# Structured Unitary Space-Time Autocoding Constellations

THOMAS L. MARZETTA BABAK HASSIBI BERTRAND M. HOCHWALD

Mathematical Sciences Research Center
Bell Laboratories, Lucent Technologies
600 Mountain Avenue
Murray Hill, NJ 07974
USA
{tlm,hassibi,hochwald}@research.bell-labs.com

**Abstract**

We recently showed that arbitrarily reliable communication is possible within a single coherence interval in Rayleigh flat fading as the symbol-duration of the coherence interval and the number of transmit antennas grow simultaneously. This effect, where the space-time signals act as their own channel codes, is called *autocoding*. For relatively short (e.g., 16-symbol) coherence intervals, a codebook of isotropically random unitary space-time signals theoretically supports transmission rates that are a significant fraction of autocapacity with an extremely low probability of error. However a constellation of the required size (typically $L = 2^{80}$) is impossible to generate and store, and due to lack of structure there is little hope of finding a fast decoding scheme. In this paper we propose a random, but highly structured, constellation that is completely specified by $\log_2 L$ independent isotropically distributed unitary matrices. The distinguishing property of this construction is that any two signals in the constellation are pairwise statistically independent and isotropically distributed. Thus, the pairwise probability of error, and hence the union bound on the block probability of error, of the structured constellation is identical to that of a fully random constellation of independent signals. As part of this work we have established a subsidiary result that is interesting in its own right: the square (or for that matter, any integer power greater than one) of an isotropically random unitary matrix is *not* isotropically random, with two exceptions: 1) a one-by-one complex unitary matrix, and 2) a two-by-two real orthogonal matrix.

*Index Terms*—wireless communication,multiple antennas, Rayleigh flat-fading, unitary space-time modulation, space-time autocoding, eigenvalues of random matrices

# 1 Introduction

Research in multiple antenna wireless entered an exciting phase with [3, 13], which predict spectacular capacities (both Shannon and outage) for single-user multiple-antenna wireless links operating in Rayleigh flat-fading, where the receiver knows the propagation matrix. In particular, the capacity grows linearly with the smaller of the number of transmit or receive antennas with no extra bandwidth or total power. Moreover, [3] discloses a practical scheme referred to as BLAST (Bell Labs Layered Space Time) for realizing a significant fraction of the capacity with small outage probability, using a divide-and-conquer strategy based on ordinary modulation and coding techniques.

BLAST, as well as certain other space-time codes [12], requires the receiver to know the propagation matrix between the transmit and receive antennas. This knowledge can be acquired by sending known training signals. The required training interval is proportional to the number of transmit antennas [8], and for many potential applications, training is an acceptable burden. However, in TDMA applications with fast fading, for example, both training and data transmission may have to occur during a relatively short interval. Because both the training interval and capacity increase linearly with the number of transmit antennas, the total throughput is maximized by choosing the number of transmit antennas such that half of the interval is used for training, and half for data transmission [8].

Ideally one would like to achieve BLAST-like transmission rates with multiple antennas while circumventing training and channel estimation. Some steps in this direction are described in [1, 9, 5, 14], based on a piecewise constant model for fading (also called block fading [10, 2]). Here the random propagation matrix (which nobody knows) remains constant for a $T$ symbol coherence interval, after which it jumps to a new independent value where it remains for another $T$ symbols, and so on. This constitutes a memoryless channel from one coherence interval to another for matrix-valued signals, which permits a direct application of Shannon theory implicitly involving coding over many coherence intervals. During every coherence interval a $T \times M$ complex matrix is transmitted and a $T \times N$ complex matrix is received, where $M$ and $N$ are the number of transmit and receive antennas respectively. It was shown [9] that capacity cannot be increased by making $M > T$, and that the capacity attaining signals are equal to the product of two independent matrices: a $T \times M$ isotropically distributed unitary matrix, and a $M \times M$ diagonal, real, nonnegative matrix. This structure motivates the use of *unitary space-time modulation* [5] involving a constellation of $L$ $T \times M$ unitary matrices $\{\Phi_0, \cdots, \Phi_{L-1}\}$, where $\Phi_\ell^\dagger \Phi_\ell = I_M$, chosen according to a design criterion that differs markedly from the familiar maximum-Euclidean-distance criterion.

Some small ($L = 64$) unitary space-time constellations are designed in [5] using a simple iterative algorithm. A systematic approach is pursued in [7], where an initial $T \times M$ unitary matrix is successively rotated $L-1$ times to generate the entire constellation of signals. The rotation matrix is $T \times T$ diagonal, with its diagonal elements equal to $L$th roots of unity, and with the initial signal comprising $M$ columns from a $T \times T$ discrete Fourier transform unitary matrix. Using iterative random search, the roots that characterize the rotation matrix are chosen to give a low raw (uncoded) block probability of error for the constellation, based on pairwise probabilities of error. The search is facilitated by the fact that the correlation between the signals, which determines the pairwise probabilities of error, has a circulant structure. Using this approach, constellations larger than $L = 2000$ have been designed. However it was not established how restrictive the circulant structure is, or whether significant improvements in performance are possible by relaxing this structure.

The recent *space-time autocoding* effect [6] implies that arbitrarily reliable communication can be achieved within a single coherence interval if $T$ and $M$ simultaneously become large. There is a positive *autocapacity*, such that for any rate $R$ less than the autocapacity, the block probability of error for a single coherence interval goes to zero as $T$ and $M$ grow large, with no knowledge of the propagation matrix available to anybody. In effect, temporal diversity - which is unreliable for stop-and-go mobiles - is replaced by spatial diversity. Thus, autocoding says that we may avoid channel coding that is normally performed over many independent coherence intervals, and shift the problem of achieving reliability to the problem of designing an effective constellation of $T \times M$ signals.

Achieving autocapacity theoretically requires unbounded $M$ and $T$, but the autocoding effect manifests itself for relatively small $T$ and $M$, and transmission rates that are a significant fraction of autocapacity can theoretically be supported with extremely small probabilities of error. For example, using $M = 7$ transmit antennas and $N = 4$ receive antennas, and with an expected SNR of 18 dB, a single user can theoretically transmit 80 bits during a single $T = 16$ symbol coherence interval (rate 5 bits/symbol) with a block probability of error less than $10^{-9}$, all without any training or knowledge of the propagation matrix. These performance predictions are obtained by applying a union bound and an expression for pairwise probability of error to a hypothetical codebook of $L = 2^{80}$ independent isotropically random $16 \times 7$ unitary matrices.

A constellation of $2^{80}$ independent matrices is impossible to generate exhaustively or to store, and because of its lack of structure there is little hope of ever finding a fast decoding scheme. This note proposes a unitary space-time constellation that, although random, is structured, and has exactly the same union-bound performance as a constellation of independent signals.

Section 2 reviews the signal model, unitary space-time modulation, space-time autocoding, and the earlier systematic construction. Section 3 explains the new construction. Section 4 reinterprets earlier systematic constructions such as [7] in light of this new construction. The mathematical results required for many of the conclusions of this paper are developed in the appendices: Appendix A obtains the eigenvalue distribution for isotropically distributed unitary matrices, and Appendix B shows that any power (larger than one) of an isotropically distributed unitary matrix is not isotropically distributed. In fact a limiting distribution is obtained for large enough finite powers.

## 2 Background and Problem Statement

A single user has access to a multiple antenna wireless link in a Rayleigh flat-fading environment with no knowledge of the propagation matrix, and the goal is to transmit a large number of bits reliably during one coherence interval. The recently discovered space-time autocoding effect implies that, for any rate less than the autocapacity, the block probability of error goes to zero as the duration of the coherence interval and the number of transmit antennas increase simultaneously. A significant fraction of the autocapacity can theoretically be realized in a typical scenario with low probability of error using a large constellation of isotropically random unitary space-time signals.

### 2.1 Signal model

There are $M$ transmit antennas and $N$ receive antennas operating in a Rayleigh flat fading environment. During a $T$-symbol coherence interval, over which the propagation coefficients are constant, a single user transmits a $T \times M$ complex matrix S, and another user receives a complex $T \times N$ matrix X,

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

where $H$ is a $M \times N$ propagation matrix, whose elements are independent $\mathcal{CN}(0,1)$, and $W$ is an independent $T \times N$ receiver noise matrix whose elements are independent $\mathcal{CN}(0,1)$. There is a power constraint

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

and $\rho$ represents the expected SNR at each receive antenna.

