Space-Time Modulation for Unknown Fading

Bertrand M. Hochwald and Thomas L. Marzetta

Mathematical Sciences Research Center
Bell Laboratories, Lucent Technologies
600 Mountain Avenue, Murray Hill, NJ 07974-0636

ABSTRACT

Recent theoretical and experimental work at Bell Laboratories has demonstrated that some widely accepted views about our ability to communicate across a wireless fading channel appear to be incorrect. It now seems that the effects of channel fading, rather than being harmful, can actually be beneficial to wireless communications. BLAST (Bell Labs Layered Space Time) is an example of a multiple-antenna communication link whose capacity in a fading environment grows linearly with the minimum of the number of antennas at the transmitter and receiver, with no increase in bandwidth or transmitted power. BLAST is intended to work in a quasi-stationary environment since it expects the fading coefficients between pairs of transmitter and receiver antennas to change slowly enough for the receiver to track them.

With mobile wireless in mind, we have taken a radically different approach where we assume that nobody knows the propagation coefficients. Our principal results are that capacity cannot be improved by making the number of transmit antennas greater than the symbol-duration of the fading coherence interval, and that the capacity-attaining signals have considerable structure that leads directly to a new signaling scheme: Unitary Space-Time Modulation.

Unitary Space-Time Modulation is ideally suited for fast fading mobile environments, and it does not require the receiver to know or to learn the propagation coefficients. The signals are oblong complex matrices whose columns are orthonormal, each column representing the signal that is fed to one of the transmit antennas during an interval where the fading is approximately constant. Large constellations of these signals are designed systematically according to a special criterion that attempts to minimize the singular values associated with pair-wise products of the signals. Our simulations demonstrate diversity gains for multiple transmit antennas compared with one transmit antenna - all without knowing or learning the propagation coefficients.

Keywords: wireless, space-time modulation, antenna array

1. INTRODUCTION

A traditional application of multiple-element antenna arrays to wireless communication has been to improve signal-to-interference ratio, for example by steering the main lobe in the direction of the desired signal and constructing nulls in the direction of undesired signals. Multiple antenna receiver arrays have also been used to mitigate the harmful effects of fading by providing so-called diversity gains.[1] More recently, research by Telatar[2] and by Foschini[3] suggests a much more exciting possibility: multiple antenna links can potentially sustain very high data rates without an accompanying increase in either bandwidth or transmit power. This theory was confirmed strikingly in recent experimental demonstrations at Bell Labs of a particular embodiment called BLAST (Bell Labs Layered Space Time). Two key assumptions underlie the BLAST effect: the complex fading coefficients between every pair of transmit and receive antennas are statistically independent with Rayleigh distributed magnitude and uniformly distributed phase, and these fading coefficients are known to the receiver. However it is not always possible for the receiver to learn the fading coefficients, particularly if either the transmitter or the receiver is mobile, or if there are a large number of transmit antennas.

Our research departs from that of Telatar and Foschini by treating the fading coefficients as random and unknown. We impose no ad-hoc training schemes, relying solely on information theory to determine the capacity of the link and the structure of the capacity attaining signals. We model the fading coefficients as random, unknown, and piecewise constant, remaining constant for T symbol periods, after which they change to new independent values

