{E}e^{\mathrm{tr}A\Phi B\Phi^\dagger}e^{-\mathrm{tr}aB\Phi^\dagger\Phi} \\
&= e^{\mathrm{tr}aB}\mathrm{E}e^{\mathrm{tr}A\Phi B\Phi^\dagger}e^{-\mathrm{tr}a\Phi B\Phi^\dagger} \\
&= e^{\mathrm{tr}aB}\mathrm{E}e^{-\mathrm{tr}(aI_T-A)\Phi B\Phi^\dagger}.
\end{aligned}$$

which transforms $A\geq 0$ to $A-aI_T<0$.

Now denoting the distribution on Φ by $p(\Phi)$, we need to compute

$$\mathrm{E}e^{\mathrm{tr}A\Phi B\Phi^\dagger} = e^{\mathrm{tr}aB}\int d\Phi\, p(\Phi)e^{-\mathrm{tr}(aI_T-A)\Phi B\Phi^\dagger}.$$

Although it is possible to give an explicit formula for the density as $p(\Phi) = c\delta(\Phi^\dagger\Phi - I_M)$, where c is an appropriate normalization constant, direct integration of the above expression using the Dirac delta representation of the density appears to be formidable. To compute this integral, it is useful to introduce the following integral formula for $p(\Phi)$

$$p(\Phi) = \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M\pi^{\frac{M}{2}(2T+M+1)}}\int d\Omega\exp\left[i\,\mathrm{tr}\Omega(\Phi^\dagger\Phi - I_M)\right], \tag{13}$$

where the integration is over the $M\times M$ Hermitian matrix variable $\Omega=\Omega^\dagger$. (For a proof see Appendix A.)

We may therefore write

$$
\begin{aligned}
\mathrm{E}e^{\mathrm{tr}A\Phi B\Phi^\dagger} &= e^{\mathrm{tr}aB}\frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M\pi^{\frac{M}{2}(2T+M+1)}}\int d\Phi e^{-\mathrm{tr}(aI_T-A)\Phi B\Phi^\dagger}\int d\Omega e^{i\mathrm{tr}\Omega(\Phi^\dagger\Phi-I_M)} \\
&= e^{\mathrm{tr}aB}\frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M\pi^{\frac{M}{2}(2T+M+1)}}\int d\Omega e^{-i\mathrm{tr}\Omega}\int d\Phi e^{-\mathrm{tr}\left[(aI_T-A)\Phi B\Phi^\dagger-\Phi i\Omega\Phi^\dagger\right]}. \quad (14)
\end{aligned}
$$

Note that the inner integration over Φ is feasible, since Φ is now an unconstrained matrix variable. To facilitate this integration, the expression tr $\left[(aI_T-A)\Phi B\Phi^\dagger-\Phi i\Omega\Phi^\dagger\right]$ can be reorganized as a quadratic form in the entries of Φ,

$$
\begin{aligned}
\mathrm{tr}\left[(aI_T-A)\Phi B\Phi^\dagger-\Phi i\Omega\Phi^\dagger\right] &= \sum_{t=1}^{T}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\phi_{tm_1}\phi^*_{tm_2}\left[(a-a_t)b_{m_1}\delta_{m_1m_2}-i\omega_{m_1m_2}\right] \\
&= \sum_{t=1}^{T}\begin{bmatrix}\phi_{t1} & \cdots & \phi_{tM}\end{bmatrix}\left[(a-a_t)B-i\Omega\right]\begin{bmatrix}\phi^*_{t1}\\ \vdots \\ \phi^*_{tM}\end{bmatrix}, \quad (15)
\end{aligned}
$$

where $\{a_t\}$ and $\{b_m\}$ denote the diagonal elements of A and B respectively. Denoting the t-th row of Φ by ϕ_t, then the integral over Φ in (14) is the product of T elementary integrals,

$$
e^{-\mathrm{tr}\left[(aI_T-A)\Phi B\Phi^\dagger-\Phi i\Omega\Phi^\dagger\right]}=\prod_{t=1}^{T}\int d\phi_t e^{-\phi_t[(a-a_t)B-i\Omega]\phi_t^*}.
$$

Note that the parameter a has been chosen so that $a-a_t>0$ and hence $(a-a_t)B>0$. This implies that each of the integrals converges absolutely for all Hermitian Ω and so the result is given by

$$
\int d\Phi e^{-\mathrm{tr}\left[(aI_T-A)\Phi B\Phi^\dagger-\Phi i\Omega\Phi^\dagger\right]}=\frac{\pi^{TM}}{\prod_{t=1}^{T}\det\left[(a-a_t)B-i\Omega\right]}. \quad (16)
$$

(See App. D for an explanation.)

We substitute (16) into (14) to obtain

$$Ee^{\mathrm{tr}A\Phi B\Phi^\dagger} = e^{\mathrm{tr}aB}\frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M\pi^{\frac{M}{2}(M+1)}}\int\frac{d\Omega e^{-i\mathrm{tr}\Omega}}{\det\left[B(a-a_1)-i\Omega\right]\ldots\det\left[B(a-a_T)-i\Omega\right]}. \quad (17)$$

Recall in our case, that $A=\mathrm{diag}\{\sigma_1,\ldots,\sigma_K,0,\ldots,0\}$, where the σ_k are the nonzero eigenvalues of $XX^\dagger$ and $K=\min(T,N)$, and that $B=\left(I_M+\frac{1}{\rho\beta}D^{-2}\right)^{-1}$. Therefore

$$\begin{aligned}\mathrm{E}_{|D}e^{\mathrm{tr}X^\dagger\Phi\left(I_M+\frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger X} &= \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M\pi^{\frac{M}{2}(M+1)}}\\ &\times\int\frac{d\Omega e^{\mathrm{tr}[Ba-i\Omega]}}{\det\left[B(a-\sigma_1)-i\Omega\right]\ldots\det\left[B(a-\sigma_K)-i\Omega\right]\det\left[Ba-i\Omega\right]^{T-K}}. \quad (18)\end{aligned}$$

For the remainder of the paper, rather than consider an arbitrary diagonal B, we confine ourselves to the case of unitary space-time modulation.

3.1 The case of isotropically-distributed USTM

When $D=I_M$, the input signal takes on the form of the (scaled) isotropically-distributed unitary matrix $S=\sqrt{T}\Phi$. In this case, we have

$$B=\frac{1}{1+\frac{1}{\rho\beta}}I_M=\alpha I_M, \quad \alpha\triangleq\frac{\rho\beta}{1+\rho\beta} \quad (19)$$

and therefore

$$\begin{aligned}\mathrm{E}_{|D}e^{\mathrm{tr}X^\dagger\Phi\left(I_M+\frac{1}{\rho\beta}D^{-2}\right)^{-1}\Phi^\dagger X} = Ee^{\mathrm{tr}\alpha X^\dagger\Phi\Phi^\dagger X} &= \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M\pi^{\frac{M}{2}(M+1)}}\\ &\times\int\frac{d\Omega e^{\mathrm{tr}(\alpha aI_M-i\Omega)}}{\det\left[\alpha(a-\sigma_1)-i\Omega\right]\ldots\det\left[\alpha(a-\sigma_K)-i\Omega\right]\det\left[\alpha a-i\Omega\right]^{T-K}}. \quad (20)\end{aligned}$$

To carry out the integration in the above expression, it is useful to introduce the eigenvalue decomposition $\Omega = U\Lambda U^\dagger$ of the Hermitian matrix Ω. This eigenvalue decomposition can be regarded as a change of variables from Ω to $\{U, \Lambda\}$. Note that the integrand depends only on the eigenvalues. It is shown in Appendix C that, for any function $g(\Omega)$ whose value depends only on the eigenvalues of Ω (and not on the eigenvectors),

$$\int d\Omega g(\Omega) = \int d\Lambda \frac{\pi^{M(M-1)/2}}{\Gamma(M+1)\cdots\Gamma(1)} \prod_{k>l}\prod_{l}(\lambda_k - \lambda_l)^2 \cdot g(\Lambda). \tag{21}$$