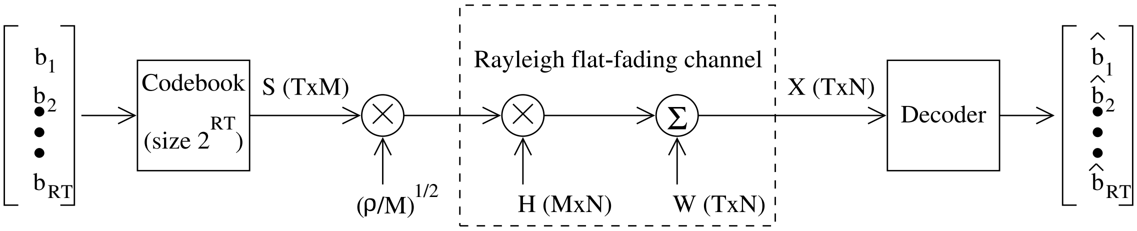


Figure 1: Wireless link comprising $M$ transmitter and $N$ receiver antennas. We wish to transmit $R \cdot T$ bits of information reliably in a *single* coherence interval $T$, where $R$ is the rate in bits/symbol.

We assume throughout the paper that the random propagation matrix is unknown to both the transmitter and the receiver.

## 2.2 Space-time autocoding and unitary space-time modulation

We wish to transmit a total of $R \cdot T$ bits, for some rate $R$, during a single coherence interval as illustrated in Figure 1. There is a positive autocapacity $C_a$ [6], such that for all $R < C_a$, there exists a code such that the block probability of error goes to zero exponentially as $T, M \to \infty$. The autocapacity, in units of bits/symbol, is given by the simple formula

$$C_a = N \, \log(1 + \rho). \tag{3}$$

Within a finite duration coherence interval $T < \infty$, bits can theoretically be transmitted at rates below the autocapacity with low probability of error using a random codebook of $L$ independent isotropically random (Section 3 describes the isotropic distribution in some detail) $T \times M$ unitary space-time signals [6], $\left\{ S_\ell = \sqrt{T}\Phi_\ell, \ \ell = 0, \cdots, L-1 \right\}$, where $L = 2^{RT}$, and where the column vectors of each $\Phi_\ell$ are orthonormal, $\Phi_\ell^\dagger \Phi_\ell = I_M$. The block probability of error $P_e$ may be upper-bounded through the union bound

$$P_e < 2^{RT} \, \mathrm{E}_{\Phi_1, \Phi_2} \left\{ P_e \left\{ \Phi_1 \text{ vs. } \Phi_2 \right\} \right\}, \tag{4}$$

where $P_e \{\Phi_1 \text{ vs. } \Phi_2\}$ is the pairwise (e.g., two-signal constellation) probability of error that is associated with any distinct pair of signals in the constellation [5], given by the exact formula

$$P_e \left\{ \Phi_1 \text{ vs. } \Phi_2 \right\} = \frac{1}{\pi} \int_0^{\pi/2} d\theta \prod_{m=1}^{M} \left[ \frac{\cos^2\theta}{\cos^2\theta + \frac{(\rho T/M)^2 (1-d_m^2)}{4(1+\rho T/M)}} \right]^N , \tag{5}$$

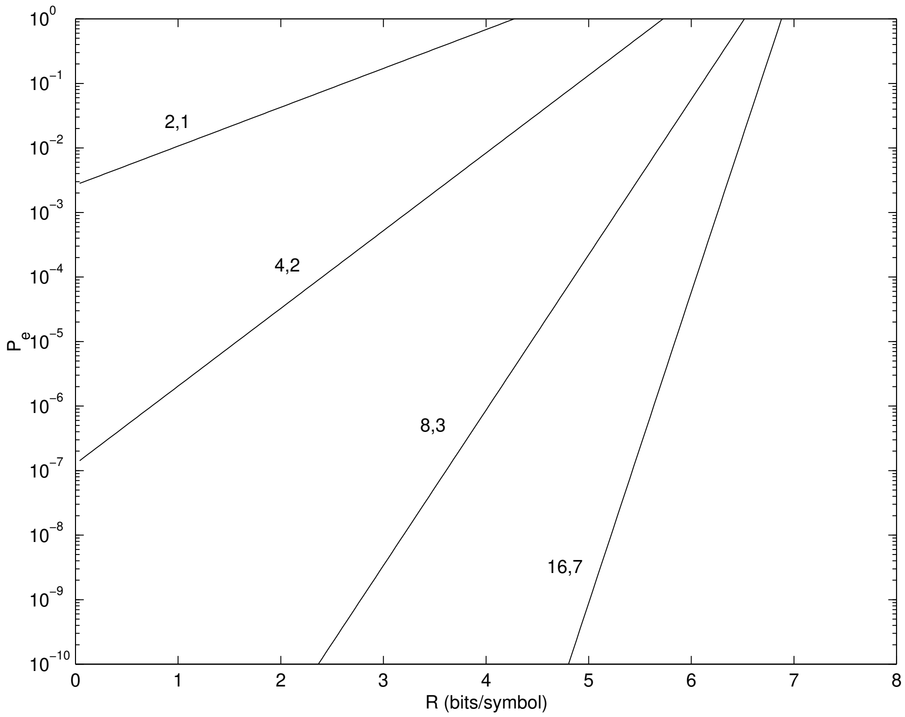


Figure 2: [6] Upper bound on block probability of error versus transmission rate (bits/symbol) for random codebook of unitary space-time signals, for $N = 4$, $\rho = 18$ dB, and $(T, M) = (2, 1), (4, 2), (8, 3), (16, 7)$. Autocapacity is equal to $24.01$ bits/symbol.

where $d_1, \cdots, d_M$ are the singular values of the $M \times M$ matrix $\Phi_2^\dagger \Phi_1$. The expectation with respect to the singular values may be brought inside the integral in (5) and, when the two signals are independent, may be obtained in closed-form; $\theta$ can be integrated numerically. We omit all the details and refer the interested reader to [6].

Figure 2 displays the bound (4) as a function of the transmission rate $R$, for an $18$ dB expected SNR, $N = 4$ receive antennas, for $T = 2, 4, 8, 16$, and for $M = 1, 2, 3, 7$ respectively. For the larger values of $T$, transmission rates as high as $25\%$ of the autocapacity $C_a = 24.01$ bits/symbol can theoretically be sustained with very low probability of error. However to realize the autocoding effect, constellations of unprecedented size ($L = 2^{80}$ for $T = 16$, $R = 5$, and $P_e < 10^{-9}$) [1] must be constructed and decoded.

[1]The union bound is almost certainly too conservative. If all of the signals were perfectly orthogonal to each other, the union bound would be $P_e = 10^{-9}$ for $R = 7.5$.

### 2.3 Earlier systematic constructions [7]

The design of constellations of unitary space-time signals when the propagation matrix is unknown involves a criterion that differs considerably from the usual maximum Euclidean distance criterion [5, 7]. We see from (5) that making the singular values of the "correlation matrix" $\Phi_2^\dagger \Phi_1$ as small as possible is beneficial. In particular, making $\Phi_2^\dagger \Phi_1 = 0$ is ideal, but this is not generally possible for all possible pairs of elements in a constellation. Constellations of unitary space-time signals have a block probability of error that is invariant to certain transformations: 1) left multiplication by a common $T \times T$ unitary matrix, $\Phi_\ell \to \Psi^\dagger \Phi_\ell$, $\ell = 0, \cdots, L-1$; 2) right multiplication by individual $M \times M$ unitary matrices, $\Phi_\ell \to \Phi_\ell \Upsilon_\ell$, $\ell = 0, \cdots, L-1$. Any constellations that are related by transformations of this type are considered to be equivalent.

The problem of constructing moderately large constellations of unitary space-time signals is addressed in [7] with the goal of achieving a low block probability of error. The construction proposed in [7] involves successive rotations of an initial signal in $T$-dimensional complex space,

$$\Phi_\ell = \Omega^\ell \Phi_0, \quad \ell = 0, \cdots, L-1, \tag{6}$$

where $\Omega$ is a $T \times T$ unitary matrix, and $\Phi_0$ is the $T \times M$ initial signal obeying $\Phi_0^\dagger \Phi_0 = I_M$. A judicious choice of $\Omega$ and $\Phi_0$, is needed to make the columns of $\Phi_\ell$ zig-zag over the surface of the $T$-dimensional complex sphere.

