symbol duration of the coherence interval. We have established the correctness of the conjecture for three special cases: 1) a coherence interval comprising one symbol, 2) a coherence interval of unlimited duration, and 3) an unlimited SNR.

As mentioned in the Introduction, there is an interesting analogy between multiple user, multiple antenna links, and coded DS-CDMA with random signatures. In the regime considered here, the signatures are chosen in an absolutely random fashion for every T-symbol coherence interval, without their values being known by any of the receivers, whether legitimate or not. It is not surprising that our results also conform with the autocoding capacity [7], [20], where, for the case of unknown CSI, the number of actively operating users at each time instant is small compared with T.

Operating at the multiple-user mode forces the signals corresponding to different users to be independent, yet all users can pre-coordinate powers, maintain the average power constraint. (TDMA is a simple example). We conjecture that operating with the optimized number of the transmitting users that maximizes throughput, optimization of such power allocation sharing (mixed strategy) would not further increase the throughput. Evidently, in case that the optimized number of transmitting users M_* is smaller than the actual number of users M, "fairness" could be imposed at no penalty in overall throughput, by permitting different sets of M_* out of M users to access the channel in some uniform preassigned order.

References

[1] I. C. Abou Faycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1290–1301, May 2001.

[2] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2992, Oct. 1998.

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

[4] 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.

[5] S. V. Hanly and D. N. Tse, "The multi-access fading channel: Shannon and delay limited capacities," in *Proc. 33d Allerton Conf.*, Monticello, IL, 1995, pp. 786–795.

[6] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links," *IEEE Trans. Inform. Theory*, submitted for publication.

[7] B. M. Hochwald, T. L. Marzetta, and B. Hassibi, "Space–time autocoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2761–2781, Nov. 2001.

[8] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. Int. Conf. Communications, ICC'95*, Seattle, WA, USA, June 18–22, 1995, pp. 331–335.

[9] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high-capacity space–time wireless," in *Proc. 37th Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sept. 22–24, 1999, pp. 958–966.

[10] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[11] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.

[12] S. Shamai (Shitz) and E. Telatar, "Some information theoretic aspects of decentralized power control in multiple access fading channels," in *Proc. Abstracts, 1999 Information Theory and Networking Workshop*, Metsovo, Greece, June 27–July 1, 1999.

[13] S. Shamai (Shitz) and S. Verdú, "The effect of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1302–1327, May 2001.

[14] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun. (ETT)*, vol. 10, no. 6, pp. 585–596, Nov./Dec. 1999.

[15] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.

[16] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1984–1991, Sept. 1999.

[17] W. Yu, W. Rhee, and J. M. Cioffi, "Multiuser transmitter optimization for vector multiple access channels," preprint, Mar. 3, 2000.

[18] L. Zheng and D. N. C. Tse, "Packing spheres in the Grassmann manifold: A geometric approach to the noncoherent multi-antenna channel," preprint, Apr. 30, 2000. See also *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 25–30, 2000, p. 364.

[19] P. Billingsley, *Convergence of Probability Measures*, 2nd ed. New York: Wiley-Interscience, 1999.

[20] T. L. Marzetta, B. M. Hochwald, and B. Hassibi, "Space–time autocoding: Arbitrarily reliable communication in a single fading interval," in *Proc. 2000 IEEE Int. Symp. Information Theory*, Sorrento, Italy, June, 25–30 2000, p. 313.

Structured Unitary Space–Time Autocoding Constellations

Thomas L. Marzetta, *Senior Member, IEEE*, Babak Hassibi, and Bertrand M. Hochwald

***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 independent isotropically random unitary space–time signals theoretically supports transmission rates that are a significant fraction of autocapacity with an extremely low probability of error. The exploitation of space–time autocoding requires the creation and decoding of extraordinarily large constellations—typically $L = 2^{80}$. In this correspondence, we make progress on the first part of the problem through 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. We establish the limitations of an earlier construction through 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 the sole exception of the one-by-one unitary matrix.**

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

I. Introduction

Research in multiple-antenna wireless entered an exciting phase with [3], [15], which predict spectacular capacities (both Shannon

Manuscript received June 13, 2000; revised June 16, 2001.

T. L. Marzetta and B. M. Hochwald are with the Mathematical Sciences Research Center, Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: tlm@research.bell-labs.com; hochwald@research.bell-labs.com).

B. Hassibi was with the Mathematical Sciences Research Center, Bell Laboratories, Lucent Technologies. He is now with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: hassibi@systems.caltech.edu).

Communicated by M. L. Honig, Associate Editor for Communications.

Publisher Item Identifier S 0018-9448(02)02001-1.