B.M.H.: hochwald@research.bell-labs.com
T.L.M.: tlm@research.bell-labs.com

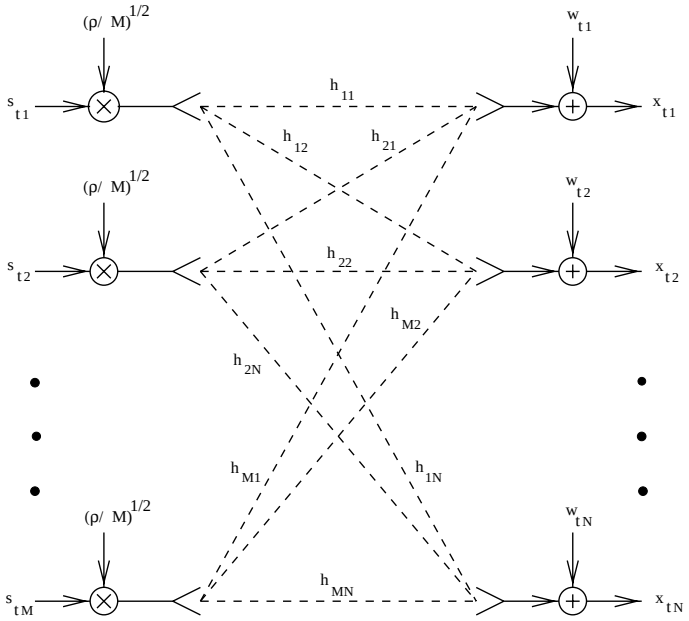


Figure 1. Wireless link comprising M transmitter and N receiver antennas. Every receiver antenna is connected to every transmitter antenna through an independent, random, unknown propagation coefficient having Rayleigh distributed magnitude and uniformly distributed phase. As M changes, the total transmit power remains fixed.

which they maintain for another T symbol periods, and so on. This mimics, in a tractable manner, some important features of an actual fading process. If channel coding is performed over multiple T-symbol blocks, then the noisy channel coding theorem of information theory can be invoked to provide the theoretical channel capacity. Computing channel capacity involves optimizing mutual information with respect to the joint probability density of the signals that are fed into the multiple transmitter antennas within each T-symbol block, an extremely high dimensional problem. We have shown[4] that the capacity attaining signals have considerable structure that enables us to solve the capacity computation problem for many cases of interest, and that leads directly to a new signaling scheme: Unitary Space-Time Modulation.

Unitary Space-Time Modulation is ideally suited for fast fading mobile environments where nobody knows the propagation coefficients, and under certain reasonable conditions these signals can theoretically attain capacity when combined with appropriate channel coding. The design of constellations of these signals requires a special criterion[5] because the fading coefficients are unknown. Indeed the traditional maximum Euclidean distance criterion is not usable. Most applications of Unitary Space-Time Modulation would require large constellations of signals. To that end, a systematic method of constructing these constellations was developed.[6]

We introduce a model for the multiple antenna wireless link in Section 2. Section 3 reviews the BLAST effect, which applies when the receiver knows the propagation coefficients. The contrary assumption - that nobody knows the propagation coefficients - drives Section 4, where the development of Unitary Space-Time Modulation is seen as its natural consequence.

2. SIGNAL MODEL: RANDOM PIECEWISE-CONSTANT FLAT FADING

Figure 1 displays a communication link comprising M transmitter antennas and N receiver antennas that operates in a Rayleigh flat-fading environment. Each receiver antenna responds to each transmitter antenna through a statistically independent complex-valued fading coefficient that is constant in frequency over the bandwidth of the signals, and that is constant for T symbol periods. The received signals are corrupted by additive noise that is statistically independent among the N receivers and the T symbol periods.

At every time, the average of the M expected transmitter powers is equal to one,

$$\frac{1}{M}\sum_{m=1}^{M}\mathrm{E}|s_{tm}|^2 = 1, \qquad t = 1,\ldots,T. \tag{1}$$

The scalar ρ denotes the expected signal-to-noise ratio at each receiver. Normalization ensures that the total expected transmitted power is independent of M for a fixed ρ.

Let

$$S = \begin{bmatrix} s_{11} & \ldots & s_{1M} \\ \vdots & & \\ s_{T1} & \ldots & s_{TM} \end{bmatrix}, \qquad X = \begin{bmatrix} x_{11} & \ldots & x_{1N} \\ \vdots & & \\ x_{T1} & \ldots & x_{TN} \end{bmatrix},$$

$$H = \begin{bmatrix} h_{11} & \ldots & h_{1N} \\ \vdots & & \\ h_{M1} & \ldots & h_{MN} \end{bmatrix}, \qquad W = \begin{bmatrix} w_{11} & \ldots & w_{1N} \\ \vdots & & \\ w_{T1} & \ldots & w_{TN} \end{bmatrix},$$

where S is the $T \times M$ matrix of transmitted signals, X is the $T \times N$ matrix of received signals, H is the $M \times N$ matrix of propagation coefficients, and W is the $T \times N$ matrix of additive noise components. Then

$$X = \sqrt{\frac{\rho}{M}}SH + W. \tag{2}$$

Both the fading coefficients and the additive receiver noise are independent zero-mean, circularly-symmetric complex Gaussian with unit variance ($\mathrm{CN}(0,1)$), i.e.,

$$p(h_{mn}) = (1/\pi)\exp\left\{-|h_{mn}|^2\right\}.$$

Conditioned on S and H, the components of the received signal X are independent, complex Gaussian, with unit variance but with nonzero means,

$$p\left(X \mid S,H\right) = \frac{\exp\left(-\mathrm{tr}\left\{\left[X-(\rho/M)^{1/2}SH\right]\left[X-(\rho/M)^{1/2}SH\right]^{\dagger}\right\}\right)}{\pi^{TN}}, \tag{3}$$

where "tr" denotes "trace". Conditioned S only, the components of X are zero-mean complex Gaussian, with a covariance structure that depends on S,

