

Space-Time Codes for Wireless Optical Communications

Shane M. Haas

*Laboratory for Information and Decisions Systems, and Research Laboratory of Electronics, Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA 02139, USA
Email: shaas@mit.edu*

Jeffrey H. Shapiro

*Laboratory for Information and Decisions Systems, and Research Laboratory of Electronics, Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA 02139, USA
Email: jhs@mit.edu*

Vahid Tarokh

*Laboratory for Information and Decisions Systems, Massachusetts Institute of Technology, 77 Massachusetts Avenue,
Cambridge, MA 02139, USA
Email: vahid@mit.edu*

Received 1 June 2001

A space-time channel coding technique is presented for overcoming turbulence-induced fading in an atmospheric optical heterodyne communication system that uses multiple transmit and receive apertures. In particular, a design criterion for minimizing the pairwise probability of codeword error in a space-time code (STC) is developed from a central limit theorem approximation. This design criterion maximizes the mean-to-standard-deviation ratio of the received energy difference between codewords. It leads to STCs that are a subset of the previously reported STCs for Rayleigh channels, namely those created from orthogonal designs. This approach also extends to other fading channels with independent, zero-mean path gains. Consequently, for large numbers of transmit and receive antennas, STCs created from orthogonal designs minimize the pairwise codeword error probability for this larger class of fading channels.

Keywords and phrases: space-time codes, optical communication, orthogonal designs.

1. INTRODUCTION

In atmospheric optical communication, lognormal fading arising from refractive-index turbulence can make the recovery of a transmitted signal extremely difficult at the receiver. As a result, the receiver must have a redundant replica of the transmitted signal for reliable communication. Space-time codes (STCs) provide both spatial and temporal redundancy, or diversity, by using multiple apertures (antennas) over several time-slots.

Tarokh et al., in [1], established space-time code design criteria for Rayleigh and Ricean fading channels. These design criteria specify the pairwise properties of codewords from the STC. In this paper, we derive a similar design criterion for the lognormal fading channel based on a central limit theorem approximation. Our criterion leads to STCs created from orthogonal designs, a subset of the previously reported STCs for Rayleigh channels. Tarokh et al., in [2], showed that such codes have a decoding algorithm requiring only linear processing at the receiver. We show that these STCs also maximize the mean-to-standard-deviation ratio of

the received energy difference between codewords, a result analogous to maximal ratio combining.

Our derivation extends to other fading channels with independent, zero-mean path gains. In other words, we show that for large numbers of transmit and receive antennas, STCs created from orthogonal designs minimize the pairwise codeword error probability regardless of the individual path-gain fading distributions.

This paper is structured as follows. Section 2 describes our channel model and the objective of space-time coding. Section 3 derives the STC design criterion for lognormal fading channels based on a central limit theorem approximation. Section 4 discusses the performance and presents an example of these STCs.

2. PROBLEM FORMULATION

Consider a line-of-sight atmospheric optical heterodyne communication system that uses multiple transmit and receive apertures, as shown in Figure 1. The space-time encoder maps a segment of bits from the information

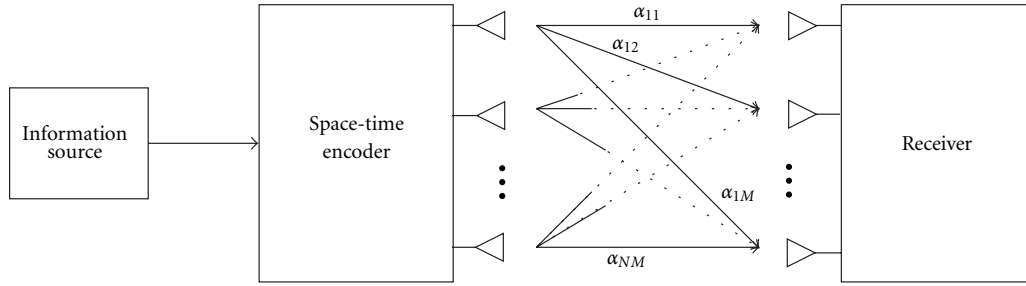


FIGURE 1: Block diagram of the transmitter, channel, and receiver.

source to a codeword of the STC. The codeword c , where $c = (c_1(1), c_1(2), \dots, c_1(T), c_2(1), \dots, c_2(T), \dots, c_N(1), \dots, c_N(T))$ is sent over T nonoverlapping adjacent discrete-time slots using N transmit apertures. During time slot t , transmit aperture n sends $c_n(t)$, a symbol from a quadrature-amplitude-modulation (QAM) signal constellation.

A laser beam propagating over a clear-weather, line-of-sight atmospheric path from transmitter exit optics to receiver entrance optics experiences amplitude and phase fluctuations due to refractive-index turbulence [3]. A propagation model, established in [4], based on the extended Huygens-Fresnel principle [5] characterizes this fading as a complex lognormal process with correlation times on the order of 10^{-3} to 10^{-2} seconds. At high data rates, a single fade can obliterate several message packets.

Because the duration of a fade is usually much longer than the length of a message packet, we will assume that the fades are constant during a codeword transmission. We will model the path gain from transmit aperture n to receive aperture m as $\alpha_{nm} = \exp(\chi_{nm} + j\phi_{nm})$. Here χ_{nm} , ϕ_{nm} are independent Gaussian random variables with moments $\text{var}(\chi_{nm}) = \sigma_\chi^2$, $E(\chi_{nm}) = -\sigma_\chi^2$, $\text{var}(\phi_{nm}) = \sigma_\phi^2 \gg 1$, and $E(\phi_{nm}) = 0$. The log-amplitude variance, σ_χ^2 , typically lies within the range 0.01 (mild fading) to 0.35 (severe fading). We also assume that the spacing between elements of the receiver aperture array is large enough to ensure that the path gains for different (n, m) values are approximately independent.

We will assume optical heterodyne reception, for which the detector output is known to consist of a frequency downshifted version of the incident optical field plus an additive white Gaussian noise [6]. We will use $w_m(t)$ to denote this additive Gaussian noise for receive aperture m during time slot t ; it is a complex-valued, zero-mean, white Gaussian random process with variance $N_0/2$ per real dimension.