There is no loss of generality in assuming that $\Omega$ is diagonal because, as shown in Appendix A, the Schur factorization [11] implies that any square unitary matrix $\Omega$ can be expressed as

$$\Omega = \Psi \Theta \Psi^\dagger, \tag{7}$$

where $\Psi$ is $T \times T$ unitary, and $\Theta$ is $T \times T$ diagonal matrix of eigenvalues of $\Omega$. The transformation, $\Phi_\ell \to \Psi^\dagger \Phi_\ell$, produces an equivalent constellation that is generated by a diagonal rotation matrix that comprises the eigenvalues of $\Omega$,

$$\Phi_\ell = \Theta^\ell \Phi_0, \quad \ell = 0, \cdots, L-1. \tag{8}$$

In [7] further structure is imposed by choosing a) $\Theta$ to be an $L$th root of the identity matrix, implying that $\Theta_{tt} = e^{i2\pi u_t/L}$, where $u_t \in \{0, 1, \cdots, L-1\}$; b) $\Phi_0$ to comprise $M$ distinct columns of a $T \times T$ discrete Fourier transform matrix. The integers $u_1, \ldots, u_T$ and the DFT columns are chosen by iterative random search with the goal of minimizing the maximum pairwise probability of error (more precisely, an upper bound on the Chernoff bound) between all distinct pairs of signals in the constellation.

Choosing $\Theta$ to be an $L$th root of the identity matrix makes the correlation between the signals, which determines the pairwise probabilities of error, have a circulant structure, i.e., $\Phi_{\ell'}^{\dagger}\Phi_{\ell}$ depends only on $(\ell' - \ell) \bmod L$. Conversely, any constellation that has circulant correlation structure is equivalent to one that has the construction (8).

The circulant structure implies that the conditional probability of error is the same for every signal in the constellation, it simplifies the iterative design since only $L-1$ rather than $(L^2 - L)/2$ correlations have to be checked, and it has some intuitive appeal. However, no indication is given in [7] as to how restrictive this structure really is, or whether significant improvements could be obtained by relaxing this structure. Moreover, the iterative optimization could never be used for an $L = 2^{80}$ constellation.

In what follows, we show that gains can indeed be obtained by relaxing the structure, and propose a method for designing constellations that can readily generate $2^{80}$ signals.

## 3 Structured Constellation with Good Average Performance

Our approach to specifying constellations of unitary space-time signals is based on the observation that the union bound (4), where the expected pairwise probability of error is identical for all distinct pairs, only requires $\Phi_{\ell'}$ and $\Phi_{\ell}$ to be pairwise independent isotropically distributed matrices for all $\ell' \neq \ell$. Any constellation having marginally isotropically random and pairwise independent signals would have exactly the same union-bound performance (as given by Fig. 2, for example) as a constellation of independent unitary space-time signals, no matter what other probability dependencies they may have. We now demonstrate a construction that has pairwise independence and is easy to generate.

Our signals are represented by $R \cdot T$ binary indices, and they are generated as follows,

$$\Phi_{\ell_1 \ell_2 \cdots \ell_{RT}} = \Omega_1^{\ell_1} \Omega_2^{\ell_2} \cdots \Omega_{RT}^{\ell_{RT}} \Phi_{00\cdots 0}, \quad \ell_1, \ell_2 \, \ldots, \ell_{RT} \in \{0, 1\}, \tag{9}$$

where the $\Omega_1, \ldots, \Omega_{RT}$ are independent $T \times T$ isotropically distributed unitary matrices. We let $\Phi_{00\cdots 0}$ be an independent $T \times M$ isotropically distributed unitary matrix.

A $T \times T$ random unitary matrix $\Omega$ is isotropically distributed if its probability density is unchanged when $\Omega$ is premultiplied by any $T \times T$ deterministic unitary matrix. From this definition, one may deduce [9] that a) there is exactly one probability density that possesses this property, with the formula given by

(A.1); b) the density is invariant to postmultiplication of $\Theta$ by any deterministic unitary matrix. Likewise, a $T \times M$ random unitary matrix $\Phi$, i.e., $\Phi^\dagger \Phi = I_M$, is isotropically distributed if its probability density is invariant to premultiplication by any deterministic $T \times T$ unitary matrix. An oblong matrix of this type has the same density as any $M$ columns of a $T \times T$ isotropically distributed unitary matrix.

We now show that the signal matrices (9) are marginally isotropically distributed and pairwise independent. They are marginally isotropically distributed because any signal is equal to the $T \times M$ isotropically distributed unitary matrix $\Phi_{00\cdots 0}$ premultiplied by an independent $T \times T$ unitary matrix. Conditioned on this $T \times T$ factor, the signal is therefore isotropically distributed and not dependent on this factor. Since the conditional density is independent of the factor it follows that the unconditional distribution is also isotropic.

Let $\ell$ denote the vector comprising the $RT$ binary indices $\{\ell_1, \cdots, \ell_{RT}\}$, and consider two distinct signals from the constellation (9), $\Phi_\ell$ and $\Phi_k$, for $k \neq \ell$. We wish to establish that the signals are independent. With $\ell = 0$ denoting the vector of all zeros, the signals may be expressed as

$$\Phi_\ell = \Xi_\ell \Phi_0, \qquad \Phi_k = \Xi_k \Phi_0,$$

where $\Xi_\ell$ and $\Xi_k$ are products of certain subsets of $\{\Omega_1, \cdots, \Omega_{RT}\}$. At least one of $\Xi_\ell$ or $\Xi_k$ has a factor $\Omega_q$ that is not contained in the other. Assume without loss of generality that $\Omega_q$ is a factor of $\Xi_\ell$, which then takes the form

$$\Xi_\ell = A \Omega_q B.$$

Consider the following argument:

- $\Omega_q$ does not appear in $A$, $B$, or $\Xi_k$, and therefore $\Omega_q$ is independent of these matrices.
- Conditioned on $A$, the product $A\Omega_q$ is isotropically distributed. Since this conditional density does not depend on $A$, the product $A\Omega_q$ is isotropic and independent of $A$. The product is also independent of $\Xi_k$.
- By a similar argument, $\Xi_\ell = A\Omega_q B$ is isotropic and independent of $B$ and $\Xi_k$.
- Finally, because $\Phi_0^\dagger \Phi_0 = I$, conditioned on $\Phi_0$, $\Phi_\ell = \Xi_\ell \Phi_0$ is $T \times M$ isotropically distributed, and since the conditional density for $\Phi_\ell$ has no dependence on either $\Phi_0$ or on $\Xi_k$, we conclude that $\Phi_\ell$ and $\Phi_k = \Xi_k \Phi_0$ are independent.

We note a final simplification that can be made to the construction (9). Specifically, the first signal $\Phi_0$ can be chosen to be a deterministic (rather than isotropically random) $T \times M$ unitary matrix without changing the block probability of error for the constellation. To see this, we introduce the $T \times (T-M)$ orthogonal complement $\Phi_{0\perp}$ to the isotropically distributed $\Phi_0$, such that together they form a $T \times T$ isotropically distributed matrix $\Psi = [\Phi_0 \; \Phi_{0\perp}]$. Every signal in the constellation can be premultiplied by $\Psi^\dagger$ without changing the block probability of error. This transformation gives

$$\begin{aligned} \Psi^\dagger \Phi_{\ell_1 \cdots \ell_{RT}} &= \Psi^\dagger \Omega_1^{\ell_1} \cdots \Omega_{RT}^{\ell_{RT}} \Psi \Psi^\dagger \Phi_0 \\ &= \left(\Psi^\dagger \Omega_1 \Psi\right)^{\ell_1} \cdots \left(\Psi^\dagger \Omega_{RT} \Psi\right)^{\ell_{RT}} \begin{bmatrix} I_M \\ 0 \end{bmatrix}. \end{aligned}$$

By our now-standard argument, the product $\Psi^\dagger \Omega_j$ is isotropic and independent of $\Psi$. Likewise, $\Psi^\dagger \Omega_j \Psi$ is isotropically distributed and independent of $\Psi$. Therefore the constellation (9) with $\Phi_0$ isotropically distributed is equivalent to a constellation where $\Phi_0^\dagger = \begin{bmatrix} I_M & 0 \end{bmatrix}$.

To summarize, a constellation of $2^{RT}$ independent isotropically random unitary space-time signals can be replaced, without altering its union bound performance, by a highly structured random constellation that is specified by only the initial $T \times M$ signal and by $RT$ isotropically distributed $T \times T$ unitary matrices.

# 4 Comments on Earlier Systematic Construction of [7]

In the previous section we presented a highly structured random constellation that is based on $R \cdot T$ independent isotropically random rotation matrices and which has the same (good) union-bound performance (4) as a fully random constellation of independent isotropically random signals. We now demonstrate why this new scheme is superior to the earlier construction (6) that is based on a single rotation matrix.