The combination of (21) with (20) gives

$$\mathrm{E}e^{\mathrm{tr}\alpha X^\dagger \Phi\Phi^\dagger X} = \frac{\Gamma(T)\cdots\Gamma(T+1-M)}{\Gamma(M+1)\cdots\Gamma(1)(2\pi)^M}\int d\lambda_1 \cdots \int d\lambda_M$$
$$\prod_{m=1}^{M}\left(\frac{e^{\alpha a - i\lambda_m}}{[\alpha(a-\sigma_1)-i\lambda_m]\ldots[\alpha(a-\sigma_K)-i\lambda_m][\alpha a - i\lambda_m]^{T-K}}\right)\prod_{l<m}(\lambda_m - \lambda_l)^2.$$

Note that, since Ω is Hermitian, the λ_m, $m = 1, \ldots, M$ are real and so the domain of integration over the λ_m is $[-\infty, \infty]$. We now perform the change of variables $\lambda_m \to \lambda_m - i\alpha a$, and introduce the factor $-i$ in the terms $(\lambda_m - \lambda_l)$. This leads to

$$\begin{aligned}\mathrm{E}e^{\mathrm{tr}\alpha X^\dagger \Phi\Phi^\dagger X} \quad = \quad & \frac{\Gamma(T)\cdots\Gamma(T+1-M)\,(-1)^{M(M-1)/2}}{\Gamma(M+1)\cdots\Gamma(1)(2\pi)^M}\int d\lambda_1\cdots\int d\lambda_M \\ & \prod_{m=1}^{M}\left[\frac{e^{-i\lambda_m}}{(-\alpha\sigma_1 - i\lambda_m)\cdots(-\alpha\sigma_K - i\lambda_m)(-i\lambda_m)^{T-K}}\right]\prod_{l<m}(-i\lambda_m + i\lambda_l)^2.\end{aligned}$$

where we should now make a mental note of the fact that the domain of integration of the λ_m is no longer $[-\infty, \infty]$, but rather $[-\infty + i\alpha a, \infty + i\alpha a]$.

Now using the integral formula developed in Appendix B we obtain

$$\int \frac{d\lambda_1}{2\pi}\cdots\int\frac{d\lambda_M}{2\pi}\prod_{m=1}^{M}\left[\frac{e^{-i\lambda_m}}{(-\alpha\sigma_1 - i\lambda_m)\cdots(-\alpha\sigma_K - i\lambda_m)(-i\lambda_m)^{T-K}}\right]\prod_{l<m}(-i\lambda_m + i\lambda_l)^2 = M!\det F,$$

where the $M \times M$ Hankel matrix F is given by

$$F \triangleq \int_{-\infty+i\alpha a}^{\infty+i\alpha a} \frac{e^{-i\lambda} d\lambda/(2\pi)}{(-\alpha\sigma_1 - i\lambda)\dots(-\alpha\sigma_K - i\lambda)(-i\lambda)^{T-K}} \begin{bmatrix} (-i\lambda)^0 \\ \vdots \\ (-i\lambda)^{M-1} \end{bmatrix} \begin{bmatrix} (-i\lambda)^0 & \dots & (-i\lambda)^{M-1} \end{bmatrix}. \tag{22}$$

We thus have

$$Ee^{\mathrm{tr}\alpha X^\dagger \Phi\Phi^\dagger X} = (-1)^{M(M-1)/2} \cdot \frac{\Gamma(T)\dots\Gamma(T+1-M)}{\Gamma(M)\dots\Gamma(1)} \det F,$$

and therefore using (12)

$$p(X) = \frac{(-1)^{M(M-1)/2}}{\pi^{TN}} \cdot \frac{\exp\left(-\mathrm{tr}X^\dagger X\right)}{(1+\rho\beta)^{MN}} \cdot \frac{\Gamma(T)\dots\Gamma(T+1-M)}{\Gamma(M)\dots\Gamma(1)} \det F, \tag{23}$$

which is the desired expression for $p(X)$.

3.2 Computing the entries of F

All that remains to give an explicit formula for $p(X)$ is to determine the entries of the Hankel matrix F. Note from (22) that the (m, n)-th entry of F is given by

$$F_{mn} = \frac{1}{2\pi} \int_{-\infty+i\alpha a}^{\infty+i\alpha a} d\lambda \frac{(-i\lambda)^{-T+K+m+n-2} e^{-i\lambda}}{(-\alpha\sigma_1 - i\lambda)\dots(-\alpha\sigma_K - i\lambda)}. \tag{24}$$

This expression may be evaluated by inspection. We introduce a function of a real variable τ,

$$f(\tau) = \frac{1}{2\pi} \int_{-\infty+i\alpha a}^{\infty+i\alpha a} d\lambda \frac{e^{-i\lambda\tau}}{(-\alpha\sigma_1 - i\lambda)\dots(-\alpha\sigma_K - i\lambda)}, \tag{25}$$

which constitutes an inverse Fourier transform. Since the region of convergence includes $\mathrm{Im}\{\lambda\} \geq \alpha a$, $f(\tau)$ is causal. We recognize that the elements of F are merely derivatives or integrals of various orders of the function $f(\tau)$, evaluated at $\tau = 1$,

$$F_{mn} = f^{(-T+K+m+n-2)}(\tau)|_{\tau=1}. \tag{26}$$

The formula for $f(\tau)$ is obtained by taking $-i2\pi$ times the sum of the residues of the K distinct poles in (25), $\lambda = i\alpha\sigma_k,\ k = 1, \cdots, K$; the factor -1 results from the clockwise encirclements of the poles,

$$f(\tau) = \begin{cases} \sum_{k=1}^{K} \frac{e^{\alpha\sigma_k\tau}}{\prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)}, & \tau \geq 0 \\ 0, & \tau < 0 \end{cases}. \tag{27}$$

The r-th derivative of $f(\tau)$ at $\tau = 1$ is

$$\frac{d^r f(\tau)}{d\tau^r}|_{\tau=1} = \sum_{k=1}^{K} \frac{(\alpha\sigma_k)^r e^{\alpha\sigma_k}}{\prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)}. \tag{28}$$

The r-th integral of $f(\tau)$ at $\tau = 1$ is

$$\begin{aligned} \left(\int_0^\tau du\right)^r f(u) &= \sum_{k=1}^{K} \frac{1}{(\alpha\sigma_k)^r \prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \left(e^{\alpha\sigma_k} - \sum_{q=0}^{r-1} \frac{(\alpha\sigma_k)^q}{q!}\right) \\ &= \sum_{k=1}^{K} \frac{e^{\alpha\sigma_k}}{(\alpha\sigma_k)^r \prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \frac{\gamma(r, \alpha\sigma_k)}{\Gamma(r)}, \end{aligned} \tag{29}$$

where $\gamma(r, s)$ is the incomplete Gamma function,

$$\gamma(r, s) = \int_0^s du\, u^{r-1} e^{-u}.$$

Finally, we combine (26) with (28) and (29) to obtain the desired formula for the entries of F.

$$F_{mn} = \sum_{k=1}^{K} \frac{e^{\alpha\sigma_k}}{(\alpha\sigma_k)^Q \prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \cdot \begin{cases} \frac{\gamma(Q,\alpha\sigma_k)}{\Gamma(Q)}, & Q \geq 1 \\ 1, & Q \leq 0 \end{cases} \tag{30}$$

where $m, n = 1, \ldots, M$ and

$$Q = T - K - m - n + 2.$$

We can also give an alternative expression for $p(X)$. To this end, note that

$$\frac{d^r f(\tau)}{d\tau^r}|_{\tau=1} = \sum_{k=1}^{K} \frac{(\alpha\sigma_k)^r e^{\alpha\sigma_k}}{\prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} = \sum_{k=1}^{K} \frac{1}{\prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \sum_{q=0}^{\infty} \frac{(\alpha\sigma_k)^{q+r}}{q!}.$$