Combining the fading and additive noise fluctuations, the signal at receive aperture $m \in \{1, \dots, M\}$ during time slot $t \in \{1, \dots, T\}$ is

$$r_m(t) = \sum_{n=1}^N \alpha_{nm} c_n(t) + w_m(t). \quad (1)$$

Given a received sequence $\{r_m(t) : 1 \leq m \leq M, 1 \leq t \leq T\}$ and knowledge of the path gains $\alpha = \{\alpha_{nm} : 1 \leq n \leq N, 1 \leq m \leq M\}$, the minimum probability of error receiver chooses

the codeword c that minimizes

$$\sum_{m=1}^M \sum_{t=1}^T \left| r_m(t) - \sum_{n=1}^N \alpha_{nm} c_n(t) \right|^2. \quad (2)$$

The exact probability of error is difficult to calculate for an STC with more than two codewords. An upper bound on this probability of error comes from the union bound

$$P_e \leq \sum_{c,e} \Pr(c \rightarrow e), \quad (3)$$

where $\Pr(c \rightarrow e)$ is the probability of decoding codeword c as codeword e in the absence of all other codewords. This sum is usually dominated by the terms of the closest, or minimum distance, codeword pairs. The union bound estimate of the codeword error probability is the sum of pairwise error probabilities of the minimum distance codeword pairs

$$P_e \approx K_{\min} \Pr(c \rightarrow e)_{\min}, \quad (4)$$

where K_{\min} is the average number of minimum-distance codeword neighbors and $\Pr(c \rightarrow e)_{\min}$ is the pairwise probability of erroneously decoding a pair of minimum distance codewords.

Given knowledge of the path gains and assuming equally-likely codewords, the pairwise probability of incorrectly decoding transmitted codeword c as codeword e is

$$\Pr(c \rightarrow e | \alpha) = Q\left(\sqrt{\frac{d^2(c, e)}{2N_0}}\right), \quad (5)$$

where

$$d^2(c, e) = \sum_{m=1}^M \sum_{t=1}^T \left| \sum_{n=1}^N \alpha_{nm} (e_n(t) - c_n(t)) \right|^2 \quad (6)$$

is the distance between codewords at the receiver, and $Q(x)$ is the area under the tail of the standard normal density function. Averaging over α , the unconditional probability of incorrectly decoding c as e is therefore,

$$\Pr(c \rightarrow e) = \int \Pr(c \rightarrow e | \alpha) p_\alpha(\alpha) d\alpha, \quad (7)$$

where $p_\alpha(\alpha)$ is the joint probability density function of the lognormal path gains.

Our ultimate objective is to construct a space-time code that minimizes the exact probability of error, P_e . In this paper, however, we will focus on minimizing $\Pr(c \rightarrow e)_{\min}$ in the union bound estimate of this probability.

3. DESIGN CRITERION

The integral in (7) is very difficult to evaluate analytically because of the lognormal density function. We will attempt to simplify its evaluation using a central limit theorem (CLT) approximation.

Rewriting (6) as

$$d^2(c, e) = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \alpha_{nm} \alpha_{km}^* A_{nk}, \quad (8)$$

where

$$A_{nk} = \sum_{t=1}^T (e_n(t) - c_n(t))(e_k(t) - c_k(t))^* \quad (9)$$

shows that $d^2(c, e)$ is the sum of MN^2 complex lognormal random variables.¹ Because the coefficients $\{A_{nk} : 1 \leq n, k \leq N\}$ and the central moments are bounded, no single term dominates the sum. Thus, we will use the central limit theorem to approximate its distribution as a truncated Gaussian with mean μ and variance σ^2 on the interval $d^2(c, e) \geq 0$. Using this approximation, we can rewrite (7) as

$$\Pr(c \rightarrow e) \approx \int_0^\infty Q\left(\sqrt{\frac{x}{2N_0}}\right) p_{X|X \geq 0}(x | X \geq 0) dx, \quad (10)$$

where

$$\begin{aligned} p_X(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(1/2\sigma^2)(x-\mu)^2}, \\ p_{X|X \geq 0}(x | X \geq 0) &= \frac{p_X(x)}{\Pr(X \geq 0)} \\ &= \frac{(1/\sqrt{2\pi\sigma^2}) e^{-(1/2\sigma^2)(x-\mu)^2}}{1 - Q(\mu/\sigma)}, \quad \text{for } x \geq 0. \end{aligned} \quad (11)$$

Define A as the matrix with A_{nk} as its nk th element. This matrix characterizes the relationship between codeword pairs of the space-time code. Our goal is to derive properties of A that minimize the CLT approximation to $\Pr(c \rightarrow e)$. We will do so by expressing (10) as a function of two normalized parameters that measure the fading strength and the signal-to-noise ratio. We then find bounds on the normalized fading strength based on the design matrix A . We conjecture that (10) is unimodal as a function of this fluctuation strength. We then show that for large numbers of transmit and receive apertures, minimizing the normalized fading strength, or equivalently choosing A to be a scaled identity

matrix, minimizes the CLT approximation to the pairwise probability of error.

3.1. Normalized parameters

Our first step in minimizing (10) is to rewrite $\Pr(c \rightarrow e)$ in terms of normalized parameters. The first normalized parameter measures the strength of the fading. Define the normalized fading strength, η , to be the standard-deviation-to-mean ratio of the energy difference between the codewords at the receiver, that is, $\eta = \sigma/\mu$. The second normalized parameter measures the total received signal-to-noise ratio and is defined as $\rho = \mu/N_0$.

With the change of variables $z = x/\mu$, (10) becomes

$$\Pr(c \rightarrow e; \eta^2, \rho) \approx \int_0^\infty Q\left(\sqrt{\frac{1}{2}\rho z}\right) p_{Z|Z \geq 0}(z) dz, \quad (12)$$

where

$$p_{Z|Z \geq 0}(z) = \frac{(1/\sqrt{2\pi\eta^2}) e^{-(1/2\eta^2)(z-1)^2}}{1 - Q(1/\eta)}, \quad \text{for } z \geq 0, \quad (13)$$

and Z is a Gaussian random variable with unit mean and variance η^2 .

3.2. Mean and variance calculations

To approximate $d^2(c, e)$ as Gaussian, we must first determine its mean μ and variance σ^2 . Notice that because $\sigma_\phi^2 \gg 1$, we have that $E(\alpha_{nm}) \approx E(\alpha_{nm}^2) \approx 0$. Also, because $E(\chi_{nm}) = -\sigma_\chi^2$, we find that $E(|\alpha_{nm}|^2) = 1$ and $E(|\alpha_{nm}|^4) = e^{4\sigma_\chi^2}$. The mean of $d^2(c, e)$ is then

$$\mu = E\left(\sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \alpha_{nm} \alpha_{km}^* A_{nk}\right) = M \operatorname{tr}(A), \quad (14)$$

where $\operatorname{tr}(A) = \sum_{n=1}^N A_{nn}$. We define the energy difference between transmitted codewords as

$$E_d = \operatorname{tr}(A) = \sum_{n=1}^N \sum_{t=1}^T |c_n(t) - e_n(t)|^2. \quad (15)$$

We can then express the total signal-to-noise ratio, ρ , as the sum of signal-to-noise ratios at each receive aperture, that is, $\rho = ME_d/N_0 = MSNR$, where $\text{SNR} = E_d/N_0$ is the signal-to-noise ratio at each receive aperture.