## 4.1 Single isotropically random rotation matrix

Consider the construction (6), with the initial signal $\Phi_0$ isotropically random unitary, and with the rotation matrix $\Omega$ independent isotropically random unitary. Because $\Phi_0$ is isotropically distributed, all of the signals in the constellation are marginally isotropically distributed. For the construction to have the same union bound performance as the construction (9), we would require every $\Omega^\ell$, $\ell = 2, \ldots, L-1$ to be marginally isotropically distributed. For then, by an argument that is parallel to that of the previous section, any two distinct signals would be pairwise independent.

It is an intuitively appealing proposition that, if $\Omega$ is isotropically distributed, then $\Omega^\ell$ is isotropically distributed for any integer $\ell \geq 2$. Indeed if this were true, the successive rotations $\Theta^\ell \Phi_0$ would zig-zag uniformly over the surface of the $T$-dimensional complex sphere, as hoped. However, this turns out not to be true. The reason follows directly from the eigenvector/eigenvalue decomposition (7). As shown in Appendix A, the eigenvectors and the eigenvalues of an isotropic matrix are independent of each other, the $T \times T$ eigenvector matrix $\Psi$ is itself isotropically distributed, and the $T$ eigenvalues $\lambda_1, \ldots, \lambda_T$ have the density

$$p_\lambda(\lambda) = \frac{1}{T!\,\pi^T} \prod_{t=1}^{T} \delta\left(|\lambda_t|^2 - 1\right) \prod_{s>t} |\lambda_s - \lambda_t|^2 . \tag{10}$$

The eigenvalues have unit magnitude, while the phases $\theta_t$ in $\lambda_t = e^{i\theta_t}$ are distributed as

$$p(\theta) = \frac{2^{T^2-T}}{T!\,(2\pi)^T} \prod_{t=1}^{T} \prod_{s>t} \sin^2\left(\frac{\theta_s - \theta_t}{2}\right). \tag{11}$$

On the other hand, the eigenvalues of $\Omega^\ell$, which again have unit magnitude and we denote $\mu_t = \lambda_t^\ell,\ t = 1, \ldots, T$ have phases $\gamma_1, \ldots, \gamma_T$ distributed as

$$p(\gamma) = \frac{2^{T^2-T}}{T!(2\pi)^T \ell^T} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \prod_{s>t} \sin^2\left[\frac{\gamma_s - \gamma_t + 2\pi(j_s - j_t)}{2\ell}\right]. \tag{12}$$

We show in Appendix B that if $T > 1$, $\Omega^\ell$ is *not* isotropically distributed for any $\ell \geq 2$. Furthermore $\Omega^\ell$ reaches a limiting density for all $\ell \geq T$, where the eigenvalue phases are independent and uniformly distributed

$$p(\gamma) = \left(\frac{1}{2\pi}\right)^T, \quad \ell \geq T. \tag{13}$$

These results are counterintuitive, as most of our experience with isotropically distributed unitary matrices is based on the real $2 \times 2$ case, where all powers *are* isotropically distributed. Even so, this familiar property does not generalize to real spaces of dimension greater than two: the $\ell$-th power of a $3 \times 3$ real isotropically distributed orthogonal matrix is not isotropic, as illustrated in Figure 3. Let $\Psi$ be a $3 \times 3$ real, isotropically distributed orthogonal matrix, and let $e_x$ be the unit vector that points in the $x$-direction. The figure shows that the product $\Psi e_x$ is equally likely to lie anywhere on the unit sphere. However, the product $\Psi^\ell e_x$, for even powers of $\ell$ is biased towards $e_x$, and for odd powers of $\ell$ is biased towards $\pm e_x$.

For two signals $\Phi_\ell$ and $\Phi_k$, the pairwise error probability depends on $\Phi_\ell^\dagger \Phi_k = \Phi_0^\dagger \Omega^{k-\ell} \Phi_0$, and if $|\ell - k| \geq T$ then $\Omega^{k-\ell}$ has independent uniformly distributed eigenvalue phases. Because $L$ is often very

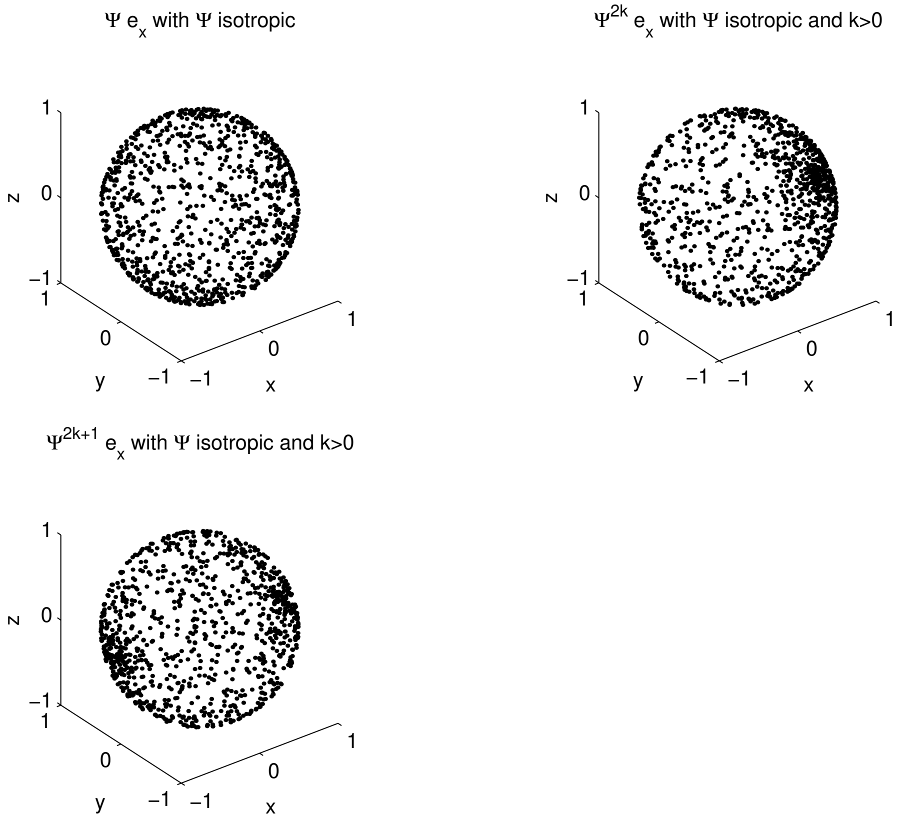


Figure 3: The power of a $3\times 3$ real, isotropically distributed orthogonal matrix is not isotropically distributed. Upper left: single application of random rotation to the unit vector $e_x$ results in a unit vector that is equally likely to lie anywhere on the unit sphere. Upper right: an even number of applications of the same random matrix is biased towards $e_x$. Lower left: an odd number of applications of the same random matrix is biased towards $\pm e_x$.

large ($L = 2^{80}$), most of the pairwise signals have an effective $\Omega^{k-\ell}$ with this phase distribution, even if $\Omega$ is itself isotropic. We therefore look briefly at the performance of a constellation that is generated by a single rotation matrix $\Omega$ that is not isotropic, but rather has independent uniformly distributed eigenvalue phases. As we show, this constellation does not perform as well as the constellation (9).

### 4.2 Rotation matrix with independent, uniform-phase eigenvalues

Consider the construction (6) where $\Phi_0$ is isotropic but $\Omega$ has eigenvalues with independent, uniform $[0, 2\pi)$ phases (its eigenvectors are still isotropic). Then all of the signals are marginally isotropically distributed and any two distinct signals have the same joint distribution. We are unable to take the expectation of the pairwise probability of error (5) analytically, so we use Monte-Carlo integration. The resulting union bound is shown in Figure 4. We used $10^5$ trials to generate each pairwise probability of error so the curves $(T, M) = (8, 3)$ and $(T, M) = (16, 7)$ are only approximate; nevertheless, a comparison with Figure 2 shows that the single-rotation matrix construction is worse than the construction (9).

## 5 Conclusions

The realization of the space-time autocoding effect requires two things: 1) a good, extraordinarily large, constellation of unitary space-time signals, and 2) a decoding algorithm that avoids exhaustive search over the constellation. We have addressed the first problem with a construction that is random but highly structured. The $L$ signals are specified by $\log_2 L$ isotropically distributed $T \times T$ unitary matrices. A random constellation of this sort has, in a deterministic sense, no guaranteed performance. However, due to the size of the constellation, we expect that any given realization will have a performance very close to the average performance of the fully random constellation, which we have shown to be provably good.