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 [14], 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 [9], and for many potential applications, training is an acceptable burden. However, in time-division multiple-access (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 [9].

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], [10], [6], [16], based on a piecewise-constant model for fading (also called block fading [12], [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 [10] 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 an $M \times M$ diagonal, real, nonnegative matrix. This structure motivates the use of *unitary space–time modulation* [6] involving a constellation of L $T \times M$ unitary matrices $\{\Phi_0, \ldots, \Phi_{L-1}\}$, 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 [6] using a simple iterative algorithm. A systematic approach is pursued in [8], 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 Lth roots of unity, and with the initial signal comprising M columns from a $T \times T$ discrete Fourier transform (DFT) 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 [7] 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 signal-to-noise ratio (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 16×7 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 II reviews the signal model, unitary space–time modulation, space–time autocoding, and the earlier systematic construction. Section III explains the new construction. Section IV reinterprets earlier systematic constructions such as [8] in light of this new construction. The mathematical results required for many of the conclusions of this correspondence are developed in the appendxes: Appendix A reviews the isotropically random unitary matrix, and presents some unusual operations involving Dirac delta functions. 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.

II. 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.

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

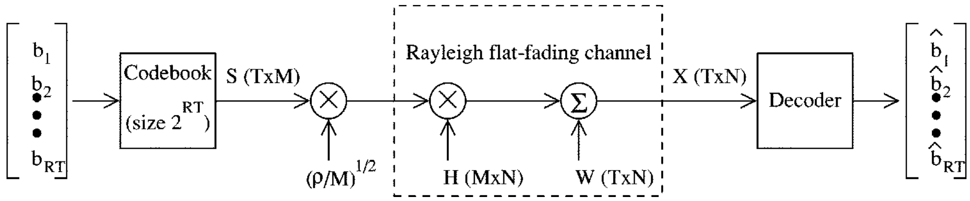

Fig. 1. Wireless link comprising M transmitter and N receiver antennas. We wish to transmit $R \cdot T$ bits of information reliably in a *single* coherence interval T, where R is the rate in bits per symbol.

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 ρ represents the expected SNR at each receive antenna.

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

B. 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 Fig. 1. There is a positive autocapacity C_a [7], 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 per 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 III describes the isotropic distribution in some detail) $T \times M$ unitary space–time signals [7], $\{S_\ell = \sqrt{T}\Phi_\ell, \ell = 0, \dots, L-1\}$, where $L = 2^{RT}$, and where the column vectors of each Φ_ℓ are orthonormal, $\Phi_\ell^\dagger \Phi_\ell = I_M$. The block probability of error P_e may be upper-bounded through the union bound

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

where $P_e\{\Phi_1 \text{ versus } \Phi_2\}$ is the pairwise (e.g., two-signal constellation) probability of error that is associated with any distinct pair of signals in the constellation [6], given by the exact formula

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

where $d_1, \dots, d_M$ 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; θ can be integrated numerically. We omit all the details and refer the interested reader to [7].

Fig. 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 we need constellations of unprecedented size ($L = 2^{80}$ for $T = 16$, and $R = 5$).

C. Earlier Systematic Constructions [8]

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 [6], [8]. 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, $\Phi_\ell \to \Psi^\dagger \Phi_\ell, \ell = 0, \dots, L-1$; 2) right multiplication by individual $M \times M$ unitary matrices, $\Phi_\ell \to \Phi_\ell \Upsilon_\ell, \ell = 0, \dots, L-1$. 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 [8] with the goal of achieving a low block probability of error. The construction proposed in [8] involves successive rotations of an initial signal in T-dimensional complex space

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

where Ω is a $T \times T$ unitary matrix, and Φ_0 is the $T \times M$ initial signal obeying $\Phi_0^\dagger \Phi_0 = I_M$. A judicious choice of Ω and Φ_0 is needed to make the columns of Φ_ℓ zig-zag over the surface of the T-dimensional complex sphere.

There is no loss of generality in assuming that Ω is diagonal because the Schur factorization [13] implies that any square unitary matrix Ω has the eigenvector–eigenvalue decomposition

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

where Ψ is $T \times T$ unitary, and Θ is a $T \times T$ diagonal matrix of eigenvalues of Ω. 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 Ω

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

In [8], further structure is imposed by choosing a) Θ to be an Lth root of the identity matrix, implying that $\Theta_{tt} = e^{i2\pi u_t/L}$, where $u_t \in \{0, 1, \dots, L-1\}$; b) Φ_0 to comprise M distinct columns of a $T \times T$ DFT matrix. The integers $u_1, \dots, u_T$ 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 Θ to be an Lth 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