The second moment of $d^2(c, e)$ is

$$\begin{aligned} E(d^4(c, e)) &= \sum_{m_1=1}^M \sum_{n_1=1}^N \sum_{k_1=1}^N \sum_{m_2=1}^M \sum_{n_2=1}^N \sum_{k_2=1}^N A_{n_1 k_1} A_{n_2 k_2} \\ &\quad \times E(\alpha_{n_1 m_1} \alpha_{k_1 m_1}^* \alpha_{n_2 m_2} \alpha_{k_2 m_2}^*). \end{aligned} \quad (16)$$

To evaluate this summation, we can split it into two cases.

¹The scaled multiplication of lognormal random variables is also a lognormal random variable.

For $m_1 \neq m_2$, we have

$$\begin{aligned} & E(\alpha_{n_1 m_1} \alpha_{k_1 m_1}^* \alpha_{n_2 m_2} \alpha_{k_2 m_2}^*) \\ &= E(\alpha_{n_1 m_1} \alpha_{k_1 m_1}^*) E(\alpha_{n_2 m_2} \alpha_{k_2 m_2}^*) \\ &= \begin{cases} 1 & \text{if } n_1 = k_1 \text{ and } n_2 = k_2, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (17)$$

When $m_1 = m_2 = m$, we find that

$$\begin{aligned} E(\alpha_{n_1 m_1} \alpha_{k_1 m_1}^* \alpha_{n_2 m_2} \alpha_{k_2 m_2}^*) &= E(\alpha_{n_1 m} \alpha_{k_1 m}^* \alpha_{n_2 m} \alpha_{k_2 m}^*) \\ &= \begin{cases} e^{4\sigma_x^2} & \text{if } n_1 = k_1 = n_2 = k_2, \\ 1 & \text{if } n_1 = k_1 \neq n_2 = k_2, \\ 1 & \text{if } n_2 = k_2 \neq n_1 = k_1, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (18)$$

From these results it follows that the second moment of $d^2(c, e)$ is

$$\begin{aligned} & E(d^4(c, e)) \\ &= \sum_{m_1=1}^M \left(e^{4\sigma_x^2} \sum_{n=1}^N A_{nn}^2 + \sum_{n=1}^N \sum_{\substack{k=1 \\ k \neq n}}^N A_{nn} A_{kk} \right. \\ & \quad \left. + \sum_{n=1}^N \sum_{\substack{k=1 \\ k \neq n}}^N A_{kn} A_{nk} + \sum_{\substack{m_2=1 \\ m_2 \neq m_1}}^M \sum_{n=1}^N \sum_{k=1}^N A_{nn} A_{kk} \right) \\ &= M \left((e^{4\sigma_x^2} - 2) \sum_{n=1}^N A_{nn}^2 + \sum_{n=1}^N \sum_{k=1}^N |A_{nk}|^2 + M(\text{tr}(A))^2 \right), \end{aligned} \quad (19)$$

whence

$$\begin{aligned} \sigma^2 &= \text{var}(d^2(c, e)) \\ &= M \left((e^{4\sigma_x^2} - 1) \sum_{n=1}^N A_{nn}^2 + 2 \sum_{n=1}^N \sum_{k=1}^{n-1} |A_{nk}|^2 \right). \end{aligned} \quad (20)$$

Notice that we have assumed that the path gains are log-normally distributed, but we have only used the fact that they are independent and identically distributed with zero mean, unit variance, and finite fourth moment. Therefore, our method and results extend to all fading distributions that satisfy these weaker conditions.

3.3. Bounds on the normalized fading fluctuation

The μ and σ that we have found are tied to the design matrix A by (14) and (20), respectively. We will now derive bounds on their ratio $\eta = \sigma/\mu$ expressed in terms of A .

A lower bound, obtained via the Cauchy-Schwarz in-

equality, is

$$\begin{aligned} \eta^2 &= \frac{M \left((e^{4\sigma_x^2} - 1) \sum_{n=1}^N A_{nn}^2 + 2 \sum_{n=1}^N \sum_{k=1}^{n-1} |A_{nk}|^2 \right)}{M^2 \left(\sum_{n=1}^N A_{nn} \right)^2} \\ &\geq \frac{(e^{4\sigma_x^2} - 1) \sum_{n=1}^N A_{nn}^2 + 2 \sum_{n=1}^N \sum_{k=1}^{n-1} |A_{nk}|^2}{MN \sum_{n=1}^N A_{nn}^2}. \end{aligned} \quad (21)$$

Equality holds in (21) when $A_{nn} = \beta$, $n = 1, \dots, N$, for some positive real number β . Furthermore, setting $A_{nk} = 0$ for $n \neq k$ minimizes the numerator in (21). Thus we get the bound

$$\eta^2 \geq \frac{e^{4\sigma_x^2} - 1}{MN}, \quad (22)$$

with equality when $A = \beta I$, where I is the $N \times N$ identity matrix. Also, $A_{nn} = \sum_{t=1}^T |c_n(t) - e_n(t)|^2 = \beta$, $n = 1, \dots, N$ implies that $\beta = E_d/N$. Orthogonal designs [2] provide a method to construct STCs that satisfy the design criterion $A = (E_d/N)I$ and provide easy decoding at the receiver. Therefore, STCs created from orthogonal designs maximize the mean-to-standard-deviation ratio of the received energy difference between codewords.