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

where I_T denotes the $T \times T$ identity matrix. The former density (3) is useful for describing the channel when H is known to the receiver (i.e., for BLAST), while the latter density (4) is appropriate for the case where H is not known to the receiver. In either case a single user, with access to all elements of the transmit antenna array, is attempting to send a message to someone who has access to all elements of the receive array. In neither case does the transmitter know the propagation coefficients.

3. PROPAGATION COEFFICIENTS KNOWN TO THE RECEIVER; BLAST

When the propagation coefficients H are known to the receiver, the received signal is (conditionally) independent over time and over receiver index. In effect we have a memoryless channel, and it is sufficient only to consider the row vectors of S and X as functions of time t,

$$p(x_t \mid s_t,H) = \frac{\exp\left(-\left[x_t-(\rho/M)^{1/2}s_tH\right]\left[x_t-(\rho/M)^{1/2}s_tH\right]^{\dagger}\right)}{\pi^{N}}. \tag{5}$$

If we ignore, for the moment, the fact that H is random, and merely assume that it is known to the receiver, then the mutual information between the transmitted and the received signal is

$$I(x_t; s_t \mid H) = \mathrm{E}_{x_t, s_t|H} \left\{ \log \left(\frac{p(x_t \mid s_t, H)}{\mathrm{E}_{s_t} \{ p(x_t \mid s_t, H) \}} \right) \right\}, \tag{6}$$

where the logarithm is base-2. The channel capacity is calculated by choosing the probability density for the transmitted signal $p(s_t)$ that maximizes the mutual information, subject to the power constraint (1), and it can be shown that this is accomplished by choosing the components of s_t to be independent, $\mathrm{CN}(0,1)$. By substituting this density into (6) we obtain the capacity (in bits per time-sample),

$$C(x_t; s_t \mid H) = \log \det \left[I_M + \frac{\rho}{M} H H^\dagger \right]. \tag{7}$$

In this case, capacity can theoretically be attained by performing channel coding over many time-samples during which H remains constant.

The channel capacity depends critically on the structure of the propagation matrix H. In keeping with the expected power constraint (1) suppose that we attempt to minimize or maximize the capacity (7) with respect to H, subject to the constraint $\mathrm{tr}\left\{ H H^\dagger \right\} = MN$. It can be shown that the minimum capacity is obtained when H has rank 1, while the maximum capacity is obtained when the $\min(M,N)$ singular values of H are all equal. The respective lower and upper bounds are then

$$\log(1 + \rho N) \leq C(x_t; s_t \mid H) \leq \min(M,N) \cdot \log \left(1 + \frac{\rho N}{\min(M,N)} \right). \tag{8}$$

The case where H has rank 1 occurs, for example, when linear transmit and receive arrays operate in free space in each other's far field. In that case there is no point in having more than one transmit antenna, and the only effect of having multiple receive antennas is to increase the effective SNR from ρ to $N\rho$. But traditionally this non-fading propagation environment was regarded as the most favorable!

Telatar[2] and Foschini[3] independently discovered that when the elements of H are random and statistically independent, H has full rank, with singular values that typically are nearly equal, which corresponds to the capacity upper bound (8). This is a revolutionary development in wireless since the capacity increases linearly with the smaller of the number of transmit or receive antennas, and this capacity increase occurs without any increase in bandwidth or power.

3.1. Shannon Capacity

Telatar[2] assumed that the propagation coefficients are random as modeled in Section 2, that they are known to the receiver, and that they change to new independent values for every sample t (i.e., $T = 1$). Then if channel coding is performed over many time samples, the occurrences of favorable values of H can compensate for the occurrences of unfavorable values of H. The mutual information is

$$I(x_t; s_t, H) = \mathrm{E}_H \left\{ I(x_t; s_t \mid H) \right\}, \tag{9}$$

and again, mutual information is maximized by the components of s_t being independent $\mathrm{CN}(0,1)$. The resulting capacity is

$$C(x_t; s_t, H) = \mathrm{E}_H \left\{ I(x_t; s_t \mid H) \right\} = \mathrm{E}_H \left\{ \log \det \left[I_M + \frac{\rho}{M} H H^\dagger \right] \right\}. \tag{10}$$

As $M \to \infty$ for fixed N, or as $N \to \infty$ for fixed M, it can be shown that the capacity (10) converges to the upper bound in (8).