If we adopt the convention that $q! = \infty$ for all $q < 0$, this last equation can be written as

$$\frac{d^r f(\tau)}{d\tau^r}|_{\tau=1} = \sum_{k=1}^{K} \frac{1}{\prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \sum_{q=0}^{\infty} \frac{(\alpha\sigma_k)^q}{(q-r)!} \tag{31}$$

On the other hand,

$$\begin{aligned} \left(\int_0^\tau du\right)^r f(u) &= \sum_{k=1}^{K} \frac{1}{(\alpha\sigma_k)^r \prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \left(e^{\alpha\sigma_k} - \sum_{q=0}^{r-1} \frac{(\alpha\sigma_k)^q}{q!} \right) \\ &= \sum_{k=1}^{K} \frac{1}{(\alpha\sigma_k)^r \prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \sum_{q=0}^{\infty} \frac{(\alpha\sigma_k)^{q+r}}{(q+r)!} \\ &= \sum_{k=1}^{K} \frac{1}{\prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \sum_{q=0}^{\infty} \frac{(\alpha\sigma_k)^q}{(q+r)!}. \end{aligned} \tag{32}$$

Combining (31) and (32) yields

$$F_{mn} = \sum_{k=1}^{K} \frac{1}{\prod_{l\neq k}(\alpha\sigma_k - \alpha\sigma_l)} \sum_{q=0}^{\infty} \frac{(\alpha\sigma_k)^q}{(q + T - (m+n+K) + 2)!}, \quad m, n = 1, \ldots, M \tag{33}$$

where we have adopted the convention $q! = \infty$, for all $q < 0$.

3.3 An Alternative Expression for $p(X)$

The expressions (23), (30) and (33) for $p(X)$ can, in principle, be used to compute the mutual information $E \log \frac{p(X|S)}{p(X)}$ via Monte Carlo integration. However, practice has shown that Monte Carlo evaluation of the mutual information using these expressions leads to numerical problems at high SNR and at large numbers of receive and transmit antennas. The main reason for this is that the entries of the Hankel matrix F (in (30) and (33)) contain exponential terms with positive exponent (the $e^{\alpha\sigma_k}$ terms) which can lead to numerical instability, overflow, etc. Therefore it would be useful to avoid such expressions.

This indeed can be done: one can derive an equivalent formulation that involves only negative exponents. To this end, let $\Phi_\perp$ denote the $T \times (T-M)$ orthogonal complement of Φ, i.e.,

$$I_T = \Phi\Phi^\dagger + \Phi_\perp\Phi_\perp^\dagger \quad \text{and} \quad \Phi_\perp^\dagger\Phi_\perp = I_{T-M}.$$

Then one can write

$$p(X) = \frac{e^{-\text{tr}X^\dagger X}}{\pi^{TN}(1+\rho\beta)^{MN}} E e^{\text{tr}\alpha XX^\dagger\Phi\Phi^\dagger} = \frac{e^{-\text{tr}(1-\alpha)X^\dagger X}}{\pi^{TN}(1+\rho\beta)^{MN}} E e^{-\text{tr}\alpha XX^\dagger\Phi_\perp\Phi_\perp^\dagger}.$$

Note that in the above expression $\Phi_\perp$ is also isotropically random. We may therefore repeat the above steps to compute the expectation on the right-hand-side. We need only note two differences: 1) α has changed to $-\alpha$ in the exponent of the isotropic term, and 2) the i.r. unitary matrix has $T-M$, rather than M, columns, which means that we only need to replace M by $T-M$ in all our expressions. This leads to

$$p(X) = \frac{(-1)^{(T-M)(T-M-1)/2}}{\pi^{TN}} \cdot \frac{\exp\left(-\text{tr}(1-\alpha)X^\dagger X\right)}{(1+\rho\beta)^{MN}} \cdot \frac{\Gamma(T)\ldots\Gamma(M+1)}{\Gamma(T-M)\ldots\Gamma(1)} \det G, \tag{34}$$

where G is a $(T-M)\times(T-M)$ Hankel matrix whose entries are given by

$$G_{mn}=\sum_{k=1}^{K}\frac{e^{-\alpha\sigma_k}}{(-\alpha\sigma_k)^Q\prod_{l\neq k}(-\alpha\sigma_k+\alpha\sigma_l)}\cdot\begin{cases}\frac{\gamma(Q,-\alpha\sigma_k)}{\Gamma(Q)}, & Q\geq 1\\ 1, & Q\leq 0\end{cases} \tag{35}$$

where we define the incomplete Gamma function for a negative-valued argument as

$$\frac{\gamma\left(Q,-\alpha\sigma_k\right)}{\Gamma(Q)}=1-e^{\alpha\sigma_k}\sum_{q=0}^{Q-1}\frac{(-\alpha\sigma_k)^q}{q!},$$

or, equivalently,

$$G_{mn}=\sum_{k=1}^{K}\frac{1}{\prod_{l\neq k}(-\alpha\sigma_k+\alpha\sigma_l)}\sum_{q=0}^{\infty}\frac{(-\alpha\sigma_k)^q}{(q+T-(m+n+K)+2)!}, \tag{36}$$

where $m,n=1,\ldots,T-M$.

4 The Mutual Information for Isotropically Distributed USTM

Having obtained various formulae for $p(X)$ we can now readily write the mutual information $I(X;\Phi)$, using (10) with $D=I_M$, and (23), as

$$\begin{aligned}I(X;\Phi) &= -TN\log(\pi e)-MN\log(1+\rho\beta)-\Big[-TN\log\pi-MN\log(1+\rho\beta)\\ &\quad+\log\frac{\Gamma(T)\cdots\Gamma(T+1-M)}{\Gamma(M)\cdots\Gamma(1)}-\mathrm{E}\left\{\mathrm{tr}X^\dagger X\right\}\log e+\mathrm{E}\left\{\log\left[(-1)^{M(M-1)/2}\det F\right]\right\}\Big]\\ &= MN\rho\beta\log e-\log\frac{\Gamma(T)\cdots\Gamma(T+1-M)}{\Gamma(M)\cdots\Gamma(1)}-\mathrm{E}\left\{\log\left[(-1)^{M(M-1)/2}\det F\right]\right\},\end{aligned} \tag{37}$$

where the expectations are taken with respect to X, conditioned on Φ_0 being transmitted. A similar expression follows from (34).

Theorem 1 (Mutual Information for Isotropically Distributed Unitary Input). *Consider the channel model*

$$X = \sqrt{\frac{\rho}{M}} SH + W,$$

where the matrix $X \in \mathcal{C}^{T\times N}$ *is the received signal,* $S \in \mathcal{C}^{T\times M}$ *is the transmitted signal, and* $W \in \mathcal{C}^{T\times N}$ *is the additive noise. The entries of* H *and* W *are assumed to be independent* $\mathcal{CN}(0,1)$ *distributed. Then if* $S = \sqrt{T}\Phi$*, where* Φ *is an isotropically-distributed unitary matrix, we have*

$$I(X;\Phi) = \rho T N \log e - \log \frac{\Gamma(T)\cdots\Gamma(T+1-M)}{\Gamma(M)\cdots\Gamma(1)} - \mathrm{E}\left\{\log\left((-1)^{M(M-1)/2}\det F\right)\right\}, \tag{38}$$

or, equivalently,

$$I(X;\Phi) = \frac{1-\beta}{1+\rho\beta}\rho T N \log e - \log \frac{\Gamma(T)\cdots\Gamma(M+1)}{\Gamma(T-M)\cdots\Gamma(1)} - \mathrm{E}\left\{\log\left((-1)^{(T-M)(T-M-1)/2}\det G\right)\right\}, \tag{39}$$

where F *and* G *are* $M\times M$ *and* $(T-M)\times(T-M)$ *Hankel matrices whose entries are given by (30) and (35), respectively.*