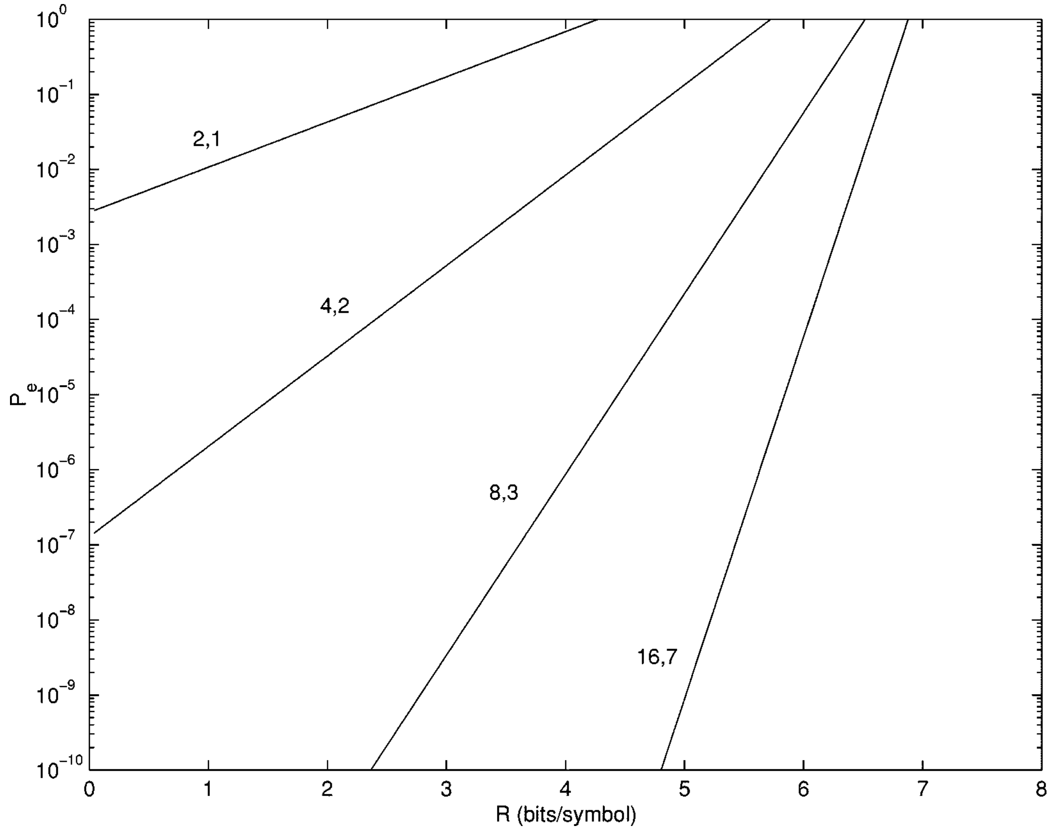

Fig. 2. Reference [7] upper bound on block probability of error versus transmission rate (bits/symbol) for random codebook of unitary space–time signals, for $N = 4$, $\rho = 18$ dB, and $(T, M) = (2, 1), (4, 2), (8, 3), (16, 7)$. Autocapacity is equal to 24.01 bits/symbol.

have to be checked, and it has some intuitive appeal. However, no indication is given in [8] 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.

III. 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 Φ_ℓ to be pairwise independent isotropically distributed matrices for all $\ell' \neq \ell$.[1] 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 indexes, and they are generated as follows:

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

[1]In fact, the random coding exponent only depends on pairwise independence [4].

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 Ω is isotropically distributed if its probability density is unchanged when Ω is premultiplied by any $T \times T$ deterministic unitary matrix. From this definition, one may deduce [10] that a) there is exactly one probability density that possesses this property, with the formula given by (A1); b) the density is invariant to postmultiplication of Ω by any deterministic unitary matrix. Likewise, a $T \times M$ random unitary matrix Φ, 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 ℓ denote the vector comprising the RT binary indexes $\{\ell_1, \ldots, \ell_{RT}\}$, and consider two distinct signals from the constellation (9), Φ_ℓ and Φ_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

$$\Phi_\ell = \Xi_\ell\Phi_0 \quad \text{and} \quad \Phi_k = \Xi_k\Phi_0$$

where Ξ_ℓ and Ξ_k are products of certain subsets of $\{\Omega_1, \ldots, \Omega_{RT}\}$. At least one of Ξ_ℓ or Ξ_k has a factor Ω_q that is not contained in the other. Assume without loss of generality that Ω_q is a factor of Ξ_ℓ (but not of Ξ_k) which then takes the form

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

Consider the following argument.