3.2. Outage Capacity; BLAST

Foschini[3] took a somewhat different approach than Telatar, by assuming that the transmitter and receiver agree in advance on some data transmission rate (and the associated modulation and coding scheme), and that a message is sent at this rate over some finite number of time samples during which H is constant and known to the receiver. Because H is random, the capacity itself is random according to (7), and there is some chance that the capacity will

be too small to support the data transmission rate. Such an event is called an outage, and no amount of coding during the transmission interval can reduce its probability of occurrence. In practice one would choose a transmission rate such that the outage probability is acceptable. In fact Foschini's calculations show that very high outage capacities occur with high probability when multiple transmit and receive antenna arrays are employed.

BLAST is a particular scheme for achieving a significant fraction of the theoretical outage capacity that uses only ordinary modulation and coding schemes. The message bits to be transmitted are divided equally among the M transmit antennas, and the modulation and coding for each transmit antenna occurs independently of the modulation and coding for the other antennas. The number of receive antennas must be greater than or equal to the number of transmit antennas, $N \geq M$. The received signal at time t is

$$x_t = (\rho/M)^{1/2} s_t H + w_t, \tag{11}$$

so every receive antenna sees a linear combination of the M transmitted signals. One obvious and simple way to recover the individual signals would be to multiply the received signal by the pseudo-inverse of H (recall that H is known to the receiver)

$$x_t H^\dagger \left(HH^\dagger\right)^{-1} = (\rho/M)^{1/2} s_t + w_t H^\dagger \left(HH^\dagger\right)^{-1}. \tag{12}$$

Foschini showed that the pseudo-inverse is well-behaved in a probabilistic sense, so the operation (12) can be performed without undue amplification of the noise. BLAST employs an improved scheme involving two operations called nulling and cancellation.

Nulling and cancellation may be explained in terms of the QR factorization of the propagation matrix H.[7] With its knowledge of H, the receiver can always perform the factorization $H = RQ$, where Q is $N \times N$ unitary, and R is an $M \times N$ matrix that is lower-triangular (i.e., $r_{mn} = 0,\ \forall\ n > m$). The nulling operation applies the conjugate-transpose of Q to the received signal vector,

$$x_t Q^\dagger \equiv y_t = (\rho/M)^{1/2}\, s_t R + w_t Q^\dagger = (\rho/M)^{1/2}\, s_t R + \nu_t. \tag{13}$$

The nulling operation, in effect, renders the propagation matrix lower-triangular with no amplification of noise. The last $N - M$ components of y_t may be discarded since they have no dependence on the transmitted signal. Because of the triangular structure of R, the M-th component of y_t depends only on the M-th component of s_t, and this signal is then demodulated and decoded without interference from the other signals. If a reliable code is used, then the associated message bits are recovered correctly with high probability, and a replica of the M-th component of s_t can be reconstructed by the receiver, and then subtracted from the $(M-1)$-th component of y_t to allow the demodulation and decoding of the $(M-1)$-th component of s_t without interference. The process of reconstructing the detected components of the signal and subtracting to eliminate interference is called cancellation, and it is applied in succession to recover all of the message bits.

The key conditions required for a realization of the BLAST effect are 1) the statistical independence of the components of the propagation matrix, and 2) the knowledge of the values of this propagation matrix by the receiver. The existence of the first condition for various environments must be investigated experimentally. One way for the receiver to learn the propagation coefficients is for the transmitter to send known training signals, from which it estimates H. It turns out that the number of training samples required for successful nulling and cancellation is independent of the number of receiver antennas, but is proportional to the number of transmitter antennas.[8]

In a fixed environment, the propagation matrix stays constant for long intervals, and there is adequate time for both training and sending a message. In a mobile environment, however, H may change over a small number of time samples. It is instructive to consider the optimum number of transmit antennas in order to maximize the transmission rate. Foschini[3] has shown that asymptotically, the outage capacity is proportional to the number of transmit antennas, $C = \alpha M$ for some constant α, if the number of receive antennas is also proportion to M. The required number of training samples is equal to βM for some constant β. If H remains constant for a total of T symbols, then the number of bits that can be transmitted, allowing for training, is equal to $\alpha M(T - \beta M)$. This is maximized with respect to M for $\beta M = T/2$, implying that half of the interval between fades should be used for training, and the remaining half for data transmission. In short, when seeking high capacity in a mobile environment, learning the propagation coefficients is a major problem.