These constellations can support transmission rates that are a significant fraction of the autocapacity with extremely low probabilities of error, and their low complexity makes them practical for the transmitter to employ. Their ultimate utility, however, depends entirely on the discovery of a good decoding algorithm.

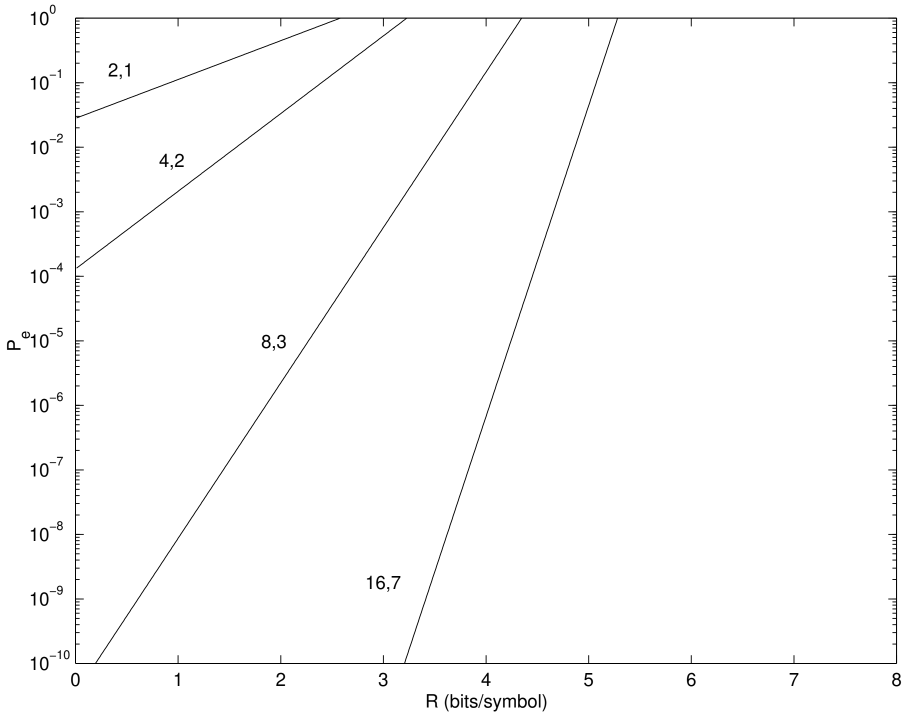


Figure 4: Upper bound on block probability of error versus transmission rate (bits/symbol) for codebook of unitary space-time signals, generated from a single rotation matrix having independent, uniform-phase eigenvalues, for $N = 4$, $\rho = 18$ dB, and $(T, M) = (2, 1), (4, 2), (8, 3), (16, 7)$.

# A Appendix: Eigenvalue Distribution of Isotropically Random Unitary Matrix

We derive the joint probability density of the eigenvectors and eigenvalues of an isotropically random unitary matrix. This density is known, see, e.g., [4], but our presentation is self-contained and particularly simple.

A $T \times T$ random unitary matrix $\Phi$ is isotropically distributed if its probability density is unchanged when $\Phi$ is pre-multiplied by any $T \times T$ deterministic unitary matrix. This operational definition leads directly to the unique probability density [9] as a function of the $T$ column vectors $\{\phi_1, \ldots, \phi_T\}$,

$$\begin{aligned} p(\Phi) &= p(\phi_1) \cdot \prod_{s=2}^{T} p\left(\phi_s \mid \phi_1, \cdots, \phi_{s-1}\right) \\ &= \left[\frac{\Gamma(T)}{\pi^T}\delta\left(\phi_1^\dagger\phi_1 - 1\right)\right] \cdot \prod_{s=2}^{T}\left[\frac{\Gamma(T+1-s)}{\pi^{T+1-s}}\delta\left(\phi_s^\dagger\phi_s - 1\right)\prod_{t=1}^{s-1}\delta\left(\phi_s^\dagger\phi_t\right)\right] \end{aligned} \tag{A.1}$$

where the Dirac delta function of a complex-valued argument is interpreted as $\delta(z) = \delta(\mathrm{Re}\{z\}) \cdot \delta(\mathrm{Im}\{z\})$. This density, defined with respect to Lebesgue measure, is invariant to postmultiplication of its argument by any deterministic unitary matrix.

We frequently will need to integrate coupled Dirac delta functions whose arguments are nonlinear functions. Let $f(x)$ be a $K$-component vector-valued nonlinear function of a $K$-dimensional real-valued vector $x$, and consider the integral

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

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

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

This expression 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.

### A.1 Independence of eigenvectors and eigenvalues; isotropically distributed eigenvectors

The orthogonality of the eigenvectors follows directly from the Schur decomposition [11] (which applies to any square matrix) of the isotropically distributed unitary matrix $\Phi$,

$$\Phi = \Psi \Lambda \Psi^\dagger,$$

where $\Psi$ is $T \times T$ unitary, and $\Lambda$ is $T \times T$ lower triangular. Since $\Psi$ and $\Phi$ are both unitary the same must be true of $\Lambda = \Psi^\dagger \Phi \Psi$, which, since $\Lambda$ is lower triangular, implies that $\Lambda$ is diagonal. Thus, the Schur decomposition of a unitary matrix coincides with its eigenvector/eigenvalue representation

$$\Phi = \Psi \Lambda \Psi^\dagger, \tag{A.3}$$

with $\Psi$ is $T \times T$ unitary, and $\Lambda$ is $T \times T$ diagonal. We first establish that $\Psi$ and $\Lambda$ are independent, and that $\Psi$ itself is isotropically distributed. We rewrite (A.3) as

$$\Phi \Psi = \Psi \Lambda,$$

and we multiply both sides of this expression by a $T \times T$ independent isotropically distributed unitary matrix $\Upsilon$

$$\Upsilon \Phi \Psi = \Upsilon \Psi \Lambda,$$

or

$$\left(\Upsilon \Phi \Upsilon^\dagger\right)(\Upsilon \Psi) = (\Upsilon \Psi)\Lambda. \tag{A.4}$$

Now $\Lambda$ is the eigenvalue matrix, and $(\Upsilon\Psi)$ the eigenvector matrix, of $\left(\Upsilon \Phi \Upsilon^\dagger\right)$.

- The product $\Upsilon\Psi$ is isotropically distributed, and it is also isotropically distributed conditioned on $\Lambda$. Consequently $\Upsilon\Psi$ is independent of $\Lambda$.

- Since $\Upsilon\Phi$ is independent of $\Upsilon$ and isotropically distributed, it follows that $\Upsilon\Phi\Upsilon^\dagger$ is isotropically distributed as well.

- Thus, (A.4) implies that the eigenvector matrix $\Upsilon\Psi$ of the particular isotropically distributed unitary matrix $\Upsilon\Phi\Upsilon^\dagger$ is itself isotropically distributed and independent of the eigenvalues. There is nothing special about $\Upsilon\Phi\Upsilon^\dagger$, so the same thing must be true for *any* isotropically distributed unitary matrix.

## A.2 Eigenvalue density

We have established that the eigenvectors of an isotropically random unitary matrix are also isotropically distributed and independent of the eigenvalues. Starting with the decomposition (A.3) we have

$$\begin{aligned} \Phi &= \Psi P^\dagger P \Lambda P^\dagger P \Psi^\dagger \\ &= \acute{\Psi} P \Lambda P^\dagger \acute{\Psi}^\dagger, \end{aligned} \tag{A.5}$$

where $P$ is any $T \times T$ permutation matrix. The unitary matrix $\acute{\Psi}$ is still isotropically distributed and independent of $P$ and of $P\Lambda P^\dagger$. In turn $P\Lambda P^\dagger$ is diagonal, with its on-diagonal elements equal to the permuted version of the diagonal elements of $\Lambda$. Consequently the eigenvalues of $\Phi$ may be ordered in any manner, and the eigenvectors will still be isotropically distributed and independent of the eigenvalues. In what follows we assume that the eigenvalues are ordered randomly.[2] It follows that, for any $T \times T$ permutation matrix $P$,

$$p_\lambda\left(P\lambda\right) = p_\lambda\left(\lambda\right). \tag{A.6}$$

We obtain the joint probability density for the vector of eigenvalues $\lambda$ by working backward from its unknown density to the known density for $\Phi$. The conditional density for $\Phi$ is

$$p(\Phi \mid \lambda, \Psi) = \delta\left(\Phi - \Psi D_\lambda \Psi^\dagger\right), \tag{A.7}$$

where $D_\lambda$ is the $T \times T$ diagonal matrix whose diagonal elements comprise the elements of the vector $\lambda$, and where $\delta\left(\Phi - \Psi D_\lambda \Psi^\dagger\right)$ is equal to the product of $T^2$ complex Dirac delta functions involving the $T^2$ complex components of $\Phi$. The expectation of this expression yields the known marginal isotropic density

[2]This can always be done by multiplying the vector of eigenvalues by a random permutation matrix.

for $\Phi$ (cf. (A.1)) in terms of the isotropic density for $\Psi$ and the unknown density for $\lambda$,

$$p(\Phi) = \int d\lambda \int d\Psi \, p(\lambda) \, p(\Psi) \, \delta \left( \Phi - \Psi D_\lambda \Psi^\dagger \right). \tag{A.8}$$

We analytically integrate with respect to $\Psi$ to obtain a linear integral equation for $p(\lambda)$ (a multivariate Fredholm equation of the first kind) that we solve by inspection.