- Ω_q does not appear in A, B, or Ξ_k, and therefore Ω_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 Ξ_k.
- By a similar argument, $\Xi_\ell = A\Omega_q B$ is isotropic and independent of B and Ξ_k.
- Finally, because $\Phi_0^\dagger \Phi_0 = I$, conditioned on Φ_0, $\Phi_\ell = \Xi_\ell \Phi_0$ is $T \times M$ isotropically distributed, and since the conditional density for Φ_ℓ has no dependence on either Φ_0 or on Ξ_k, we conclude that Φ_ℓ 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 Φ_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 Φ_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 Ψ. Likewise, $\Psi^\dagger \Omega_j \Psi$ is isotropically distributed and independent of Ψ. Therefore, the constellation (9) with Φ_0 isotropically distributed is equivalent to a constellation where $\Phi_0^\dagger = [I_M \quad 0]$.

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.

IV. Comments on Earlier Systematic Construction of [8]

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.

A. Single Isotropically Random Rotation Matrix

Consider the construction (6), with the initial signal Φ_0 isotropically random unitary, and with the rotation matrix Ω independent isotropically random unitary. Because Φ_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 Ω^ℓ, $\ell = 2, \ldots, L-1$ 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 Ω is isotropically distributed, then Ω^ℓ is isotropically distributed for any integer $\ell \geq 2$. Indeed, if this were true, the successive rotations $\Omega^\ell \Phi_0$ 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). The eigenvectors and the eigenvalues of an isotropic matrix are independent of each other, the $T \times T$ eigenvector matrix Ψ is itself isotropically distributed, and the T eigenvalues $\lambda_1, \ldots, \lambda_T$ have the density [5]

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

The eigenvalues have unit magnitude, while the phases θ_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 Ω^ℓ, 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

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

For all $T > 1$, Ω^ℓ is *not* isotropically distributed for any $\ell \geq 2$. Furthermore, Ω^ℓ reaches a limiting density for all $\ell \geq T$, where the eigenvalue phases are independent and uniformly distributed

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

See [11] or, alternatively, Appendix B for derivations of these results.

These results are highly counterintuitive. For example, the ℓth power of a 3×3 real isotropically distributed orthogonal matrix is not isotropic, as illustrated in Fig. 3. Let Ψ be a 3×3 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 Ψe_x is equally likely to lie anywhere on the unit sphere. However, the product $\Psi^\ell e_x$, for even powers of ℓ is biased toward e_x, and for odd powers of ℓ is biased toward $\pm e_x$.

For two signals Φ_ℓ and Φ_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 large ($L = 2^{80}$), most of the pairwise signals have an effective $\Omega^{k-\ell}$ with this phase distribution, even if Ω is itself isotropic. We, therefore, look briefly at the performance of a constellation that is generated by a single rotation matrix Ω 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).

B. Rotation Matrix With Independent, Uniform-Phase Eigenvalues

Consider the construction (6) where Φ_0 is isotropic but Ω 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 Fig. 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 Fig. 2 shows that the single-rotation matrix construction is worse than the construction (9).