In fact, there is no fundamental reason why resources should be expended to learn the propagation coefficients, as we see in the next section.

4. MOBILE ENVIRONMENT: NOBODY KNOWS THE PROPAGATION COEFFICIENTS

When nobody knows the random propagation coefficients, the channel is completely described by the conditional probability density (4). An inspection of (2) discloses two sources of noise: additive noise W, and multiplicative noise H. Conditioned on the transmitted signal, the received signals are zero-mean, independent and identically distributed from one receiver to another. The transmitted signal S manifests itself only through the temporal correlation of the received signal.

The mutual information between the transmitted and received signals is

$$I(X;S) = \mathrm{E}_{X,S}\left\{\log\left(\frac{p(X \mid S)}{\mathrm{E}_S\left\{p(X \mid S)\right\}}\right)\right\}. \tag{14}$$

We assume that the fading coefficients change to new independent realizations every T symbol periods. By performing channel coding over multiple fading intervals, the intervals of favorable fading compensate for the intervals of unfavorable fading. Data can be transmitted reliably at any rate less than the channel capacity, where the capacity is the least upper bound on the mutual information between X and S, or

$$C(X;S) = \sup_{p(S)} I(X;S), \tag{15}$$

subject to the average power constraint (1).

The above maximization is formidable, involving the joint probability density of TM complex scalars. The special case of one transmit antenna and one receive antenna, with the fading coefficient independent for every symbol period (i.e., $T = M = N = 1$) is treated in[9] where it is shown that the capacity attaining signal is real, nonnegative, with a discrete probability distribution that was determined numerically. As happens sometimes in mathematics, the larger and more general problem turns out to be more tractable. By exploiting certain symmetries: that, on average, all time samples are equally good, and all transmit antennas are equally good, much can be said about the structure of the capacity attaining signals.[4]

4.1. Coherence Interval Determines Ultimate Capacity

The effect of the transmitted signals on the conditional probability density of the received signals (4) is through the $T \times T$ matrix $SS^\dagger$. Any joint density of the elements of $SS^\dagger$ can be realized with at most $M = T$ transmit antennas. As a consequence, the capacity obtained with any $M > T$ transmit antennas can also be obtained by $M = T$ transmit antennas.

If the coherence interval T is ten symbols then there is no point in using more than ten transmitters. In the extreme case where the fading is independent for every symbol (i.e., $T = 1$) a single transmitter antenna gives the same capacity as any multi-element transmitter antenna. In short, there is a fundamental limitation on capacity in a mobile environment that is determined by the symbol duration of the coherence interval. Of course capacity will continue to increase as the number of receive antennas N increases, but the improvement is only logarithmic in N. We previously argued heuristically, in the context of BLAST, that there should be such a limitation, but we now see that this exists irrespective of any training considerations.

In what follows we assume that $M \leq T$.

4.2. Capacity Attaining Density

Considerations of symmetry reveal that the capacity attaining signal matrix S has considerable structure. In particular, the signal matrix that achieves capacity can always be factored as $S = \Phi V$, where Φ is an $T \times M$ isotropically distributed unitary matrix, and V is an independent $M \times M$ real, nonnegative, diagonal matrix.

An isotropically distributed unitary matrix has a probability density that is unchanged when the matrix is premultiplied by any deterministic unitary matrix. More details, including the probability density of these matrices, may be found in.[4] The signal that is fed into the m-th transmit antenna at symbol t is $s_{tm} = v_m \phi_{tm}$, where v_m is a nonnegative real random scalar that is independent of the random unit vector $\{\phi_{1m}, \ldots, \phi_{Tm}\}$. The signals that are fed into the different antennas are mutually orthogonal with respect to t.