As noted earlier, the Schur decomposition for any square $\Phi$ is

$$\Phi = \Xi \left[ D_v + \tilde{R} \right] \Xi^\dagger, \tag{A.9}$$

where $\Xi$ is unitary, $D_v$ is a diagonal matrix whose diagonal entries we denote by the $T$-dimensional vector $v$, and $\tilde{R}$ is strictly lower-triangular, i.e., $\tilde{r}_{st} = 0, \ t \geq s$. Because $\Phi$ is unitary $\tilde{R}$ is zero, but we carry $\tilde{R}$ with us anyway because we are specifying a density on $\Phi$ and need to keep track of the delta functions that disallow nonunitary $\Phi$.

The substitution of (A.9) into (A.8) gives

$$p_\Phi \left( \Xi \left[ D_v + \tilde{R} \right] \Xi^\dagger \right) = \int d\lambda \int d\Psi p(\lambda) \, p(\Psi) \, \delta \left( \Xi \left[ D_v + \tilde{R} \right] \Xi^\dagger - \Psi D_\lambda \Psi^\dagger \right). \tag{A.10}$$

The isotropic density $p(\Phi)$ is invariant to pre- and post-multiplication of its argument by unitary matrices, which implies that

$$p_\Phi \left( \Xi \left[ D_v + \tilde{R} \right] \Xi^\dagger \right) = p_\Phi \left( D_v + \tilde{R} \right). \tag{A.11}$$

The integral with respect to $\Psi$ in (A.10) can be simplified as follows,

$$\begin{aligned} \int d\Psi \, p(\Psi) \, \delta \left( \Xi \left[ D_v + \tilde{R} \right] \Xi^\dagger - \Psi D_\lambda \Psi^\dagger \right) &= \int d\Psi \, p(\Psi) \, \delta \left( D_v + \tilde{R} - \Xi^\dagger \Psi D_\lambda \Psi^\dagger \Xi \right) \\ &= \int d\Psi \, p(\Psi) \, \delta \left( D_v + \tilde{R} - \Psi D_\lambda \Psi^\dagger \right), \end{aligned} \tag{A.12}$$

where 1) we have used the fact that, for any $T \times T$ complex matrix $X$ and for any $T \times T$ unitary matrix $\Xi$,

$$\delta(X) = \delta \left( \Xi^\dagger X \right) = \delta \left( \Xi^\dagger X \Xi \right) \tag{A.13}$$

and 2) we made the coordinate transformation $\Xi^\dagger \Psi \to \Psi$ and used the fact that $p(\Psi)$ is invariant to rotations.

Substituting (A.11) and (A.12) into (A.10) yields

$$p_\Phi\left(D_v+\tilde{R}\right)=\int d\lambda\int d\Psi p(\lambda)\,p(\Psi)\,\delta\left(D_v+\tilde{R}-\Psi D_\lambda\Psi^\dagger\right). \tag{A.14}$$

Because $\Phi$ is unitary, the probability density $p_\Phi(\cdot)$ is supported at $|v_1|=\ldots=|v_T|=1$ and $\tilde{R}=0$. To compute $p_\Phi(D_v+\tilde{R})$ we apply the formula (A.2) to the delta functions in the isotropic density (A.1). To compute the needed gradient at $\tilde{R}=0$, we expand

$$\left(D_v+\tilde{R}\right)^\dagger\left(D_v+\tilde{R}\right)-I\approx D_v^\dagger D_v-I+D_v^\dagger\tilde{R}+\tilde{R}^\dagger D_v$$

to first-order terms in $\tilde{R}$. Substituting into (A.1) gives

$$\begin{aligned}p_\Phi\left(D_v+\tilde{R}\right)&=c\cdot\prod_{t=1}^{T}\delta\left(|v_t|^2-1\right)\prod_{s>t}\delta\left(v_s^*\tilde{r}_{st}+v_t\tilde{r}_{ts}^*\right)\\&=c\cdot\prod_{t=1}^{T}\delta\left(|v_t|^2-1\right)\prod_{s>t}\delta\left(\tilde{r}_{st}\right),\end{aligned} \tag{A.15}$$

where we have used the strict lower-triangularity of $\tilde{R}$, and $c$ is the normalizing constant for the isotropic density. We have thus computed the left-hand side of (A.14).

Consider next the right-hand side of (A.14). At $\tilde{R}=0$, the $T^2$ complex delta functions imply that $D_v=\Psi D_\lambda\Psi^\dagger$, while the isotropic density for $\Psi$ implies that $\Psi^\dagger\Psi=I$. There are $T!$ families of solutions for $\Psi$ and $\lambda$ to these sets of equations,

$$\Psi=D_zP_k^\dagger,\qquad\lambda=P_kv,\quad k=1,\cdots,T!\,, \tag{A.16}$$

where $P_1,\ldots,P_{T!}$ are $T\times T$ permutation matrices, and where $D_z$ is a diagonal matrix whose diagonal is given by a vector $z$. To assist in computing the gradients needed for a change of variables in the delta functions in (A.1), we represent $\Psi$ as

$$\Psi=\left(D_z+\tilde{\Psi}\right)P_k^\dagger,\quad k=1,\cdots,T!, \tag{A.17}$$

where $\tilde{\Psi}$ is zero along its diagonal. To a first order in $\tilde{\Psi}$, this gives

$$\Psi^\dagger\Psi-I\approx P_k\left(D_z^\dagger D_z-I+D_z^\dagger\tilde{\Psi}+\tilde{\Psi}^\dagger D_z\right)P_k^\dagger,$$

which we apply in (A.2) and (A.1) to obtain

$$p_\Psi\left(\left[D_z+\tilde{\Psi}\right]P_k^\dagger\right)=c\cdot\prod_{t=1}^{T}\delta\left(|z_t|^2-1\right)\prod_{s>t}\delta\left(z_s^*\tilde{\psi}_{st}+z_t\tilde{\psi}_{ts}^*\right). \tag{A.18}$$

To a first order in $\tilde{\Psi}$, we also have that

$$\begin{aligned} D_v+\tilde{R}-\Psi D_\lambda\Psi^\dagger &= D_v+\tilde{R}-\left(D_z+\tilde{\Psi}\right)P_k^\dagger D_\lambda P_k\left(D_z+\tilde{\Psi}\right)^\dagger \\ &\approx D_v-D_{P_k^\dagger\lambda}+\tilde{R}-D_zD_{P_k^\dagger\lambda}\tilde{\Psi}^\dagger-\tilde{\Psi}D_{P_k^\dagger\lambda}D_z^\dagger, \end{aligned}$$

where we have used $P_k^\dagger D_\lambda P_k=D_{P_k^\dagger\lambda}$, and which gives

$$\begin{aligned} &\delta\left(D_v+\tilde{R}-\Psi D_\lambda\Psi^\dagger\right) \\ =\ &\prod_{t=1}^{T}\delta\left(\lambda_t-[P_kv]_t\right)\cdot\prod_{s>t}\delta\left(\tilde{r}_{st}-z_sv_s\tilde{\psi}_{ts}^*-z_t^*v_t\tilde{\psi}_{st}\right)\cdot\prod_{t>s}\delta\left(z_sv_s\tilde{\psi}_{ts}^*+z_t^*v_t\tilde{\psi}_{st}\right) \end{aligned} \tag{A.19}$$