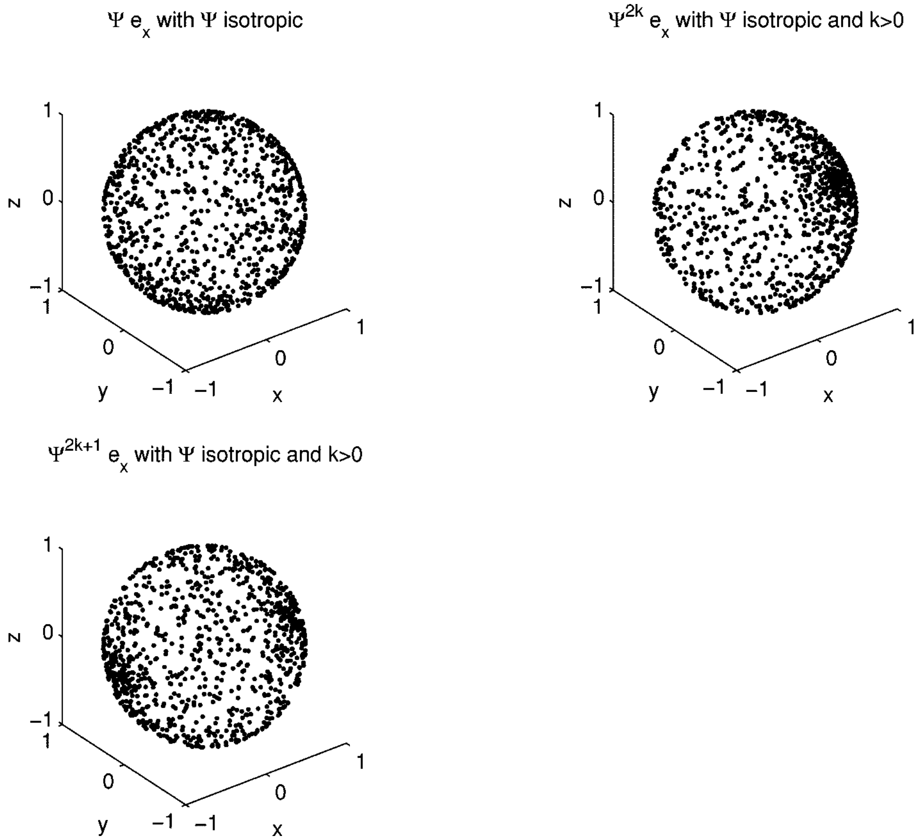

Fig. 3. The power of a 3×3 real, isotropically distributed orthogonal matrix is not isotropically distributed. Upper left: single application of random rotation to the unit vector e_x results in a unit vector that is equally likely to lie anywhere on the unit sphere. Upper right: an even number of applications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

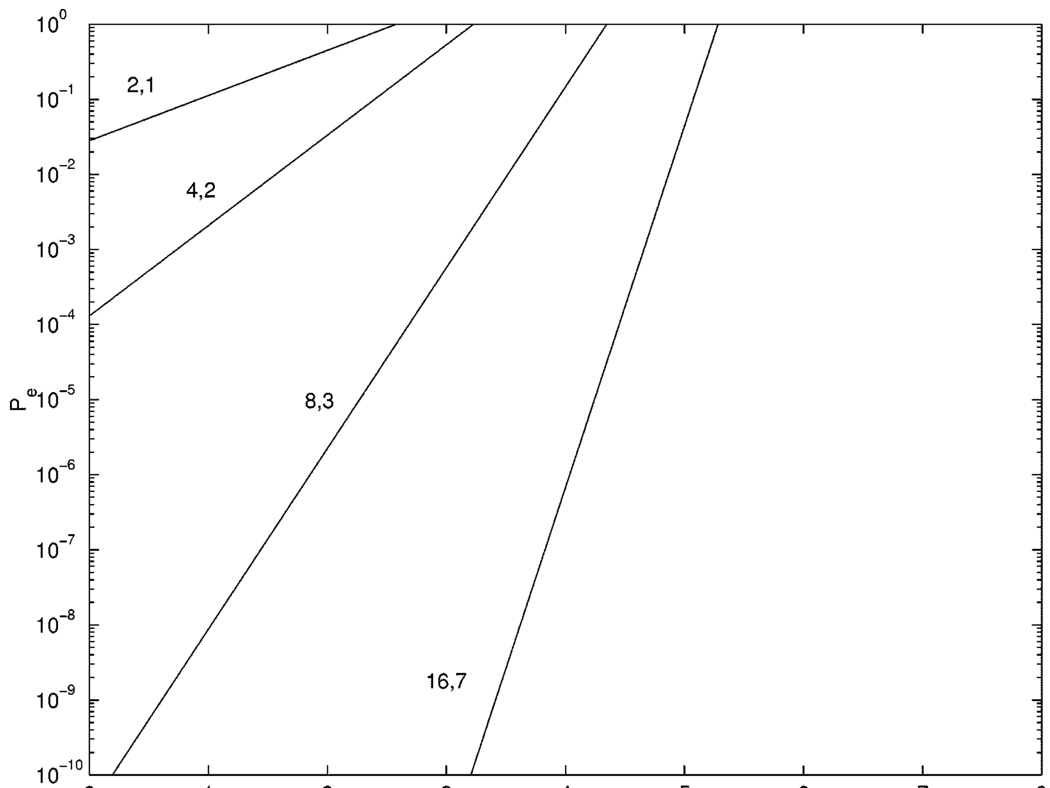

Fig. 4. Upper bound on block probability of error versus transmission rate (bits/symbol) for codebook of unitary space–time signals, generated from a single rotation matrix having independent, uniform-phase eigenvalues, for $N = 4$, $\rho = 18$ dB, and $(T, M) = (2, 1), (4, 2), (8, 3), (16, 7)$.

V. Conclusion

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. 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.