5 Numerical Results

The formulae (38) and (39) are expressions for mutual information for an isotropically random unitary space-time input. At present, however, we are unable to evaluate analytically the expectation with respect to X. This may be done by Monte-Carlo integration. We generate L independent realizations $\{X_1,\cdots,X_L\}$, where the X_l are distributed according to (11), with $D = I_M$ and $S = \sqrt{T}\Phi_0$, $\Phi_0^\dagger = \begin{bmatrix} I_M & 0 \end{bmatrix}$. As mentioned earlier, because the entries of G have exponential terms with negative exponent, whereas those of F have exponential terms with positive exponent, Monte-Carlo simulation of (39) is much more numerically stable. The Monte-Carlo estimate of mutual information is

$$
\begin{aligned}
\hat{I}(X;S) &= \frac{1}{L}\sum_{l=1}^{L}\log\frac{p(X_l \mid \Phi_0)}{p(X_l)} \\
&= \frac{1}{L}\sum_{l=1}^{L}\log\left(\frac{\Gamma(T-M)\cdots\Gamma(1)}{\Gamma(T)\cdots\Gamma(M+1)}\frac{\exp\left(\alpha\,\mathrm{tr}\,X_l X_l^\dagger(-I_T+\Phi_0\Phi_0^\dagger)\right)}{(-1)^{(T-M)(T-M-1)/2}\det G_l}\right), \qquad (40)
\end{aligned}
$$

where we have used (11) for $p(X_l \mid \Phi_0)$, and (34) for $p(X_l)$. We avoid the form (39), where the expectation of $\log p(X \mid \Phi_0)$ is taken analytically, and instead, for greater stability, we apply Monte-Carlo integration to a log-likelihood ratio.

5.1 Mutual Information as a Function of T

Figure 1 shows the mutual information per symbol (i.e., (40) is normalized through a division by T) as a function of $T = 1, \cdots, 25$, for $M = N = 1, 2, 4$, and $\rho = -6.0,\ 6.0,\ 18.0$ dB, with $L = 4 \times 10^4$ Monte-Carlo trials. The $M = N = 1$ curves match similar curves obtained in [5], where the expectation with respect to X was performed by numerical quadrature. The dashed lines represent the perfect knowledge capacity. The figure shows the rate of convergence of the mutual information to the perfect knowledge capacity as $T \to \infty$. As can be seen, convergence is faster at higher SNR's.

5.2 Mutual Information as a Function of M and N

Figure 2 shows the mutual information per symbol as a function of $M = 1, \cdots, 8$, and $N = 1, \cdots, 8$, for $T = 8$, and $\rho = -6.0, 6.0, 18.0$ dB, and $L = 2500$. As can be seen, at very low SNR (here -6.0 db) unitary space-time modulation is unable to utilize multiple transmit antennas efficiently, and it is best to reserve all of the power for a single antenna. This contrasts sharply with the high SNR regime (here 18 db), where the mutual information peaks at about $M = \min(T/2, N)$ transmit antennas, thereby confirming the high SNR

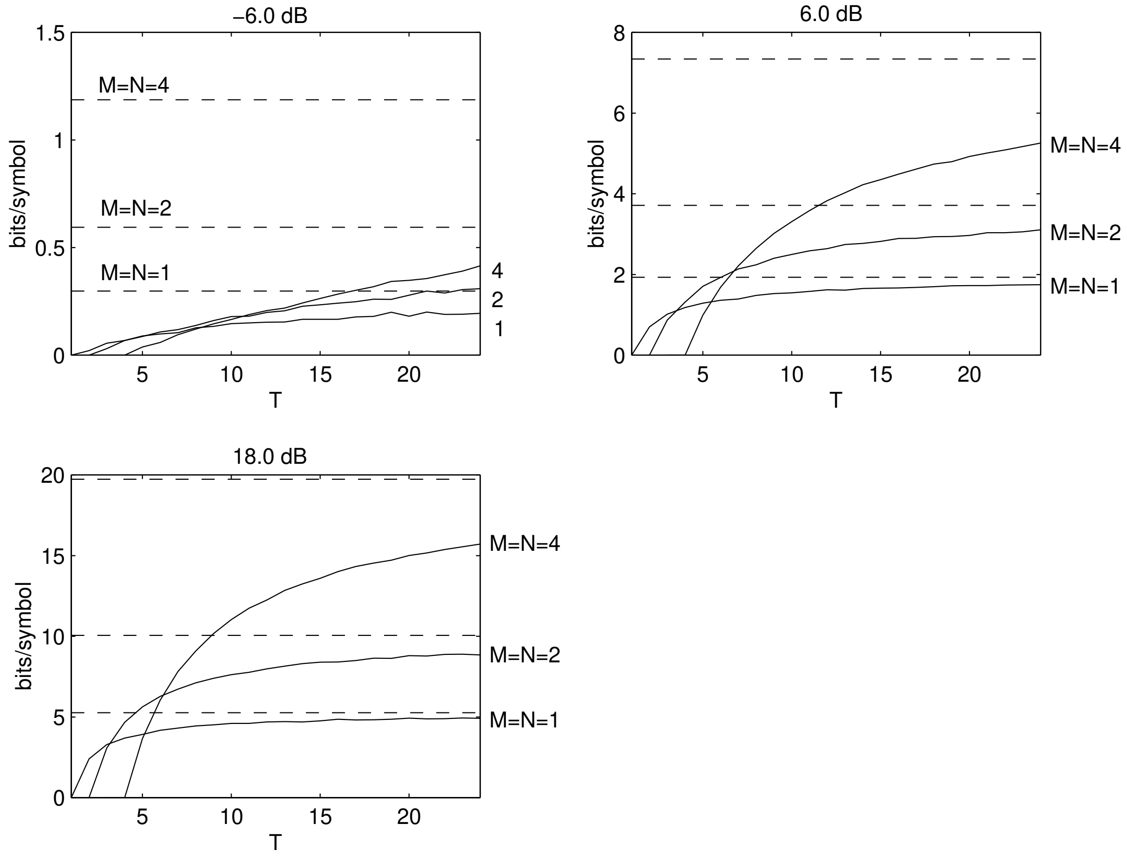

Figure 1: Mutual information per symbol versus duration of coherence interval T, for $M = N = 1, 2, 4$ transmit and receive antennas, and for $\rho = -6.0,\ 6.0,\ 18.0$ dB.