We start the upper bound derivation by noticing that A is positive semi-definite [7] because it has an $N \times T$ square-root matrix B with nt th element $c_n(t) - e_n(t)$ such that $A = BB^T$ [1]. Let $\lambda_1, \dots, \lambda_N$ denote the nonnegative eigenvalues of A . For $e^{4\sigma_x^2} - 2 \geq 0$, an upper bound on η is found as follows:

$$\begin{aligned} \eta^2 &= \frac{(e^{4\sigma_x^2} - 2) \sum_{n=1}^N A_{nn}^2 + \sum_{n=1}^N \sum_{k=1}^N |A_{nk}|^2}{M \left(\sum_{n=1}^N A_{nn} \right)^2} \\ &\leq \frac{(e^{4\sigma_x^2} - 2) \sum_{n=1}^N \sum_{k=1}^N |A_{nk}|^2 + \sum_{n=1}^N \sum_{k=1}^N |A_{nk}|^2}{M \left(\sum_{n=1}^N A_{nn} \right)^2}, \end{aligned} \quad (23)$$

with equality when A is a diagonal matrix. Using $\text{tr}(A^2) = \sum_{n=1}^N \sum_{k=1}^N |A_{nk}|^2 = \sum_{n=1}^N \lambda_n^2$, this upper bound becomes

$$\eta^2 \leq \frac{e^{4\sigma_x^2} - 1}{M} \sum_{n=1}^N \left(\frac{\lambda_n}{\sum_{k=1}^N \lambda_k} \right)^2 \leq \frac{e^{4\sigma_x^2} - 1}{M}, \quad (24)$$

with equality when A is a diagonal matrix of rank one. The last inequality follows from

$$\sum_{n=1}^N \left(\frac{\lambda_n}{\sum_{k=1}^N \lambda_k} \right)^2 \leq \sum_{n=1}^N \frac{\lambda_n}{\sum_{k=1}^N \lambda_k} = 1, \quad (25)$$

which is met with equality when exactly one of the eigenvalues is nonzero.

For $e^{4\sigma_x^2} - 2 < 0$, an upper bound on η is found by suppressing the first term in σ^2 :

$$\eta^2 \leq \frac{\sum_{n=1}^N \sum_{k=1}^N |A_{nk}|^2}{M \left(\sum_{n=1}^N A_{nn} \right)^2} = \frac{1}{M} \sum_{n=1}^N \left(\frac{\lambda_n}{\sum_{k=1}^N \lambda_k} \right)^2 \leq \frac{1}{M}. \quad (26)$$

Equality in (26) requires A to be rank one with all its diagonal terms equal to zero, an impossibility if A is positive semi-definite.

The bounds on η are then

$$\frac{e^{4\sigma_x^2} - 1}{MN} \leq \eta^2 \leq \frac{\max\{1, e^{4\sigma_x^2} - 1\}}{M}. \quad (27)$$

The lower bound is achieved when $A = (E_d/N)I$. If $e^{4\sigma_x^2} - 1 \geq 1$, the upper bound is achieved when A has only one nonzero diagonal element. The upper bound is unachieved when $e^{4\sigma_x^2} - 1 < 1$.

3.4. Minimizing the probability of codeword error

To our knowledge, the codeword probability of error in (12) does not have a closed-form solution. In this section, we will analyze its asymptotic behavior, and conjecture that it is unimodal as a function of η^2 , that is, it has only one extremum, a maximum, for a fixed ρ .

First, we will fix a value for η and examine the behavior of $\Pr(c \rightarrow e; \eta^2, \rho)$ as we vary ρ . We saw in Section 3.3 that η is closely tied to the STC design matrix; therefore, fixing a value of η is in essence fixing a design matrix.

For small values of ρ , the probability of codeword error approaches one-half, that is,

$$\lim_{\rho \rightarrow 0} \Pr(c \rightarrow e; \eta^2, \rho) = \frac{1}{2}. \quad (28)$$

As ρ increases without bound, $\Pr(c \rightarrow e; \eta^2, \rho)$ decays as $1/\rho$, namely,

$$\begin{aligned} \Pr(c \rightarrow e; \eta^2, \text{large } \rho) &\approx \int_0^\infty Q\left(\sqrt{\frac{1}{2}\rho z}\right) p_{Z|Z \geq 0}(0) dz \\ &= \frac{e^{-1/2\eta^2}}{\sqrt{2\pi\eta^2(1 - Q(1/\eta))}} \int_0^\infty Q\left(\sqrt{\frac{1}{2}\rho z}\right) dz \\ &= \left(\frac{e^{-1/2\eta^2}}{\sqrt{2\pi\eta^2(1 - Q(1/\eta))}}\right) \frac{1}{\rho}, \end{aligned} \quad (29)$$

where $\int_0^\infty Q(\sqrt{(1/2)\rho z}) dz = 1/\rho$ using integration by parts.

We will now fix the total receiver signal-to-noise ratio, ρ , and determine the probability of codeword error for different values of normalized fading fluctuation, η , or equivalently, for different design matrices. As η approaches zero, the Gaussian probability density function in (12) becomes sharply peaked around the value $z = 1$. This sampling-like behavior results in

$$\lim_{\eta \rightarrow 0} \Pr(c \rightarrow e; \eta^2, \rho) = Q\left(\sqrt{\frac{\rho}{2}}\right). \quad (30)$$

Furthermore, for any fixed value of ρ ,

$$\lim_{\eta \rightarrow \infty} \Pr(c \rightarrow e; \eta^2, \rho) = 0, \quad (31)$$

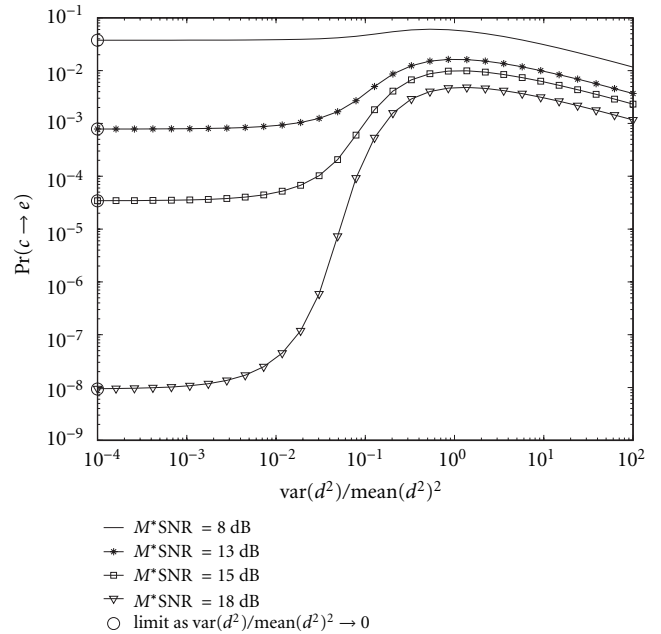


FIGURE 2: The probability of codeword error, $\Pr(c \rightarrow e; \eta^2, \rho)$, as a function of η^2 for $\rho = 8, 13, 15, 18$ dB.

because the Gaussian density approaches zero for large values of η .

The behavior of (12) for intermediate values of η is more difficult to evaluate analytically. We will, therefore, make the following conjecture as supported by numerical evaluations of $\Pr(c \rightarrow e; \eta^2, \rho)$.

Conjecture 1. For $0 < \eta < \infty$, $\Pr(c \rightarrow e; \eta^2, \rho)$ has only one extremum, a maximum, for a given value of ρ . Plots of (12) for different values of ρ are shown in Figure 2 to support this conjecture.