Appendix A
The Isotropically Random Unitary Matrix; Some Unusual Operations Involving Dirac Delta Functions

A $T \times T$ random unitary matrix Φ is isotropically distributed if its probability density is unchanged when Φ is pre-multiplied by any $T \times T$ deterministic unitary matrix. This operational definition leads directly to the unique probability density [10] as a function of the T column vectors $\{\phi_1, \ldots, \phi_T\}$

$$\begin{aligned} p(\Phi) &= p(\phi_1) \cdot \prod_{s=2}^{T} p(\phi_s | \phi_1, \ldots, \phi_{s-1}) \\ &= \left[\frac{\Gamma(T)}{\pi^T} \delta\left(\phi_1^\dagger \phi_1 - 1\right) \right] \\ &\quad \cdot \prod_{s=2}^{T} \left[\frac{\Gamma(T+1-s)}{\pi^{T+1-s}} \, \delta\left(\phi_s^\dagger \phi_s - 1\right) \prod_{t=1}^{s-1} \delta\left(\phi_s^\dagger \phi_t\right) \right] \end{aligned} \tag{A1}$$

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 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

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

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

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

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

$$\delta(f(x)) = \frac{\delta(x - x_0)}{|\det\{\nabla f^T(x)\}|_{x=x_0}}. \tag{A2}$$

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.

Appendix B
Probability Density of Eigenvalues of ℓth Power of Isotropic Unitary Matrix

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

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

where $\{\lambda_1, \ldots, \lambda_T\}$ are distributed according to (10). Let the eigenvalues of Φ^ℓ be denoted by $\mu_t = \lambda_t^\ell, \ t = 1, \ldots, T$. Their conditional density is

$$\begin{aligned} p_{\mu|\lambda}(\mu|\lambda) &= \prod_{t=1}^{T} \delta(\mu_t - \lambda_t^\ell) \\ &= \prod_{t=1}^{T} \delta\left(\prod_{j_t=0}^{\ell-1} \left[\lambda_t - e^{i2\pi j_t/\ell} \mu_t^{1/\ell} \right] \right). \end{aligned} \tag{B1}$$

We obtain the marginal density for μ by taking the expectation of the conditional density

$$\begin{aligned} p_\mu(\mu) &= \int d\lambda p_\lambda(\lambda) \prod_{t=1}^{T} \delta\left(\prod_{j_t=0}^{\ell-1} \left[\lambda_t - e^{i2\pi j_t/\ell} \mu_t^{1/\ell} \right] \right) \\ &= \int d\lambda p_\lambda(\lambda) \sum_{j_1=0}^{\ell-1} \cdots \\ &\qquad \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \left[\frac{\delta\left(\lambda_t - e^{i2\pi j_t/\ell} \mu_t^{1/\ell}\right)}{\prod_{k_t \neq j_t} |e^{i2\pi j_t/\ell} - e^{i2\pi k_t/\ell}|^2} \right]. \end{aligned} \tag{B2}$$

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

$$\begin{aligned} \prod_{k \neq j} \left| e^{i2\pi j/\ell} - e^{i2\pi k/\ell} \right| &= \prod_{k \neq j} \left| 1 - e^{i2\pi(k-j)/\ell} \right| \\ &= \prod_{k=1}^{\ell-1} \left| 1 - e^{i2\pi k/\ell} \right| \\ &= \prod_{k=1}^{\ell-1} \left| w - e^{i2\pi k/\ell} \right|_{w=1} \\ &= \left| \frac{w^\ell - 1}{w - 1} \right|_{w=1} \\ &= \left| w^{\ell-1} + w^{\ell-2} + \cdots + 1 \right|_{w=1} \\ &= \ell. \end{aligned} \tag{B3}$$

The combination of (B2), (B3), and (10) gives

$$\begin{aligned} p(\mu) &= \frac{1}{\ell^{2T}} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} p_\lambda\left(e^{i2\pi j_1/\ell} \mu_1^{1/\ell}, \ldots, e^{i2\pi j_T/\ell} \mu_T^{1/\ell} \right) \\ &= \frac{1}{T! \pi^T \ell^{2T}} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \delta(|\mu_t|^{2/\ell} - 1) \\ &\quad \cdot \prod_{s>t} \left| e^{i2\pi j_s/\ell} \mu_s^{1/\ell} - e^{i2\pi j_t/\ell} \mu_t^{1/\ell} \right|^2. \end{aligned} \tag{B4}$$

The transformation (A2) simplifies the Dirac delta functions to

$$\begin{aligned} \delta(|\mu_t|^{2/\ell} - 1) &= \delta\left(\frac{|\mu_t|^2 - 1}{\ell} \right) \\ &= \ell \delta(|\mu_t|^2 - 1) \end{aligned} \tag{B5}$$

which when substituted into (B4) gives

$$p(\mu) = \frac{1}{T!\pi^T \ell^T} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \delta(|\mu_t|^2 - 1) \prod_{s>t} \left| e^{i2\pi j_s/\ell \mu_s^{1/\ell}} - e^{i2\pi j_t/\ell \mu_t^{1/\ell}} \right|^2. \quad \text{(B6)}$$