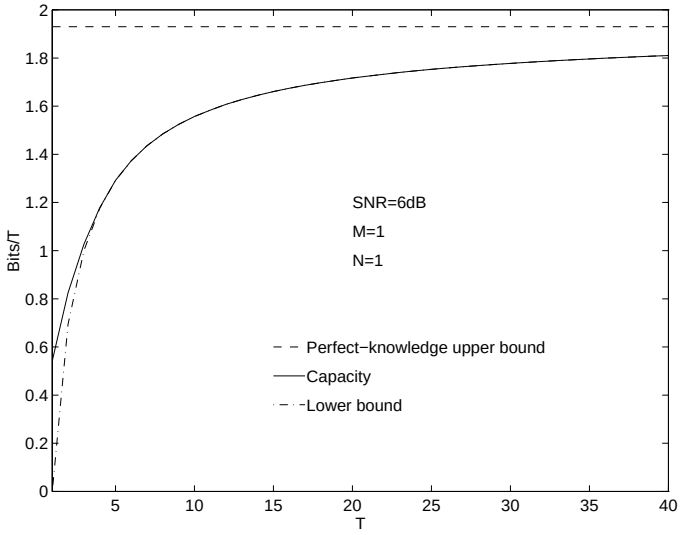


Figure 2. Normalized capacity, and upper and lower bounds, versus coherence interval T (SNR=6dB, one transmitter antenna, one receiver antenna). The lower bound and capacity meet at $T = 4$. The capacity approaches the perfect-knowledge capacity as $T \to \infty$.

Thus, the problem of maximizing $I(X;S)$ with respect to the joint probability density of the $T \cdot M$ complex elements of S reduces to the simpler problem of maximizing $I(X;S)$ with respect to the joint probability density of the M nonnegative real diagonal elements of V.

It furthmore can be shown[4,5] asymptotically, either as $T \gg M$, or as $\rho \to \infty$ for $T > M$, that $V \to \sqrt{T} I_M$, so that all of the information is sent on the directions of the orthonormal column vectors of Φ.

Figure 2 shows the normalized capacity (15) in units of bits per symbol as a function of T (the solid curve), for $M = N = 1$ and for 6 dB expected SNR. The upper dotted curve is the capacity that would be obtained if H were known to the receiver (10), while the lower dotted curve is the mutual information (14) that is obtained by setting $V = \sqrt{T}$. As T becomes big, the capacity for unknown H approaches the capacity for known H, an intuitively reasonable result since there would be time for training. For $T \geq 4$, we note that $V = \sqrt{T}$, and all information is sent on the direction of the T-component unit vector Φ.

4.3. Unitary Space-Time Modulation

The previous results suggest the use of a constellation of L signals $S_1 = \sqrt{T}\Phi_1, \ldots, S_L = \sqrt{T}\Phi_L$, where $\Phi_\ell^\dagger \Phi_\ell = I_M$, a scheme that we call Unitary Space-Time Modulation, to transmit binary information over this link. It is shown[5] that the maximum likelihood decoder for a constellation of Unitary Space-Time signals is

$$\Phi_{\mathrm{ml}} = \arg \max_{\Phi_\ell = \Phi_1, \ldots, \Phi_L} \mathrm{tr}\{X^\dagger \Phi_\ell \Phi_\ell^\dagger X\},$$

and that the two-signal probability of mistaking S_1 for S_2 or vice-versa is

$$P_e = \sum_j \mathrm{Res}_{\omega = i a_j} \left\{ -\frac{1}{\omega + i/2} \prod_{\substack{m=1 \\ d_m < 1}}^{M} \left[\frac{1 + \rho T/M}{(\rho T/M)^2 (1 - d_m^2)(\omega^2 + a_m^2)} \right]^N \right\}, \tag{16}$$

where $1 \geq d_1 \geq \ldots \geq d_M \geq 0$ are the singular values of the $M \times M$ correlation matrix $\Phi_2^\dagger \Phi_1$, and

$$a_m \equiv \sqrt{\frac{1}{4} + \frac{1 + \rho T/M}{(\rho T/M)^2 (1 - d_m^2)}}.$$

The probability of error P_e decreases as any d_m decreases, and has Chernoff upper bound

$$P_e \leq \frac{1}{2} \prod_{m=1}^{M} \left[\frac{1}{1 + \frac{(\rho T/M)^2 (1-d_m^2)}{4(1+\rho T/M)}} \right]^N . \tag{17}$$

The probability of error is lowest when $d_1 = \ldots = d_M = 0$ and highest when $d_1 = \ldots = d_M = 1$. We obtain $d_1 = \ldots = d_M = 0$ when the columns of S_2 are all orthogonal to all the columns of S_1. The ideal constellation $S_1, \ldots, S_L$ therefore has all the columns of S_ℓ orthogonal to all the columns of $S_{\ell'}$ for $\ell' \neq \ell = 1, \ldots, L$. However, because the columns of each S_ℓ are within themselves orthogonal to one another, all the pairwise $d_1, \ldots, d_M$ cannot be made zero if $L > T/M$.