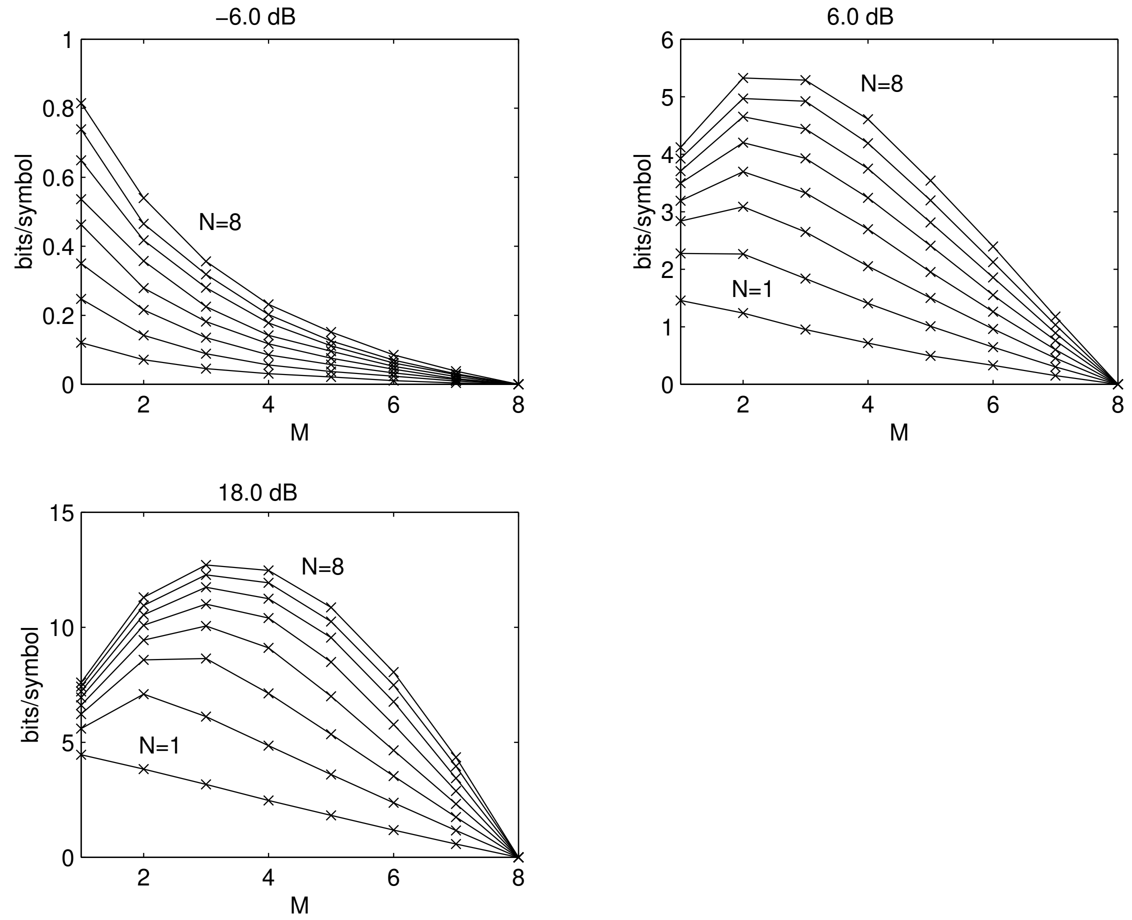

Figure 2: Mutual information per symbol versus number of transmit antennas M, for $N = 1, \cdots, 8$ receive antennas, for $T = 8$, and $\rho = -6.0, 6.0, 18.0$ dB. At this low SNR it is best to feed all power to a single antenna. At high SNR it is best to use $M = \min(T/2, N)$ transmit antennas.

results of [7]. Intermediate SNR values (here 6.0 db) present an intermediate stage. We conclude that, as the SNR increases, more and more transmit antennas should be powered successively.

6 Discussion and Conclusions

We have obtained a closed-form expression in the form of a determinant, for the probability density of the received signal, when the input is isotropically random unitary. This expression, when combined with a simple Monte-Carlo integration enables us to compute mutual information for all cases of interest. At sufficiently low SNR we have found that mutual information is maximized by allocating all power to a single transmit antenna. At sufficiently high SNR we have confirmed the result of [7] that $M = \min(T/2, N)$ transmit antennas maximizes mutual information.

In future work we will attempt to compute the mutual information for arbitrary D, rather than $D = I_M$. This should lead to further insight into the problem and to tighter lower bounds on the actual channel capacity.

Our method of taking expectations with respect to the isotropically random unitary density has other interesting ramifications. In particular, it enables us to compute any desired moments for an isotropically random unitary matrix. Also we can obtain a variety of new bounds on the random coding exponent for space-time autocoding [2], which should provide improved estimates, compared with the simple union bound, of the autocoding performance of codebooks of isotropically random unitary signals [9, 10].

[2]This was pointed out independently to us by Prof. S. Shamai.

A Integral Formula for $p(\Phi)$

Let us begin by obtaining the formula for the density of a $T \times M$ i.d. unitary matrix $\Phi = \begin{bmatrix} \phi_1 & \ldots & \phi_M \end{bmatrix}$, where the $\phi_m, m = 1, \ldots M$ denote the $T \times 1$ column vectors of Φ. This density was derived in [5], but we include it here to make our treatment self-contained and because our derivation will be useful for obtaining the desired Fourier integral expression.

We first note that, due to the unitary constraint $\Phi^\dagger \Phi = I_M$, the density must have the form $p(\Phi) = f(\Phi)\delta(\Phi^\dagger\Phi - I_M)$, where $f(\Phi)$ is some smooth function and where

$$\delta(\Phi^\dagger\Phi - I_M) = \prod_{m=1}^{M} \delta(\phi_m^\dagger \phi_m - 1) \prod_{m>n} \delta\left(\mathrm{Re}(\phi_m^\dagger \phi_n)\right) \delta\left(\mathrm{Im}(\phi_m^\dagger \phi_n)\right). \tag{A.1}$$

Now for Φ to be isotropically-distributed, the density $p(\Phi)$ must be invariant to pre- and post-multiplication by any $T \times T$ and $M \times M$ unitary matrices, say U and V. Clearly, $\delta(\Phi^\dagger\Phi - I_M)$ is invariant under such a transformation. For $f(\cdot)$ to be invariant, we require that $f(U\Phi V) = f(\Phi)$. But since U and V are arbitrary this implies that $f(\Phi)$ cannot be a function of the singular vectors of Φ and must therefore be only a function of the singular values. But since Φ is unitary, its singular values are all unity, which implies that $f(\Phi)$ must be a constant. In other words, we have $p(\Phi) = c\delta(\Phi^\dagger\Phi - I_M)$, for some normalizing constant c.

To compute the normalizing constant c, we can rewrite the density as

$$\begin{aligned} p(\Phi) = p(\phi_1, \ldots, \phi_M) &= p(\phi_1)p(\phi_2, \ldots, \phi_M|\phi_1) \\ &= p(\phi_1) \prod_{m=2}^{M} p(\phi_m|\phi_1, \ldots, \phi_{m-1}). \end{aligned} \tag{A.2}$$

Now since ϕ_1 is $T \times 1$ isotropically-distributed unitary, we have $p(\phi_1) = c\delta(\phi_1^\dagger\phi_1 - 1)$, for some constant c. To determine c it is useful to represent the delta function via its Fourier transform, $\delta(x) = \frac{1}{2\pi}\int d\omega e^{i\omega x}$, and to proceed as follows:

$$\begin{aligned}
1 = \int d\phi_1 p(\phi_1) &= \frac{c}{2\pi}\int d\phi_1 \int d\omega e^{i\omega(\phi_1^\dagger\phi_1 - 1)} \\
&= \frac{c}{2\pi}\int d\phi_1 e^{\alpha\phi_1^\dagger\phi_1} e^{-\alpha\phi_1^\dagger\phi_1}\int d\omega e^{i\omega(\phi_1^\dagger\phi_1 - 1)}, \quad \alpha > 0 \\
&= \frac{ce^{\alpha}}{2\pi}\int d\omega e^{-i\omega}\int d\phi_1 e^{-(\alpha - i\omega)\phi_1^\dagger\phi_1} \\
&= \frac{ce^{\alpha}}{2\pi}\int d\omega e^{-i\omega}\cdot\frac{\pi^T}{(\alpha - i\omega)^T} = \frac{c\pi^T}{\Gamma(T)},
\end{aligned}$$

which implies that

$$p(\phi_1) = \frac{\Gamma(T)}{\pi^T}\delta(\phi_1^\dagger\phi_1 - 1).$$

Now we also have

$$p(\phi_m|\phi_1,\ldots,\phi_{m-1}) = c\delta(\phi_m^\dagger\phi_m - 1)\prod_{n=1}^{m-1}\delta(\phi_m^\dagger\phi_n),$$

for some other normalizing constant c, and where $\delta(\phi_m^\dagger\phi_n) = \delta\left(\mathrm{Re}(\phi_m^\dagger\phi_n)\right)\delta\left(\mathrm{Im}(\phi_m^\dagger\phi_n)\right)$. Proceeding with a similar argument to the one presented above to determine c, or, alternatively, by simply noting that, in the subspace that is orthogonal to $\{\phi_1,\cdots,\phi_{M-1}\}$, $\phi_{m|\phi_1,\ldots,\phi_{m-1}}$ has the same distribution as a $T+1-M$-dimensional isotropically-distributed unitary vector, we may write

$$p(\phi_m|\phi_1,\ldots,\phi_{m-1}) = \frac{\Gamma(T+1-M)}{\pi^{T+1-m}}\delta(\phi_m^\dagger\phi_m - 1)\prod_{n=1}^{m-1}\delta(\phi_m^\dagger\phi_n).$$

Using the expressions for $p(\phi_1)$ and $p(\phi_m|\phi_1,\ldots,\phi_{m-1})$, and using (A.2), we may write

$$p(\Phi) = \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{\pi^{\frac{M(2T-M+1)}{2}}}\prod_{m=1}^{M}\delta(\phi_m^\dagger\phi_m - 1)\prod_{m>n}\delta(\phi_m^\dagger\phi_n) \tag{A.3}$$

$$= \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{\pi^{\frac{M(2T-M+1)}{2}}}\delta(\Phi^\dagger\Phi - I_M), \tag{A.4}$$

which is the desired density function for $p(\Phi)$.