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 γ. The integration over y gives the joint density for the eigenvalue phases

$$p(\gamma) = \frac{2^{T^2-T}}{T!(2\pi)^T \ell^T} \sum_{j_1=0}^{\ell-1} \cdots \sum_{j_T=0}^{\ell-1} \prod_{t=1}^{T} \prod_{s>t} \sin^2 \left[\frac{\gamma_s - \gamma_t + 2\pi(j_s - j_t)}{2\ell} \right]. \quad \text{(B7)}$$

Merely by inspection, one cannot verify that this complicated expression differs from the density for $\ell = 1$, though one could establish this fact numerically. 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 $\ell < T$. These results have also been established in [11]; our treatment is more direct, and it avoids using the Haar measure.

A. Eigenvalue Density for $\ell \geq T$

It is more convenient to work with the characteristic function

$$\begin{aligned} C(k) &= \mathrm{E}\left\{ e^{ik^\dagger \gamma} \right\} \\ &= \mathrm{E}\left\{ e^{i\ell k^\dagger \theta} \right\} \\ &= \frac{1}{T!(2\pi)^T} \int d\theta\, e^{i\ell k^\dagger \theta} \prod_t \prod_{s>t} \left| e^{i\theta_s} - e^{i\theta_t} \right|^2 \end{aligned} \quad \text{(B8)}$$

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 (B8) may be regarded as a trigonometric polynomial in θ_1, and the integral with respect to θ_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 γ are independent and uniformly distributed.

B. 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 θ 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 $\{0, 1, \ldots, T-1\}$, and $a_q = \pm 1$ depending on whether the permutation is even or odd. We substitute this expression into (B8) to give

$$\begin{aligned} &C(-1, 1, 0, \ldots, 0) \\ &\quad = \frac{1}{T!(2\pi)^T} \int d\theta\, e^{i\ell(-\theta_1+\theta_2)} \sum_q \sum_r a_q a_r e^{i(q-r)^\dagger \theta}. \end{aligned} \quad \text{(B9)}$$

It is apparent that the nonvanishing terms of (B9) satisfy the following:

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

We substitute the equivalent expressions for $\{r_t\}$ into (B9) to give

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

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,

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

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

$$\begin{aligned} C(-1, 1, 0, \ldots, 0) &= \frac{(T-2)!}{T!} \sum_{q_1=\ell}^{T-1} (-1) \\ &= -\frac{T-\ell}{T(T-1)} \\ &\neq 0, \qquad \forall \ell = 1, \ldots, T-1. \end{aligned} \quad \text{(B10)}$$

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

This yields the following lemma.

Lemma 1: The eigenvalue phases of the ℓth 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. Information Theory*, Ulm, Germany, 1997, p. 473.

[2] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2992, Oct. 1998.

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

[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968, sec. 6.2.

[5] V. L. Girko, *An Introduction to Statistical Analysis of Random Arrays*. Utrecht, The Netherlands: VSP Int. Sci. Pub., 1998.

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

[7] B. M. Hochwald, T. L. Marzetta, and B. Hassibi, "Space–time autocoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2761–2781, Nov. 2001.

[8] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space–time constellations," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1962–1973, Sept. 2000.

[9] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high-capacity space–time wireless," in *Proc. 37th Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sept. 22–24, 1999, pp. 958–966.

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

[11] E. M. Rains, "High powers of random elements of compact lie groups," *Probab. Theory Relat. Fields*, vol. 107, pp. 219–241, 1997.

[12] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.

[13] G. Strang, *Linear Algebra and Its Applications*, 3rd ed. San Diego, CA: Harcourt Brace Jovanovich, 1988.

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

[15] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, Nov. 1999. Also AT&T Bell Labs. Tech. Memo, 1995.

[16] L. Zheng and D. N. C. Tse, "Packing spheres in the Grassman manifold: A geometric approach to the noncoherent multi-antenna channel," in *Proc. 37th Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sept. 22–24, 1999, pp. 861–870.