Let

$$\left\| \Phi_{\ell'}^\dagger \Phi_\ell \right\| = \sqrt{\frac{d_1^2 + \ldots + d_M^2}{M}} \tag{18}$$

denote the root-mean-square norm of the singular values of the correlation matrix (this is a scaled Frobenius norm). This norm lies between zero and one, with small norm generally implying small probability of error. For a given constellation, let

$$\delta = \max_{1 \leq \ell < \ell' \leq L} \left\| \Phi_\ell^\dagger \Phi_{\ell'} \right\| . \tag{19}$$

In general, constellations should have δ as small as possible.

This design criterion is markedly different from the familiar maximum-Euclidean-distance criterion, and it arises entirely because the fading coefficients are unknown to the receiver. Because of this, antipodal pairs of signals $\pm S$ are indistinguishable, for example. In constructing a constellation of unitary space-time signals, we note that the probability of error is invariant to two types of transformations: 1) left multiplication by a common $T \times T$ unitary matrix, $\Phi_\ell \to \Psi^\dagger \Phi_\ell, \ \ell = 1, \cdots, L$; 2) right multiplication by individual $M \times M$ unitary matrices, $\Phi_\ell \to \Phi_\ell \Upsilon_\ell, \ \ell = 1, \cdots, L$.[5]

For the case of a single transmit antenna $M = 1$, the signals are T-dimensional vectors, there is only one singular value for each pair of signals, and it is equal to the magnitude of the correlation between the unit vectors. Figure 3 illustrates the pairwise correlations of a constellation of 32 such signals for $T = 5$. A simple iterative algorithm was used to design the constellation[5] and it achieves $d_{\max} = .515$.

4.4. Systematic Design of Unitary Space-Time Signals

The size of the constellation must be at least $L = 2^{RT}$ where R is the transmission rate in bits per time sample. Even moderate values of R and T lead to very large constellations, involving a difficult design problem, and a cumbersome storage problem. A systematic method of designing Unitary Space-Time constellations was developed[6] to address these issues.

The systematically designed constellations have considerable structure,

$$\Phi_\ell = \Theta^{\ell - 1} \Phi_1, \quad \ell = 1, \cdots, L \tag{20}$$

where Θ is a $T \times T$ diagonal unitary matrix that is an L-th root of the identity matrix I_T,

$$[\Theta]_{tt} = e^{i \frac{2\pi}{L} u_t}, \quad t = 1, \cdots, T, \tag{21}$$

and the $\{u_t\}$ are integers such that $0 \leq u_1, \cdots, u_T \leq L - 1$. Beginning with the first signal Φ_1, the unitary matrix Θ successively rotates the columns of Φ_1 in the T-dimensional complex space, while maintaining their orthonormality. In a manner that cannot easily be visualized, the columns evidently move in a complicated zig-zag fashion over the surface of the high-dimensional complex sphere, thereby generating a constellation having a low value of δ (19).

The $\{u_t\}$ are chosen by iterative random selection. The optimization is facilitated by the fact that the correlation between members of the constellation has a Toeplitz structure, so only $L - 1$ correlations have to be checked, as opposed to $(L^2 - L)/2$. The choice of the initial signal Φ_1 is less critical; an effective procedure is to choose randomly M columns from the $T \times T$ FFT unitary matrix.

Figure 4 shows the bit-error rate for systematically designed constellations of Unitary Space-Time signals, for M =1, 2, and 3 transmit antennas, and $T = 8$. The constellations have size $L = 2^8$. For sufficiently high SNR's, adding more transmit antennas dramatically reduces the probability of error.