A.1 The Integral Formula

To obtain an integral formula for $p(\Phi)$, we shall replace the delta functions in (A.3) by their Fourier integrals. Thus,

$$\delta(\phi_m^\dagger\phi_m - 1) = \frac{1}{2\pi}\int d\omega_{mm} e^{i\omega_{mm}(\phi_m^\dagger\phi_m - 1)}, \tag{A.5}$$

and

$$\begin{aligned}
\delta(\phi_m^\dagger\phi_n) &= \delta\left(\mathrm{Re}(\phi_m^\dagger\phi_n)\right)\cdot\delta\left(\mathrm{Im}(\phi_m^\dagger\phi_n)\right) \\
&= \frac{1}{(2\pi)^2}\int d\alpha_{mn} d\beta_{mn} e^{i\left(\alpha_{mn}\mathrm{Re}(\phi_m^\dagger\phi_n) + \beta_{mn}\mathrm{Im}(\phi_m^\dagger\phi_n)\right)} \\
&= \frac{1}{(2\pi)^2}\int d\alpha_{mn} d\beta_{mn} e^{i\left(\frac{\alpha_{mn}-i\beta_{mn}}{2}\phi_m^\dagger\phi_n + \frac{\alpha_{mn}+i\beta_{mn}}{2}\phi_n^\dagger\phi_m\right)}.
\end{aligned} \tag{A.6}$$

Now if we define the off-diagonal parts of the Hermitian matrix Ω via $\omega_{mn} = \frac{\alpha_{mn}-i\beta_{mn}}{2}$ and $\omega_{nm} = \omega_{mn}^* = \frac{\alpha_{mn}+i\beta_{mn}}{2}$, we clearly have $d\omega_{mn} = \frac{d\alpha_{mn}d\beta_{mn}}{4}$, so that

$$\delta(\phi_m^\dagger\phi_n) = \frac{1}{\pi^2}\int d\omega_{mn} e^{i\left(\omega_{mn}\phi_m^\dagger\phi_n + \omega_{nm}\phi_n^\dagger\phi_m\right)}. \tag{A.7}$$

Combining (A.5) and (A.7) with (A.3), we arrive at

$$p(\Phi) = \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{\pi^{\frac{M(2T-M+1)}{2}}}\cdot\frac{1}{(2\pi)^M}\int\prod_{m=1}^{M} d\omega_{mm} e^{i\omega_{mm}\left[(\Phi^\dagger\Phi)_{mm} - \delta_{mm}\right]}$$

$$
\begin{aligned}
& \cdot \frac{1}{\pi^{M(M-1)}} \int \prod_{m>n} d\omega_{mn} e^{i\left[\omega_{mn}((\Phi^\dagger\Phi)_{mn}-\delta_{nm})+\omega_{nm}((\Phi^\dagger\Phi)_{nm}-\delta_{mn})\right]} \\
= \quad & \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M \pi^{\frac{M}{2}(2T+M+1)}} \int d\Omega e^{i \mathrm{tr} \Omega^T (\Phi^\dagger\Phi - I_M)} \\
= \quad & \frac{\Gamma(T)\ldots\Gamma(T+1-M)}{2^M \pi^{\frac{M}{2}(2T+M+1)}} \int d\Omega e^{i \mathrm{tr} \Omega (\Phi^\dagger\Phi - I_M)}, \qquad \text{(A.8)}
\end{aligned}
$$

where the integration is over the Hermitian matrix Ω. But this simply the desired result (13).

B A Useful Integral Formula

In the next result we generalize a formula of Wigner [11] for positive functions $f(\cdot)$ to arbitrary ones.

Lemma 1 (An Integral Formula). *Consider the single-variable function $f(\cdot)$. Then we have the following identity for the M-fold integral*

$$
\int \left(\prod_{m=1}^{M} d\lambda_m f(\lambda_m) \right) \prod_{l<m} (-i\lambda_m + i\lambda_l)^2 = M! \det F, \qquad \text{(B.1)}
$$

where F is an $M \times M$ Hankel matrix given by

$$
F \triangleq \int d\lambda f(\lambda) \begin{bmatrix} (-i\lambda)^0 \\ \vdots \\ (-i\lambda)^{M-1} \end{bmatrix} \begin{bmatrix} (-i\lambda)^0 & \ldots & (-i\lambda)^{M-1} \end{bmatrix}. \qquad \text{(B.2)}
$$

Proof: Note that the product $\prod_{l<m}(-i\lambda_m + i\lambda_l)$ can be written as the determinant of a Vandermonde

matrix

$$\prod_{l<m}(-i\lambda_m+i\lambda_l)=\det\begin{bmatrix}(-i\lambda_1)^0 & \dots & (-i\lambda_M)^0\\ \vdots & \ddots & \vdots\\ (-i\lambda_1)^{M-1} & \dots & (-i\lambda_M)^{M-1}\end{bmatrix}.$$

Assume now that F is nonsingular,[3] and define the polynomials $g_1(\cdot),\dots,g_M(\cdot)$ via

$$\begin{bmatrix}g_1(\lambda)\\ \vdots\\ g_M(\lambda)\end{bmatrix}\triangleq F^{-1}\begin{bmatrix}(-i\lambda)^0\\ \vdots\\ (-i\lambda)^{M-1}\end{bmatrix} \tag{B.3}$$

With this definition of the $g_m(\cdot)$, it is important to note that

$$\begin{aligned}I_M=F^{-1}F &= \int d\lambda\, f(\lambda)\, F^{-1}\begin{bmatrix}(-i\lambda)^0\\ \vdots\\ (-i\lambda)^{M-1}\end{bmatrix}\begin{bmatrix}(-i\lambda)^0 & \dots & (-i\lambda)^{M-1}\end{bmatrix}\\ &= \int d\lambda f(\lambda)\begin{bmatrix}g_1(\lambda)\\ \vdots\\ g_M(\lambda)\end{bmatrix}\begin{bmatrix}(-i\lambda)^0 & \dots & (-i\lambda)^{M-1}\end{bmatrix},\end{aligned}$$

which implies that the $g_m(\cdot)$ and λ^{n-1} satisfy the following *bi-orthogonality* condition

$$\int d\lambda\, f(\lambda)\, g_m(\lambda)\,(-i\lambda)^{n-1}=\delta_{mn},\quad m,n=1,\dots M. \tag{B.4}$$

[3]We shall momentarily show that if F is singular then the integral is zero.

Denoting the integral in the statement of the lemma by J, we can now write

$$\begin{aligned}
J &= \int \left(\prod_{m=1}^{M} d\lambda_m f(\lambda_m) \right) \det F \det F^{-1} \det \begin{bmatrix} (-i\lambda_1)^0 & \ldots & (-i\lambda_M)^0 \\ \vdots & \ddots & \vdots \\ (-i\lambda_1)^{M-1} & \ldots & (-i\lambda_M)^{M-1} \end{bmatrix}^2 \\
&= \det F \int \left(\prod_{m=1}^{M} d\lambda_m f(\lambda_m) \right) \det \left[\begin{bmatrix} g_1(\lambda_1) & \ldots & g_1(\lambda_M) \\ \vdots & \ddots & \vdots \\ g_M(\lambda_1) & \ldots & g_M(\lambda_M) \end{bmatrix} \begin{bmatrix} (-i\lambda_1)^0 & \ldots & (-i\lambda_M)^0 \\ \vdots & \ddots & \vdots \\ (-i\lambda_1)^{M-1} & \ldots & (-i\lambda_M)^{M-1} \end{bmatrix} \right].
\end{aligned}$$

If we now expand the above two determinants as sums of the products $g_{m_1}(\lambda_1)g_{m_2}(\lambda_2)\ldots g_{m_M}(\lambda_M)$ and $(-i\lambda_1)^{n_1-1}(-i\lambda_2)^{n_2-1}\ldots(-i\lambda_M)^{n_M-1}$, where the $\{m_i\}$ and $\{n_i\}$ denote some permutation of the integers $\{1,\ldots,M\}$, we note that, due to the biorthogonality condition, whenever the $\{m_i\}$ and $\{n_i\}$ are different permutations their product integrates to zero. When they are the same permutation they integrate to unity. Since there are $M!$ such permutations, we conclude that

$$J = M! \det F,$$

which is the desired result.