Assuming that $\Pr(c \rightarrow e; \eta^2, \rho)$ is unimodal in η^2 , its minimum must occur on the boundary of the allowable range for η in (27). In other words, if

$$\begin{aligned} \Pr\left(c \rightarrow e; \frac{e^{4\sigma_x^2} - 1}{MN}, \text{MSNR}\right) \\ < \Pr\left(c \rightarrow e; \frac{\max\{1, e^{4\sigma_x^2} - 1\}}{M}, \text{MSNR}\right), \end{aligned} \quad (32)$$

then the optimal design criterion, in terms of minimizing the pairwise probability of codeword error, is $A = (E_d/N)I$, because this design matrix meets the lower bound of η with equality. When (32) does not hold, and $e^{4\sigma_x^2} - 1 \geq 1$, then the optimal design criterion is to choose A to be all zero except for a single nonzero diagonal element. This design matrix, however, violates the CLT assumption that no single term dominates the summation in (8). Figure 3 shows the bounds on η and the probability of codeword error curve.

In the central limit theorem regime, the values of M and N must be large in order for $d^2(c, e)$ to be approximately Gaussian. We have also observed through numerical

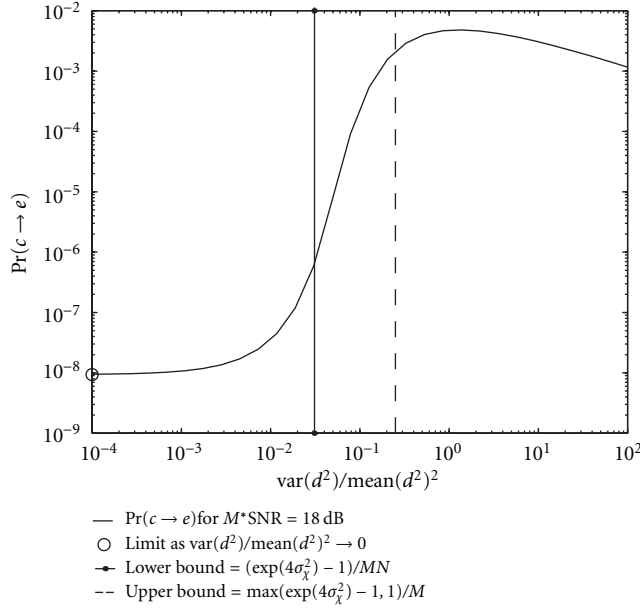


FIGURE 3: The probability of codeword error, $\Pr(c \rightarrow e; \eta^2, \rho)$, as a function of η^2 for $\rho = 18$ dB, with $M = N = 4$ and $\sigma_\chi^2 = 0.1$. The smallest achievable error probability occurs when $\eta^2 = (e^{4\sigma_\chi^2} - 1)/MN$, or equivalently, when $A = (E_d/N)I$.

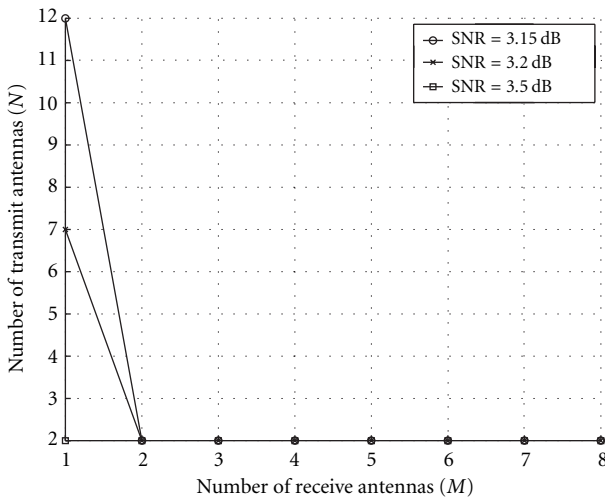


FIGURE 4: The smallest values of M and N such that (32) holds in mild fading ($\sigma_\chi^2 = 0.01$).

evaluation that the value of η that maximizes $\Pr(c \rightarrow e; \eta^2, \rho)$ increases with increasing ρ . As a result, increasing N and M cause the bounds on η given in (27) to become tighter, and the mode of $\Pr(c \rightarrow e; \eta^2, \rho)$ to increase until (32) eventually holds. These threshold values of M and N are plotted in Figures 4, 5, and 6 for different values of SNR and fading environments. For a given SNR, these plots show the smallest number of transmit and receive apertures required for $A = (E_d/N)I$ to be the optimal design matrix. From these

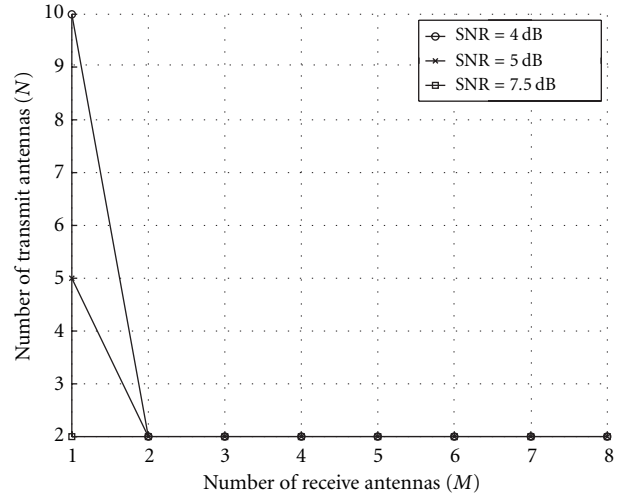


FIGURE 5: The smallest values of M and N such that (32) holds in moderate fading ($\sigma_\chi^2 = 0.1$).

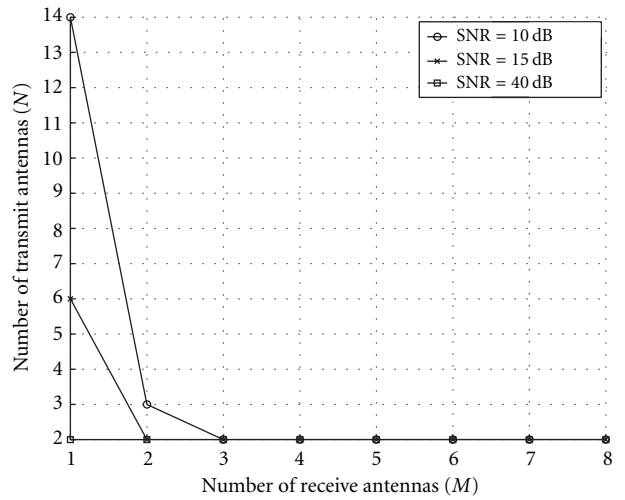


FIGURE 6: The smallest values of M and N such that (32) holds in severe fading ($\sigma_\chi^2 = 0.35$).

plots, we conclude that in the central limit theorem regime (large values of M and N), $A = (E_d/N)I$ is the optimal design matrix.