A New View of Performance Analysis of Transmit Diversity Schemes in Correlated Rayleigh Fading

Siwaruk Siwamogsatham, *Student Member, IEEE*, Michael P. Fitz, *Member, IEEE*, and Jimm H. Grimm, *Member, IEEE*

***Abstract*—This correspondence provides a new formulation for the pairwise error probability for any coherently demodulated system in arbitrarily correlated Rayleigh fading. The novelty of the result is that the error probability expression can be described as a function of the eigenvalues of a "signal"-only matrix. We also provide the relationship between the pole location of the characteristic function and the resulting error probability. This result allows us to approximate and bound the desired probability. A new simple bound on the pairwise error probability is derived that is better than the standard Chernoff bound and asymptotically tight with the signal-to-noise ratio (SNR) to the true probability.**

***Index Terms*—Asymptotic bounds, diversity, pairwise error probability, performance analysis, Rayleigh fading.**

I. Introduction

Performance analysis of digital communications systems in fading channels has been an area of long-time interest. Results were first obtained in [1]. An elegant unified technique was presented in [2], [3] for signals experiencing complex Gaussian fading and advanced textbooks have significant sections discussing related results [4]. Improvements and generalizations have continued to be derived until today (e.g., [5], [6]).

In this correspondence, we present some new results on performance analysis of transmit diversity schemes or a generic coherent demodulation scheme in independent and correlated Rayleigh-fading channels. Unlike the traditional analysis [7] which derived the performance as a function of the eigenvalues of a certain matrix that is composed of signal and noise components, here we are able to describe the pairwise error probability (PWEP) as a function of the eigenvalues of a signal-only matrix by employing an appropriate matrix identity. This result provides a nice intuition to the current problem as it allows us to describe the roots of the associated characteristic function as a function of signal and noise power. We also derive the relationship between the pole location of the characteristic function and the resulting probability density function (pdf) of a generic quadratic form of a zero-mean complex Gaussian random vector (CGRV). The result suggests that any movement of a given pole toward the origin always produces a larger error probability, and hence the desired probability can be lower- and upper-bounded or approximated by appropriately moving the poles of the characteristic function. Since we can explicitly describe these poles as a function of signal and noise power, this result directly leads to a new upper bound of the PWEP that is better than the standard Chernoff bound and asymptotically tight with SNR to the true PWEP. In addition, this bound often can be put in a product form and used with a transfer function union bound [8].

The correspondence is organized as follows. Section II provides a formulation of a generic quadratic form of a zero-mean CGRV and derives a useful property on the pdf of this quadratic form. Section III gives the description of the system under consideration. Section IV derives the PWEP for the system of interest in terms of the exact calculation, a simple approximation, and an asymptotic bound.

II. Quadratic Forms of a CGRV

In digital communications, performance analysis generally involves the evaluation of the probability distribution of a generic quadratic form of a CGRV. With an appropriate formulation of a quadratic form of CGRV, the analysis is unified for all varieties of applications (e.g., see [7], [4], [9], [10]). In this section, we formally describe the pdf of the quadratic form of a CGRV, and derive a useful property on the density function which will lead to later results of interest.

A quadratic form of an $N \times 1$ CGRV $\vec{Z}$ is a real-valued random variable given as

$$Q_z = \vec{Z}^H \boldsymbol{K} \vec{Z} \tag{1}$$

where $\boldsymbol{K}$ is a certain $N \times N$ Hermitian matrix. The form in (1) often arises in performance analysis of digital communication systems and a wide variety of results exist to characterize $\vec{Z}$ and $\boldsymbol{K}$ for different applications. For a zero-mean CGRV $\vec{Z}$, the characteristic function (ChF) of Q_z is given by [11]

$$\Phi_{Q_z}(t) = E\left[e^{jQ_z t}\right] = \frac{1}{\det(\boldsymbol{I}_N - jt\boldsymbol{C}_z\boldsymbol{K})} \tag{2}$$

where $\boldsymbol{C}_z$ is the covariance matrix of $\vec{Z}$ and $\boldsymbol{I}_N$ is an $N \times N$ identity matrix. Defining $\boldsymbol{W}_z = \boldsymbol{C}_z\boldsymbol{K}$, $\Phi_{Q_z}(t)$ can be expressed as a function of the nonzero eigenvalues of $\boldsymbol{W}_z$. After some linear algebra manipulation, we may express $\Phi_{Q_z}(t)$ in the form

$$\Phi_{Q_z}(t) = \frac{1}{\prod_{i=1}^{N^{(w)}} (jb_i)(t - jp_i)} \tag{3}$$

Manuscript received July 15, 1999; revised July 21, 2001. This work was supported in part by the National Science Foundation under NCR-9706372 and CCR-0073505 and by Ericsson Inc.

S. Siwamogsatham and M. P. Fitz are with the Department of Electrical Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: siwamogs@ee.eng.ohio-state.edu; fitz.7@osu.edu).

J. H. Grimm is with Wiscom Technologies, Clark, NJ 07066-1247 USA (e-mail: jimm@ieee.org).

Communicated by M. L. Honig, Associate Editor for Communications.

Publisher Item Identifier S 0018-9448(02)01999-5.