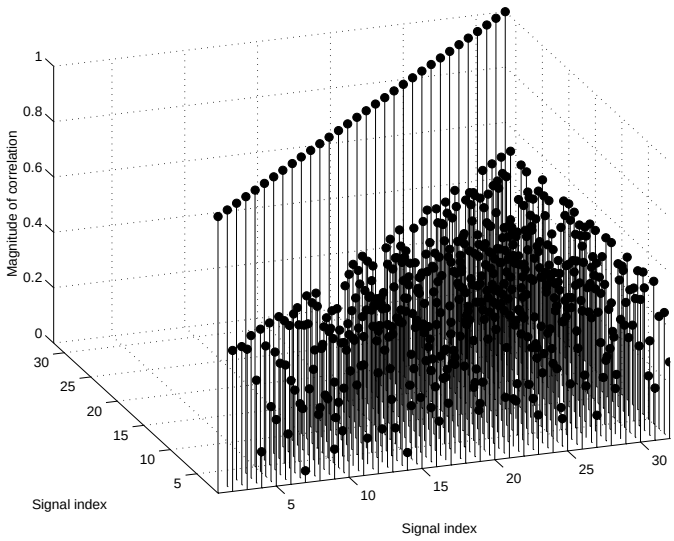


Figure 3. Magnitudes of pair-wise correlations for a constellation of 32 Unitary Space Time signals for $T = 5$ and $M = 1$. The diagonal entries with value 1.0 represent each signal correlated with itself. The maximum off-diagonal correlation is equal to .515

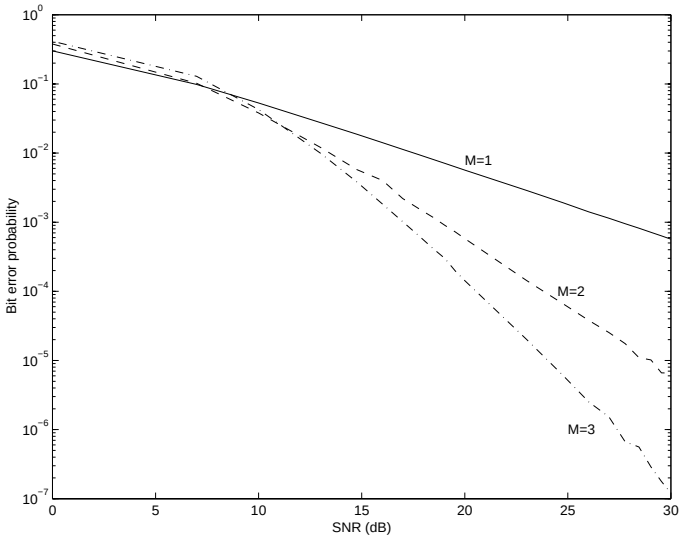


Figure 4. Bit error rate for $M = 1, 2$, and 3 transmitter antennas versus SNR with $N = 1$ receiver antenna, $T = 8$ and rate $R = 1$ bit/time-sample.

5. CONCLUSIONS

Among the technologies that drive wireless communications, multiple antenna arrays are unique in their potential to give order-of-magnitude improvements. Unitary Space-Time Modulation addressess the mobile environment where the propagation coefficients cannot be estimated reliably. Our current research is focussed on fast decoding methods for large signal constellations.

REFERENCES

1. E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inf. Th.* **44**(6), pp. 2619–2692, 1998.
2. I. E. Telatar, "Capacity of multi-antenna *g*aussian channels," tech. rep., AT&T Bell Laboratories, 1995.
3. G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment," *Bell Labs Tech. J.* **1**(2), pp. 41–59, 1996.
4. T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in *r*ayleigh flat fading," *IEEE Trans. Inf. Th.* **45**(1), pp. 139–157, 1999.
5. B. M. Hochwald and T. L. Marzetta, "Unitary *s*pace-*t*ime modulation for multiple-antenna communications in rayleigh flat fading," tech. rep., Bell Laboratories, Lucent Technologies, 1998.
6. B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of *u*nitary *s*pace-*t*ime constellations," tech. rep., Bell Laboratories, Lucent Technologies, 1998.
7. T. L. Marzetta, "The QR factorization, and its role in a new high capacity, layered space-time architecture for wirelsss communication," tech. rep., Bell Laboratories, Lucent Technologies, 1996.
8. T. L. Marzetta, "Estimating channel characteristics for a high capacity, layered space-time architecture for wireless communications," tech. rep., Bell Laboratories, Lucent Technologies, 1996.
9. I. C. AbouFaycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete-time Rayleigh fading channels," *Proc. Int. Symp. Info. Th.* , p. 473, (Ulm, Germany), 1997.