4. PERFORMANCE

In this section, we address the validity of the central limit theorem approximation and the performance of STCs on log-normal channels.

4.1. Performance bounds for orthogonal design STCs

We will now derive the pairwise probability of decoding codeword c as codeword e assuming that the space-time code satisfies the design criterion $A = (E_d/N)I$, but without using

the central limit theorem approximation. Under this design criterion, $d^2(c, e)$ becomes

$$d^2(c, e) = \sum_{m=1}^M \sum_{n=1}^N \frac{E_d}{N} |\alpha_{nm}|^2 = \frac{E_d}{N} \sum_{k=1}^{MN} e^{2\chi_k}, \quad (33)$$

where χ_k , $k = 1, \dots, MN$, are independent, identically distributed Gaussian random variables with $\text{var}(\chi_k) = \sigma_\chi^2$ and $E(\chi_k) = -\sigma_\chi^2$. Define $\chi = (\chi_1, \dots, \chi_{MN})$. The probability of decoding c as e is then

$$\Pr(c \rightarrow e) = \int_{-\infty}^{\infty} \Pr(c \rightarrow e | \chi) p_\chi(\chi) d\chi, \quad (34)$$

where $p_\chi(\chi)$ is the multivariate Gaussian probability density function for χ . Using the bound $Q(x) \leq \exp(-x^2/2)/2$ gives

$$\begin{aligned} \Pr(c \rightarrow e) &\leq \int_{-\infty}^{\infty} \frac{1}{2} \frac{1}{(2\pi\sigma_\chi^2)^{MN/2}} \\ &\quad \times \exp\left(-\frac{1}{N} \frac{E_d}{4N_0} \sum_{k=1}^{MN} e^{2\chi_k} - \frac{1}{2\sigma_\chi^2} \sum_{k=1}^{MN} (\chi_k + \sigma_\chi^2)^2\right) d\chi \\ &= \frac{1}{2} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_\chi^2}} \exp\left(-\frac{1}{N} \frac{E_d}{4N_0} e^{2x}\right) \right. \\ &\quad \left. \times \exp\left(-\frac{1}{2\sigma_\chi^2} (x + \sigma_\chi^2)^2\right) dx \right)^{MN} \\ &= \frac{1}{2} \left[\text{Fr}\left(\frac{1}{N} \frac{E_d}{4N_0}, 0; \sigma_\chi\right) \right]^{MN} \\ &= \frac{1}{2} \left[\text{Fr}\left(\frac{\text{SNR}}{4N}, 0; \sigma_\chi\right) \right]^{MN}, \end{aligned} \quad (35)$$

where $\text{Fr}(a, 0; b)$ is the lognormal density frustration function given by

$$\text{Fr}(a, 0; b) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi b^2}} \exp(-ae^{2x}) \exp\left(-\frac{1}{2b^2} (x + b^2)^2\right) dx. \quad (36)$$

Using the bound $Q(x) \geq \exp(-x^2)/4$, gives a similar lower bound

$$\Pr(c \rightarrow e) \geq \frac{1}{4} \left[\text{Fr}\left(\frac{\text{SNR}}{2N}, 0; \sigma_\chi\right) \right]^{MN}. \quad (37)$$

A closed form evaluation of the frustration function does not exist; therefore, we use a saddle-point integration method developed by Halme in [8] to numerically evaluate it. For the design criterion $A = (E_d/N)I$, Figure 7 compares the probability of codeword error in (34), the central

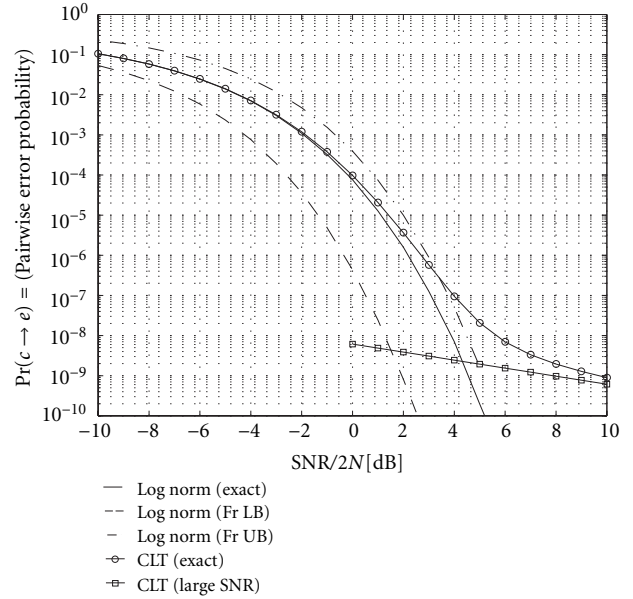


FIGURE 7: Comparison of pairwise codeword probability of error for $A = (E_d/N)I$ STCs with and without the central limit theorem approximation; $MN = 16$, $\sigma_\chi^2 = 0.1$.

limit theorem approximation probability of codeword error in (12), its asymptotic behavior in (29), and the frustration function bounds in (35) and (37). This figure shows that for small values of SNR, or typical values of error probability, the CLT approximation seems valid for $MN = 16$ in moderate fading. Asymptotically, however, the CLT probability of codeword error decays slower than the actual error probability. From (29), we know that the CLT probability of codeword error decays as $1/\text{SNR}$, whereas the frustration function bounds suggest the actual curve decays faster. This discrepancy arises from the dissimilarities in the tails of the Gaussian distribution and the actual distribution as emphasized by large values of SNR.

To measure the validity of the central limit theorem approximation, we examined the difference in SNR between the error probability expression in (34) and its approximation in (12) at a given error probability. For example, in Figure 7 for an error probability of $\Pr(c \rightarrow e) = 10^{-6}$, the CLT approximation requires 0.5 dB more SNR than the actual lognormal curve. Figure 8 shows this spurious SNR for different aperture products (MN) in different fading environments ($\sigma_\chi^2 = 0.01, 0.1, 0.35$). From Figure 8, we see that the CLT approximation is accurate to fractions of a dB in mild fading environments ($\sigma_\chi^2 = 0.01$) for all values of $MN \geq 2$. A larger number of apertures is required for more severe fading (roughly, $MN > 16$ for $\sigma_\chi^2 = 0.1$ and $MN > 64$ for $\sigma_\chi^2 = 0.35$).