All that remains to be shown is that when F is singular the integral is zero. Suppose that F has $R > 0$ zero eigenvalues, so that its singular value decomposition takes on the form

$$F = U \begin{bmatrix} \Sigma & \\ & 0_{R\times R} \end{bmatrix} V^\dagger.$$

Now defining the polynomials $q_1(\cdot), \ldots, q_M(\cdot)$ and $r_1(\cdot), \ldots, r_M(\cdot)$

$$\begin{bmatrix} q_1(\lambda) \\ \vdots \\ q_M(\lambda) \end{bmatrix} \triangleq U^\dagger \begin{bmatrix} (-i\lambda)^0 \\ \vdots \\ (-i\lambda)^{M-1} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} r_1(\lambda) \\ \vdots \\ r_M(\lambda) \end{bmatrix} \triangleq V^T \begin{bmatrix} (-i\lambda)^0 \\ \vdots \\ (-i\lambda)^{M-1} \end{bmatrix}$$

it is straightforward to see that

$$\int d\lambda f(\lambda) q_m(\lambda) r_n(\lambda) = \begin{cases} \sigma_n \delta_{mn} & \text{if } m, n \le M - R \\ 0 & \text{if } m > M - R \text{ or } n > M - R \end{cases} \tag{B.5}$$

Using an argument similar to the one presented above, one can write

$$J = \int \left(\prod_{m=1}^{M} d\lambda_m f(\lambda_m) \right) \det \left[\begin{bmatrix} q_1(\lambda_1) & \ldots & q_1(\lambda_M) \\ \vdots & \ddots & \vdots \\ q_M(\lambda_1) & \ldots & q_M(\lambda_M) \end{bmatrix} \right] \left[\begin{bmatrix} r_1(\lambda_1) & \ldots & r_1(\lambda_M) \\ \vdots & \ddots & \vdots \\ r_M(\lambda_1) & \ldots & r_M(\lambda_M) \end{bmatrix} \right].$$

Now expanding the above two determinants and using (B.5) it follows that all terms integrate to zero, so that $J = 0$.

■

C Jacobian of the Eigenvalue Decomposition

The expression (20) constitutes an integral with respect to a Hermitian matrix of a function that depends only on the eigenvalues of the matrix. Our object is to reduce this expression to an integral with respect to the eigenvalues alone. This is a well known result in random matrix theory (see, e.g., [12, 13]), but we shall include it here for completeness.

We begin by reviewing some unorthodox integrations of coupled Dirac delta functions whose arguments are nonlinear functions [10]. Let $f(x)$ be a K-component vector-valued nonlinear function of a K-dimensional real-valued vector x, and consider the integral

$$\int dx\, h(x)\delta\left(f(x)\right) = \int dx\, h(x) \prod_{k=1}^{K} \delta\left(f_k(x)\right), \tag{C.1}$$

where $h(x)$ is scalar-valued. Furthermore suppose that the function has only a single zero $f(x_0) = 0$. Clearly, the support for the integral occurs at $x = x_0$. A change of coordinates $y = f(x)$ gives

$$\begin{aligned} \int dx\, h(x) \prod_{k=1}^{K} \delta\left(f_k(x)\right) &= \int dy \left[\frac{h(x)}{\left|\det\left\{\nabla f^T(x)\right\}\right|}\right]_{x=f^{-1}(y)} \cdot \prod_{k=1}^{K} \delta\left(y_k\right) \\ &= \frac{h\left(x_0\right)}{\left|\det\left\{\nabla f^T(x)\right\}\right|_{x=x_0}}, \end{aligned}$$

where $[\nabla f^T]_{k\ell} = \partial f_\ell / \partial x_k$. Formally, this means that

$$\delta\left(f(x)\right) = \frac{\delta\left(x - x_0\right)}{\left|\det\left\{\nabla f^T(x)\right\}\right|_{x=x_0}}. \tag{C.2}$$

Note that the integral (C.1) depends only on first derivatives of $f(x)$ at x_0. In practice, therefore, one can always linearize the arguments of the delta functions as follows,

$$\begin{aligned} \int dx\, h(x)\delta\left(f(x)\right) &= \int dx\, h(x)\delta\left(f(x_0) + \left[\nabla f^T(x)\right]_{x=x_0}^T \cdot (x - x_0)\right) \\ &= \int \frac{dx}{\left|\det\left\{\nabla f^T(x)\right\}\right|_{x=x_0}} h(x)\delta(x - x_0) \\ &= \frac{h\left(x_0\right)}{\left|\det\left\{\nabla f^T(x)\right\}\right|_{x=x_0}}. \end{aligned} \tag{C.3}$$

The above integration technique accommodates multiple zeros by taking a sum of delta functions at the individual roots, and dividing by the appropriate Jacobian determinants. It is modified for complex variables by squaring the determinants.

Let $g(\Omega)$ be a function of a $M \times M$ Hermitian matrix Ω. Denote by $p_0(\Psi)$ the $M \times M$ isotropically random unitary density, i.e.,

$$p_0(\Psi) = c_0 \delta(\Psi^\dagger \Psi), \quad c_0 = \frac{\Gamma(M) \cdots \Gamma(1)}{\pi^{M(M+1)/2}}. \tag{C.4}$$

Let v be a M-dimensional real vector, and consider the following integral, with respect to arbitrary complex Ψ, Hermitian Ω, and *ordered* v, i.e., $v_1 \geq v_2 \geq \cdots \geq v_M$,

$$\int dv \int d\Psi\, p_0(\Psi) g\left(\Psi D_v \Psi^\dagger\right) = \int d\Omega\, g(\Omega) \int dv \int d\Psi\, \delta\left(\Omega - \Psi D_v \Psi^\dagger\right) \cdot p_0(\Psi), \tag{C.5}$$

where D_v denotes the diagonal matrix whose diagonal elements comprise v, and where the Dirac delta function involves only the on- and below-diagonal elements of Ω. We represent Ω by its eigenvectors and eigenvalues, $\Omega = \Phi D_\lambda \Phi^\dagger$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$,

$$\begin{aligned} \text{(C.5)} &= \int d\Omega\, g(\Omega) \int dv \int d\Psi\, \delta\left(\Phi D_\lambda \Phi^\dagger - \Psi D_v \Psi^\dagger\right) \cdot p_0(\Psi) \\ &= \int d\Omega\, g(\Omega) \int dv \int d\Psi\, \delta\left(D_\lambda - \Phi^\dagger \Psi D_v \Psi^\dagger \Phi\right) \cdot p_0(\Psi) \\ &= \int d\Omega\, g(\Omega) \int dv \int d\Psi\, \delta\left(D_\lambda - \Psi D_v \Psi^\dagger\right) \cdot p_0(\Psi), \end{aligned} \tag{C.6}$$

where the first step uses the property that, for a unitary $M \times M$ Φ, and a Hermitian Y, $\delta(Y) = \delta(\Phi Y) = \delta(Y \Phi^\dagger)$, and where the second step involves the coordinate transformation $\Phi^\dagger \Psi \rightarrow \Psi$ combined with the fact that $p_0(\Psi)$ is the isotropically random unitary density. Together, the Dirac delta function in (C.6), and