It's time to assemble the pieces. We substitute (A.17), (A.18) and (A.19) into the right-hand side of (A.14), and we replace the integral with respect to $\Psi$ by integrals with respect to $z$ and $\tilde{\Psi}$, and sum over the $T!$ values of $k$. We simultaneously integrate with respect to $\tilde{\psi}_{st}$ and $\tilde{\psi}_{ts}$ for $s>t$ by considering two coupled delta functions: one from (A.18), and one from (A.19), which give

$$\begin{aligned} z_s^*\tilde{\psi}_{st}+z_t\tilde{\psi}_{ts}^* &= 0, \quad s>t \\ z_sv_s\tilde{\psi}_{ts}^*+z_t^*v_t\tilde{\psi}_{st} &= 0, \quad t>s. \end{aligned} \tag{A.20}$$

We exchange $s$ and $t$ in the second equation and take its complex conjugate to obtain

$$\begin{bmatrix} z_s^* & z_t \\ z_t^*v_t^* & z_sv_s^* \end{bmatrix}\begin{bmatrix} \tilde{\psi}_{st} \\ \tilde{\psi}_{ts}^* \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{A.21}$$

The determinant of the matrix in this equation is $|z_s|^2v_s^*-|z_t|^2v_t^*$. To integrate over $\tilde{\psi}_{st}$ and $\tilde{\psi}_{ts}$ within the right-hand side of (A.14), we divide by the magnitude-squared of the determinant (squared because the variables are complex-valued), and we evaluate the integrand for $\tilde{\psi}_{st}=0$ and $\tilde{\psi}_{ts}=0$. As a result of these integrations, the variable $z_t$ appears only in the term $\delta\left(|z_t|^2-1\right)$, the integral of which is equal to $\pi$. After substituting (A.15), (A.18) and (A.19) into (A.14), and integrating over $\tilde{\Psi}$ and $z$, (A.14) becomes

$$\begin{aligned}
c \cdot \prod_{t=1}^{T} \delta\left(|v_t|^2 - 1\right) \prod_{s>t} \delta\left(\tilde{r}_{st}\right) &= c \cdot \pi^T \sum_{k=1}^{T!} \int d\lambda \, p_\lambda\left(\lambda\right) \prod_{t=1}^{T} \delta\left(\lambda_t - [P_k v]_t\right) \prod_{s>t} \frac{\delta\left(\tilde{r}_{st}\right)}{|v_s - v_t|^2} \\
&= c \cdot \pi^T \sum_{k=1}^{T!} p_\lambda\left(P_k v\right) \prod_{t=1}^{T} \prod_{s>t} \frac{\delta\left(\tilde{r}_{st}\right)}{|v_s - v_t|^2} \\
&= c \cdot \pi^T T! \, p_\lambda(v) \prod_{t=1}^{T} \prod_{s>t} \frac{\delta\left(\tilde{r}_{st}\right)}{|v_s - v_t|^2},
\end{aligned}$$

where we have used (A.6) in the final step. The eigenvalue probability density is obtained by inspection,

$$p_\lambda(\lambda) = \frac{1}{T! \, \pi^T} \prod_{t=1}^{T} \delta\left(|\lambda_t|^2 - 1\right) \prod_{s>t} |\lambda_s - \lambda_t|^2 . \tag{A.22}$$

We obtain the joint probability density for the phases of the eigenvalues $\theta$ by transforming coordinates with $\lambda_t = q_t^{1/2} e^{i\theta_t}$, whose Jacobian is equal to $1/2$. This gives

$$\begin{aligned}
p_{\theta,q}(\theta, q) &= \frac{1}{T! \, (2\pi)^T} \prod_{t=1}^{T} \delta\left(q_t - 1\right) \prod_{s>t} \left|e^{i\theta_s} - e^{i\theta_t}\right|^2 \\
&= \frac{1}{T! \, (2\pi)^T} \prod_{t=1}^{T} \delta\left(q_t - 1\right) \prod_{s>t} \left[4 \sin^2\left(\frac{\theta_s - \theta_t}{2}\right)\right],
\end{aligned} \tag{A.23}$$

or

$$p_\theta(\theta) = \frac{2^{T^2 - T}}{T! \, (2\pi)^T} \prod_{t=1}^{T} \prod_{s>t} \sin^2\left(\frac{\theta_s - \theta_t}{2}\right). \tag{A.24}$$

# B Appendix: Probability Density of Eigenvalues of $\ell$th Power of Isotropic Unitary Matrix

For a $T \times T$ isotropically random unitary matrix $\Phi$, and for any positive integer $\ell$, the eigenvector/eigenvalue decomposition is

$$\Phi^\ell = \Psi D_\lambda^\ell \Psi^\dagger,$$

where $\{\lambda_1, \cdots, \lambda_T\}$ are distributed according to (A.22). Let the eigenvalues of $\Phi^\ell$ be denoted by $\mu_t = \lambda_t^\ell,\ t = 1, \cdots, T$. Their conditional density is

$$\begin{aligned} p_{\mu|\lambda}(\mu \mid \lambda) &= \prod_{t=1}^{T} \delta\left(\mu_t - \lambda_t^\ell\right) \\ &= \prod_{t=1}^{T} \delta\left(\prod_{j_t=0}^{\ell-1}\left[\lambda_t - e^{i2\pi j_t/\ell}\mu_t^{1/\ell}\right]\right). \end{aligned} \tag{B.1}$$

We obtain the marginal density for $\mu$ by taking the expectation of the conditional density,

$$\begin{aligned} p_\mu(\mu) &= \int d\lambda\, p_\lambda(\lambda) \prod_{t=1}^{T} \delta\left(\prod_{j_t=0}^{\ell-1}\left[\lambda_t - e^{i2\pi j_t/\ell}\mu_t^{1/\ell}\right]\right) \\ &= \int d\lambda\, p_\lambda(\lambda) \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T}\left[\frac{\delta\left(\lambda_t - e^{i2\pi j_t/\ell}\mu_t^{1/\ell}\right)}{\prod_{k_t \neq j_t}\left|e^{i2\pi j_t/\ell} - e^{i2\pi k_t/\ell}\right|^2}\right]. \end{aligned} \tag{B.2}$$

The Jacobian factors that appear with the Dirac delta functions are evaluated as follows,

$$\begin{aligned} \prod_{k\neq j}\left|e^{i2\pi j/\ell} - e^{i2\pi k/\ell}\right| &= \prod_{k\neq j}\left|1 - e^{i2\pi(k-j)/\ell}\right| \\ &= \prod_{k=1}^{\ell-1}\left|1 - e^{i2\pi k/\ell}\right| \\ &= \prod_{k=1}^{\ell-1}\left|w - e^{i2\pi k/\ell}\right|_{w=1} \\ &= \left|\frac{w^\ell - 1}{w-1}\right|_{w=1} \\ &= \left|w^{\ell-1} + w^{\ell-2} + \cdots + 1\right|_{w=1} \\ &= \ell. \end{aligned} \tag{B.3}$$

The combination of (B.2),(B.3), and (A.22) gives

$$\begin{aligned} p(\mu) &= \frac{1}{\ell^{2T}} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} p_\lambda\left(e^{i2\pi j_1/\ell}\mu_1^{1/\ell}, \cdots, e^{i2\pi j_T/\ell}\mu_T^{1/\ell}\right) \\ &= \frac{1}{T!\pi^T\ell^{2T}} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \delta\left(|\mu_t|^{2/\ell} - 1\right) \prod_{s>t}\left|e^{i2\pi j_s/\ell}\mu_s^{1/\ell} - e^{i2\pi j_t/\ell}\mu_t^{1/\ell}\right|^2 \end{aligned} \tag{B.4}$$

The transformation (A.2) simplifies the Dirac delta functions to

$$\begin{aligned}
\delta\left(|\mu_t|^{2/\ell} - 1\right) &= \delta\left(\frac{|\mu_t|^2 - 1}{\ell}\right) \\
&= \ell\,\delta\left(|\mu_t|^2 - 1\right), \qquad \text{(B.5)}
\end{aligned}$$

which when substituted into (B.4) gives

$$p(\mu) = \frac{1}{T!\pi^T \ell^T} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \delta\left(|\mu_t|^2 - 1\right) \prod_{s>t} \left| e^{i2\pi j_s/\ell}\, \mu_s^{1/\ell} - e^{i2\pi j_t/\ell}\, \mu_t^{1/\ell} \right|^2 . \qquad \text{(B.6)}$$

The change-of-variables $\mu_t = y_t^{1/2} e^{i\gamma_t}$, combined with the appropriate Jacobian, gives the joint density for $y$ and $\gamma$. The integration over $y$ gives the joint density for the eigenvalue phases,

$$p(\gamma) = \frac{2^{T^2-T}}{T!(2\pi)^T \ell^T} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \prod_{s>t} \sin^2\left[\frac{\gamma_s - \gamma_t + 2\pi(j_s - j_t)}{2\ell}\right]. \qquad \text{(B.7)}$$

Merely by inspection, one cannot verify that this complicated expression differs from the density for $\ell = 1$. In the following we show that, for $\ell \geq T$ the eigenvalue phases are independent and uniformly distributed, and that a different density is obtained for every $\ell < T$.