4.2. A lower bound on the probability of codeword error

In Section 4.1, we derived lower and upper bounds on the probability of incorrectly decoding codeword c as codeword e under the design criterion $A = (E_d/N)I$ without using the

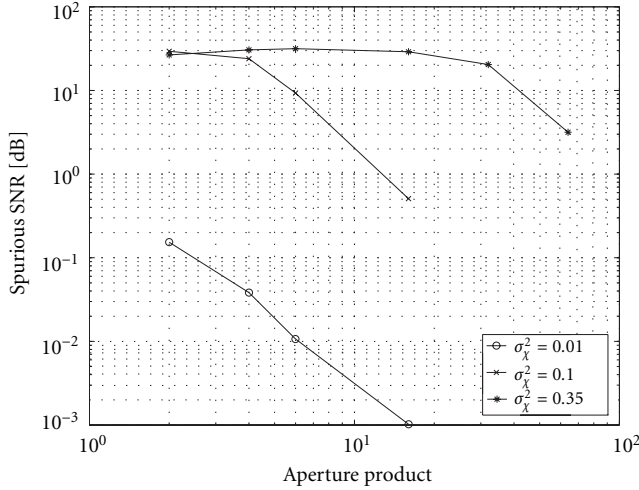


FIGURE 8: Difference in SNR required to achieve a 10^{-6} pairwise codeword error probability between the actual lognormal error expression in (34) and its central limit theorem approximation in (12).

central limit theorem approximation for $d^2(c, e)$. In this section, we derive a lower bound on this probability of error without using the central limit theorem approximation that is valid for an arbitrary design matrix A . Using the Cauchy-Schwarz inequality on (6) gives

$$\begin{aligned} d^2(c, e) &\leq \sum_{m=1}^M \sum_{t=1}^T \sum_{n=1}^N |\alpha_{nm}|^2 \sum_{k=1}^N |e_k(t) - c_k(t)|^2 \\ &= \sum_{m=1}^M \sum_{n=1}^N E_d |\alpha_{nm}|^2 \\ &= E_d \sum_{k=1}^{MN} e^{2\chi_k}, \end{aligned} \quad (38)$$

where we have renumbered the sum of the MN independent lognormal random variables as in Section 4.1. Following a similar derivation to that in Section 4.1, a lower bound on the probability of error for any design matrix is

$$\Pr(c \rightarrow e) \geq \frac{1}{4} \left[\text{Fr} \left(\frac{\text{SNR}}{2}, 0; \sigma_\chi \right) \right]^{MN}. \quad (39)$$

For a large number of transmit apertures, N , this bound can be quite loose, confer the orthogonal design bound in (37).

4.3. Infinite transmit diversity performance limit

If we fix the energy difference between codewords, E_d , and have enough receive apertures, M , such that (32) holds, then $A = (E_d/N)I$ minimizes the pairwise error probability, and this design matrix gives $\eta^2 = (\exp(4\sigma_\chi^2) - 1)/MN$. As we increase the number of transmit apertures, N , we see that η approaches zero, and hence (30) provides a performance limit

for infinite transmit diversity, that is,

$$\lim_{N \rightarrow \infty} \Pr \left(c \rightarrow e; \frac{e^{4\sigma_\chi^2} - 1}{MN}, \text{MSNR} \right) = Q \left(\sqrt{\frac{\text{MSNR}}{2}} \right). \quad (40)$$

These limits appear as circles in Figure 2 for $\text{MSNR} = 8, 13, 15, 18$ dB.

One can view this limit as the error probability of a one transmit, M receive aperture system with no fading. In other words, the large number of transmit apertures mitigates the fading, and the only uncertainty in the decision process arises from the additive white Gaussian noise.

4.4. An orthogonal design example: the Alamouti scheme

Alamouti in [9] proposed a simple transmit diversity technique using two transmit apertures ($N = 2$), two time-slots ($T = 2$), M receive apertures, and a complex QAM signal \mathcal{S} constellation of size 2^b . During the first time-slot, $2b$ bits arrive, determining two signal constellation points, s_1 and s_2 that are transmitted simultaneously on the first and second apertures, respectively. During the second time-slot, the first aperture transmits $-s_2^*$, while the second sends s_1^* . In other words, this STC consists of all the codewords of the form $c = (c_1(1), c_1(2), c_2(1), c_2(2)) = (s_1, -s_2^*, s_2, s_1^*)$ where s_1 and s_2 range over all possible signal constellation points. Tarokh in [2] showed that the Alamouti scheme is an example of a STC created from a complex orthogonal design.

The design matrix of this STC for two codewords $c = (c_1, -c_2^*, c_2, c_1^*)$ and $e = (e_1, -e_2^*, e_2, e_1^*)$ satisfies our design criteria $A = (E_d/2)I$, where $E_d = 2|c_1 - e_1|^2 + 2|c_2 - e_2|^2$ is the energy difference between the codewords. The performance of this code for pairs of codewords is shown in Figure 7 for eight receive apertures ($M = 8$) in moderate fading ($\sigma_\chi^2 = 0.1$).

Orthogonal designs have the property that the symbol sequences on each aperture are orthogonal, that is, $\sum_{t=1}^T c_n(t)c_k^*(t) = 0$ for $n \neq k$. As a result, space-time codes created from orthogonal designs, such as the Alamouti scheme, have a simple decoding algorithm. Rewriting the decision metric in (2) as

$$\begin{aligned} \hat{c} &= \underset{c \in \mathcal{S}^{NT}}{\text{argmin}} \sum_{m=1}^M \sum_{t=1}^T \left(|r_m(t)|^2 + \sum_{n=1}^N \sum_{k=1}^N \alpha_{nm} \alpha_{km}^* c_n(t) c_k^*(t) \right. \\ &\quad \left. - 2 \text{Re} \left\{ r_m^*(t) \sum_{n=1}^N \alpha_{nm} c_n(t) \right\} \right) \\ &= \underset{c \in \mathcal{S}^{NT}}{\text{argmin}} \sum_{n=1}^N \sum_{t=1}^T \left[\left(\sum_{m=1}^M |\alpha_{nm}|^2 \right) |c_n(t)|^2 \right. \\ &\quad \left. - 2 \text{Re} \left\{ \left(\sum_{m=1}^M r_m^*(t) \alpha_{nm} \right) c_n(t) \right\} \right] \end{aligned} \quad (41)$$

shows that joint detection of $(c_1(1), \dots, c_N(T))$ is equivalent to decoding each individual symbol, $c_n(t)$, separately. The