$p_0(\cdot)$, comprise coupled delta functions whose solution is of the form,

$$v = \lambda, \quad \Psi = D_z, \quad |z_1| = |z_2| = \cdots = |z_M| = 1. \tag{C.7}$$

Therefore we let $\Psi = D_z + \tilde{\Psi}$, where the diagonal elements of $\tilde{\Psi}$ are equal to zero. We then have, recalling that higher-order terms in the arguments of the delta functions may be discarded,

$$\begin{aligned} p_0(\Psi) &= p_0\left(D_z + \tilde{\Psi}\right) \\ &= c_0 \prod_{n=1}^{M} \delta\left(|z_n|^2 - 1\right) \prod_{m>n} \delta\left(z_m^* \tilde{\psi}_{mn} + z_n \tilde{\psi}_{nm}^*\right), \end{aligned} \tag{C.8}$$

$$\begin{aligned} \delta\left(D_\lambda - \Psi D_v \Psi^\dagger\right) &= \delta\left(D_\lambda - D_v - \tilde{\Psi} D_v D_z^\dagger - D_z D_v \tilde{\Psi}^\dagger\right) \\ &= \prod_{n=1}^{M} \delta\left(v_n - \lambda_n\right) \prod_{m>n} \delta\left(z_n^* v_n \tilde{\psi}_{mn} + z_m v_m \tilde{\psi}_{nm}^*\right). \end{aligned} \tag{C.9}$$

The substitution of (C.8) and (C.9) into (C.6) yields

$$\begin{aligned} \text{(C.6)} \quad = \quad & \int d\Omega\, g(\Omega) \int dv \int dz \int d\tilde{\Psi}\, c_0 \cdot \prod_{n=1}^{M} \delta\left(|z_n|^2 - 1\right) \delta\left(v_n - \lambda_n\right) \\ \cdot \quad & \prod_{m>n} \delta\left(z_m^* \tilde{\psi}_{mn} + z_n \tilde{\psi}_{nm}^*\right) \delta\left(z_n^* v_n \tilde{\psi}_{mn} + z_m v_m \tilde{\psi}_{nm}^*\right). \end{aligned} \tag{C.10}$$

The integral with respect to v is trivial; to integrate with respect to $\tilde{\psi}_{mn}$ and $\tilde{\psi}_{nm}$, we need the determinant for the coupled system of equations,

$$\begin{bmatrix} z_m^* & z_n \\ z_n^* v_n & z_m v_m \end{bmatrix} \begin{bmatrix} \tilde{\psi}_{mn} \\ \tilde{\psi}_{nm}^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{C.11}$$

which is equal to $v_m|z_m|^2 - v_n|z_n|^2 = v_m - v_n$. We then set $\tilde{\psi}_{mn}$ and $\tilde{\psi}_{nm}$ equal to zero, and divide by $(v_m - v_n)^2$ to give

$$\begin{aligned}
\text{(C.10)} &= \int d\Omega\, g(\Omega) \int dv \delta(v-\lambda) \int dz\, c_0 \cdot \prod_{n=1}^{M} \delta\left(|z_n|^2 - 1\right) \prod_{m>n} \frac{1}{(v_m - v_n)^2} \\
&= \int d\Omega\, g(\Omega)\, c_0 \pi^M \prod_{n=1}^{M} \prod_{m>n} \frac{1}{(\lambda_m - \lambda_n)^2}.
\end{aligned} \tag{C.12}$$

The combination of (C.5), (C.6), (C.10), and (C.12) yields

$$\frac{1}{M!} \int dv_{\text{unordered}} \int d\Psi\, p_0(\Psi) g\left(\Psi D_v \Psi^\dagger\right) = \int d\Omega\, g(\Omega)\, c_0 \pi^M \prod_{n=1}^{M} \prod_{m>n} \frac{1}{(\lambda_m - \lambda_n)^2}, \tag{C.13}$$

where we divide by $M!$ since the components of v are now unordered. An inspection of this expression yields the Jacobian for the coordinate transformation from Ω to its unordered eigenvalues λ and its eigenvectors Φ,

$$d\Omega = \frac{\pi^{M(M-1)/2}}{\Gamma(M+1)\cdots\Gamma(1)} p_0(\Phi)\, d\Phi \prod_{n=1}^{M} \prod_{m>n} (\lambda_m - \lambda_n)^2\, d\lambda. \tag{C.14}$$

D An Exponential Quadratic Integral

Let P and Q be $M \times M$ Hermitian matrices such that $P > 0$. In this appendix we verify the integral formula

$$\int dy e^{-y^*(P+iQ)y} = \frac{\pi^M}{\det(P+iQ)}. \tag{D.1}$$

Note that this result only requires a positivity assumption on P (and not on Q).

Since $P > 0$, for *any* square-root factorization $P = P^{1/2} P^{*/2}$, the matrix $P^{1/2}$ will be nonsingular.

Consider the change of variables $z = P^{*/2}y$. Then clearly,

$$dz = dy \left|\det P^{*/2}\right|^2 = dy \det P,$$

and the integral becomes

$$\int dy e^{-y^*(P+iQ)y} = \frac{1}{\det P} \int dz e^{-z^*(I_M + iP^{-1/2}QP^{-*/2})}.$$

Consider now the eigenvalue decomposition of the Hermitian matrix $P^{-1/2}QP^{-*/2} = U\Lambda U^*$, where U is unitary and Λ is diagonal with real diagonal entries $\lambda_1, \ldots, \lambda_M$. The change of variable $w = U^*z$ yields $dw = dz$ (since the transformation matrix is unitary) and so

$$\int dy e^{-y^*(P+iQ)y} = \frac{1}{\det P} \int dw\, e^{-w^*(I_M + i\Lambda)w} = \frac{1}{\det P} \prod_{m=1}^{M} \int dw_m e^{-(1+i\lambda_m)|w_m|^2}.$$

It is straightforward to show that $\int dw_m e^{-(1+i\lambda_m)|w_m|^2} = \frac{\pi}{1+i\lambda_m}$ so that

$$\begin{aligned}
\int dy e^{-y^*(P+iQ)y} = \frac{\pi^M}{\det P} \prod_{m=1}^{M} \frac{1}{1+i\lambda_m} &= \frac{\pi^M}{\det P} \cdot \frac{1}{\det(I_M + i\Lambda)} \\
&= \frac{\pi^M}{\det P} \cdot \frac{1}{\det(I_M + iU\Lambda U^*)} \\
&= \frac{\pi^M}{\det(P + iQ)},
\end{aligned}$$

which is the desired result.

References

[1] J. Salz and A. Wyner, "On data transmission over cross coupled linear channels," tech. rep., AT&T Bell Laboratories, 1985.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585–595, Nov. 1999.

[3] 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.

[4] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time Rayleigh fading channels," in *IEEE Int. Symp. Info. Theory*, p. 473, June 1997. Also submitted to *IEEE Trans. Info. Theory*.

[5] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Info. Theory*, vol. 45, pp. 139–157, 1999.

[6] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communication in Rayleigh flat-fading," *IEEE Trans. Info. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[7] L. Zheng and D. Tse, "Packing spheres in the Grassman manifold: a geometric approach to the non-coherent multi-antenna channel," *submitted to IEEE Trans. Info. Theory*, 2000.

[8] B. Hochwald, T. Marzetta, T. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Info. Theory*, vol. 46, no. 6, pp. 1962–1974, 2000.

[9] B. Hassibi, B. Hochwald, and T. Marzetta, "Space-time autocoding," *submitted to IEEE Trans. Info. Theory*, 1999.

[10] T. L. Marzetta, B. Hassibi, and B. M. Hochwald, "Structured unitary space-time autocoding constellations," *submitted to IEEE Trans. Info. Theory*, 2000.

[11] E. Wigner, "Random matrices in physics," *SIAM Review*, vol. 9, pp. 1–23, 1967.

[12] M. Mehta, *Random Matrices and Statistical Theory*. New York: Academic Press, 1967.

[13] R. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: John Wiley and Sons, 1982.