### B.1 Eigenvalue density for $\ell \geq T$

It is more convenient to work with the characteristic function

$$\begin{aligned}
C(k) &= \mathrm{E}\left\{e^{ik^\dagger \gamma}\right\} \\
&= \mathrm{E}\left\{e^{i\ell k^\dagger \theta}\right\} \\
&= \frac{1}{T!(2\pi)^T} \int d\theta\, e^{i\ell k^\dagger \theta} \prod_t \prod_{s>t} \left|e^{i\theta_s} - e^{i\theta_t}\right|^2, \qquad \text{(B.8)}
\end{aligned}$$

which, because of the periodicity of the density, need only be considered for integer values of the components of the vector $k$. For all $\ell \geq T$ we wish to show that $C(k) = 0$ for all $k$ not identically equal to 0. Without loss of generality, suppose that $k_1 \neq 0$. Then the integrand of (B.8) may be regarded as a trigonometric polynomial in $\theta_1$, and the integral with respect to $\theta_1$ takes the form

$$\int d\theta_1 \, e^{i\ell k_1 \theta_1} \sum_{j=-(T-1)}^{(T-1)} a_j e^{i\, j\, \theta_1}.$$

Since $\ell \geq T$ and $|j| < T$, for all nonzero integer values of $k_1$, $\ell k_1 + j \neq 0$, so the integral vanishes. It follows that $C(k) = 0$, unless $k \equiv 0$, so the eigenvalues phases $\gamma$ are independent and uniformly distributed.

## B.2 Eigenvalue phases are not independent for $\ell < T$

The converse also holds: for all $1 \leq \ell < T$, the eigenvalue phases are *not* independent and uniformly distributed. We need only show that $C(k) \neq 0$ for some nonzero value of $k$. In the following we establish this fact for $k_1 = -1$, $k_2 = 1$, $k_3 = k_4 = \cdots = k_T = 0$. We use the fact that the probability density for $\theta$ is proportional to the magnitude squared of a Vandermonde determinant, i.e.,

$$\begin{aligned} \prod_t \prod_{s>t} \left| e^{i\theta_s} - e^{i\theta_t} \right|^2 &= \left| \det \begin{bmatrix} 1 & & 1 \\ e^{i\theta_1} & & e^{i\theta_T} \\ \vdots & \cdots & \vdots \\ e^{i(T-1)\theta_1} & & e^{i(T-1)\theta_T} \end{bmatrix} \right|^2 \\ &= \left| \sum_q a_q e^{iq^\dagger \theta} \right|^2, \end{aligned}$$

where the $T$ components of $q$ range over all permutations of $\{0, 1, \cdots, T-1\}$, and $a_q = \pm 1$ depending on whether the permutation is even or odd. We substitute this expression into (B.8) to give

$$C(-1, 1, 0, \cdots, 0) = \frac{1}{T!\,(2\pi)^T} \int d\theta \, e^{i\ell(-\theta_1 + \theta_2)} \sum_q \sum_r a_q \, a_r \, e^{i(q-r)^\dagger \theta}. \tag{B.9}$$

It is apparent that the nonvanishing terms of (B.9) satisfy the following:

$$\begin{aligned} -\ell + q_1 - r_1 &= 0 \\ \ell + q_2 - r_2 &= 0 \\ q_t - r_t &= 0, \; t = 3, \cdots, T. \end{aligned}$$

We substitute the equivalent expressions for $\{r_t\}$ into (B.9) to give

$$C(-1,1,0,\cdots,0) = \frac{1}{T!}\sum_{q_1=\ell}^{T-1}\sum_{q_2=0}^{T-1-\ell}\sum_{q_3}\cdots\sum_{q_T} a_{q_1,q_2,q_3,\cdots,q_T}a_{q_1-\ell,q_2+\ell,q_3,\cdots,q_T}.$$

Recall that the $a$'s are nonzero only when their $T$ arguments are distinct. Given the last $T-2$ arguments of the $a$'s, there are only two possible values for the first two arguments. Either $q_1 = q_1 - \ell$ and $q_2 = q_2 + \ell$ (an impossibility since $\ell \neq 0$) or $q_2 = q_1 - \ell$, which corresponds to exchanging the first two arguments of the $a$'s. Hence

$$\begin{aligned} a_{q_1,q_2,q_3,\cdots,q_T}a_{q_1-\ell,q_2+\ell,q_3,\cdots,q_T} &= a_{q_1,q_1-\ell,q_3,\cdots,q_T}a_{q_1-\ell,q_1,q_3,\cdots,q_T} \\ &= -1. \end{aligned}$$

We apply this result, and sum over the $(T-2)!$ values of $\{q_3,\cdots,q_T\}$ to obtain

$$\begin{aligned} C(-1,1,0,\cdots,0) &= \frac{(T-2)!}{T!}\sum_{q_1=\ell}^{T-1}(-1) \\ &= -\frac{T-\ell}{T(T-1)} \\ &\neq 0, \quad \forall\, \ell = 1,\cdots,T-1, \end{aligned} \tag{B.10}$$

Therefore we conclude that, for $\ell = 1,\cdots,T-1$, the eigenvalue phases are *not* independent and uniformly distributed, and furthermore that the eigenvalue density is different for every $\ell = 1,\cdots,T-1$.

This yields the following lemma.

**Lemma 1** *The eigenvalue phases of the $\ell$-th power of a $T \times T$ isotropically distributed unitary matrix have a uniform joint distribution if and only if $\ell \geq T$.*

# References


[1] I. C. Abou Faycal, M. D. Trott and S. Shamai (Shitz), "The Capacity of Discrete-Time Rayleigh Fading Channels," in *Proc. Int. Symp. Info. Th.*, Ulm, Germany, 1997, p. 473.

[2] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading Channels: Information-Theoretic and Communications Aspects", *IEEE Trans. Info. Theory*, vol. 44, no. 6, pp. 2619–2992, October 1998.

[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] V.L. Girko, *An Introduction to Statistical Analysis of Random Arrays*, VSP International Science Pub., 1998.

[5] B. M. Hochwald and T. L. Marzetta, "Unitary Space-Time Modulation for Multiple-Antenna Communications in Rayleigh Flat Fading", *IEEE Trans. Info. Theory*, vol. 46, no. 2, pp. 543–564, March 2000.

[6] B. M. Hochwald, T. L. Marzetta, and B. Hassibi, "Space-Time Autocoding", *submitted to IEEE Trans. Info. Theory; see also http://mars.bell-labs.com*.

[7] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic Design of Unitary Space-Time Constellations", *to appear in IEEE Trans. Info. Theory, Sept. 2000; see also http://mars.bell-labs.com*.

[8] T. L. Marzetta, "BLAST Training: Estimating Channel Characteristics for High-Capacity Space-Time Wireless", *Proc. 37th Annual Allerton Conference on Communications, Control, and Computing*, pp. 958–966, Monticello, IL, Sept. 22–24, 1999.

[9] 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, no. 1, pp. 139–157, January 1999.

[10] L. H. Ozarow, S. Shamai (Shitz) and A. D. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio", *IEEE Trans. Veh. Tech.*, vol. 43, pp. 359–378, May 1994.

[11] G. Strang, *Linear Algebra and its Applications*, third edition, Harcourt Brace Jovanovich, Inc., San Diego, 1988.

[12] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Analysis and Code Construction", *IEEE Trans. Info. Theory*, vol 44, no. 2, pp. 744–765, March 1998.

[13] I. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels, *European Transactions on Telecommunications*, Vol. 10, No. 6, Nov. 1999; also *AT&T Bell Laboratories Technical Memo*, 1995.

[14] L. Zheng and D. N. C. Tse, "Packing Spheres in the Grassman Manifold: A Geometric Approach to the Noncoherent Multi-Antenna Channel", *Proc. 37th Annual Allerton Conference on Communications, Control, and Computing*, pp. 861–870, Monticello, IL, Sept. 22-24, 1999.