structure of the Alamouti STC allows for further simplification, and the decision rules become

$$\begin{aligned} \hat{s}_1 &= \operatorname{argmin}_{s \in \mathcal{S}} \left[\left(-1 + \sum_{n=1}^2 \sum_{m=1}^M |\alpha_{nm}|^2 \right) |s|^2 \right. \\ &\quad \left. + \left| s - \sum_{m=1}^M \left(r_m(1) \alpha_{1m}^* + r_m^*(2) \alpha_{2m} \right) \right|^2 \right], \\ \hat{s}_2 &= \operatorname{argmin}_{s \in \mathcal{S}} \left[\left(-1 + \sum_{n=1}^2 \sum_{m=1}^M |\alpha_{nm}|^2 \right) |s|^2 \right. \\ &\quad \left. + \left| s - \sum_{m=1}^M \left(r_m(1) \alpha_{2m}^* + r_m^*(2) \alpha_{1m} \right) \right|^2 \right]. \end{aligned} \quad (42)$$

5. CONCLUSIONS

In this paper, we presented a framework for developing space-time codes for an atmospheric optical heterodyne communication system. Through a central limit theorem approximation, we found that the design criterion $A = (E_d/N)I$ minimized the pairwise probability of codeword error for large numbers of apertures. Although developed for lognormal fading, this method generalizes to other fading distributions in which the fades are zero-mean and independent.

Our design criterion also satisfies the rank and determinant criteria presented in [1] for Rayleigh channels. Furthermore, orthogonal designs provide a method of constructing STCs that satisfy our criterion, and require only linear processing at the receiver [2].

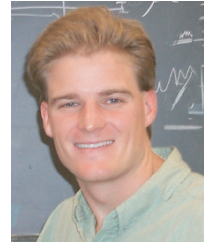
ACKNOWLEDGMENT

This research was supported in part by Defense Advanced Research Projects Agency Grant MDA972-00-1-0012.

REFERENCES

- [1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [3] J. H. Shapiro and R. C. Harney, "Simple algorithms for calculating optical communication performance through turbulence," in *Proc. SPIE Contr. and Commun. Tech. in Laser Sys.*, vol. 295, pp. 41–54, 1981.
- [4] J. H. Shapiro and R. C. Harney, "Burst-mode atmospheric optical communication," in *Proc. Nat. Telecommun. Conf. Record*, pp. 27.5.1–27.5.7, New York, NY, USA, 1980.
- [5] J. H. Shapiro, "Imaging and optical communication through atmospheric turbulence," in *Laser Beam Propagation in the Atmosphere*, J. W. Strohbehn, Ed., chapter 6, Springer, Berlin, Germany, 1978.
- [6] R. M. Gagliardi and S. Karp, *Optical Communications*, John Wiley & Sons, New York, NY, USA, 1976.
- [7] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, NY, USA, 1988.
- [8] S. J. Halme, B. K. Levitt, and R. S. Orr, "Bounds and approximations for some integral expression involving lognormal statistics," Technical report, MIT Res. Lab. Electron. Quart. Prog. Rept., 1969.
- [9] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

Shane M. Haas is a Ph.D. candidate in Electrical Engineering at the Massachusetts Institute of Technology. He received B.S. and M.S. degrees in Electrical Engineering in 1998, 1999 and B.S. and M.A. degrees in Mathematics from the University of Kansas with honors and highest distinction in 1998, 1999, respectively. His doctoral research concerns the theoretical and experimental study of coding for and the capacity of optical communication through the turbulent atmosphere. Mr. Haas is a National Defense Science and Engineering Graduate Fellow, National Science Foundation Fellow, and a Goldwater Scholar. In 1998, he was named the Outstanding Senior in both the University of Kansas School of Engineering and the Department of Electrical Engineering and Computer Science. As an undergraduate, he was supported by the National Science Foundation Research Experience for Undergraduates program for four consecutive years. From 1997 to 2000, he worked at Flint Hills Scientific, L.L.C., developing an adaptive digital signal processing algorithm to predict and detect the onset of epileptic seizures. He is currently a graduate research and teaching assistant at the Laboratory of Information and Decision Systems and the Research Laboratory of Electronics at MIT.



Jeffrey H. Shapiro received the S.B., S.M., E.E., and Ph.D. degrees in Electrical Engineering from the Massachusetts Institute of Technology in 1967, 1968, 1969, and 1970, respectively. As a graduate student he was a National Science Foundation Fellow, a Teaching Assistant, and a Fannie and John Hertz Foundation Fellow. His doctoral research was a theoretical study of adaptive techniques for improved optical communication through atmospheric turbulence. From 1970 to 1973, Dr. Shapiro was an Assistant Professor of Electrical Sciences and Applied Physics at Case Western Reserve University. From 1973 to 1985, he was an Associate Professor of Electrical Engineering at MIT, and in 1985, he was promoted to Professor of Electrical Engineering. From 1989 until 1999 Dr. Shapiro served as Associate Department Head. In 1999, he became the Julius A. Stratton Professor of Electrical Engineering. In 2001, Dr. Shapiro was appointed director of MIT's Research Laboratory of Electronics. Dr. Shapiro's research interests have centered on the application of communication theory to optical systems. He is best known for his work on the generation, detection, and application of squeezed-state light beams, but he has also published extensively in the areas of atmospheric optical communication and coherent laser radar. Dr. Shapiro is a fellow of the Institute of Electrical and Electronics Engineers, of the Optical Society of America, and of the Institute of Physics, and he is a member of SPIE (The International Society for Optical Engineering). He has been an Associate Editor of the *IEEE Transactions on Information Theory* and the *Journal of the Optical Society of America*.



Vahid Tarokh received his Ph.D. degree in Electrical Engineering from the University of Waterloo, Ontario, Canada in 1995. From August 1995 to May 1996, he was employed by the Coordinated Science Laboratory, of the University of Illinois Urbana-Champaign, as a visiting Professor. He then joined the AT&T Labs-Research, where he was employed as a Senior Member of Technical Staff, Principal Member of Technical Staff, and the Head of the Department of Wireless Communications and Signal Processing until August 2000. In the fall of 2000, Dr. Tarokh joined Department of Electrical Engineering and Computer Sciences of MIT as an Associate Professor, where he is currently employed. Dr. Tarokh received numerous awards including the 1987 Gold Tablet of The Iranian Math Society, the 1995 Governor General of Canada's Academic Gold Medal, the 1999 IEEE Information Theory Society Prize Paper Award, (jointly with A. R. Calderbank and N. Seshadri), and more recently the 2001 Alan T. Waterman Award.

