Efficient Space-Time Block Codes Designed by a Genetic Algorithm

Don Torrieri
U.S. Army Research Laboratory
dtorr@arl.army.mil

Matthew C. Valenti
West Virginia University
mvalenti@csee.wvu.edu

Abstract—Space-time block codes with orthogonal structures typically provide full-diversity reception and simple receiver processing. However, rate-1 orthogonal codes for complex constellations have not been found for more than two transmit antennas. By using a genetic algorithm, rate-1 space-time block codes that accommodate very simple receiver processing at the cost of reduced diversity are designed. Simulation results show that when the spectral efficiency is specified and an efficient outer code is used, the evolved codes provide better performance over the Rayleigh fading channel than minimum-decoding-complexity quasi-orthogonal space-time block codes at typical operating signal-to-noise ratios. When the fading is more severe than Rayleigh fading, evolved codes outperform orthogonal space-time block codes.

I. Introduction

Space-time codes transmitted by multiple antennas improve the performance of a communication system in a fading environment without the need for multiple receive antennas or channel-state information at the transmitter [1, 2]. An orthogonal space-time block code (STBC), such as the Alamouti code, provides full diversity at full transmission rate and maximum-likelihood decoding that entails only linear processing. However, rate-1 orthogonal STBCs for complex constellations exist only for two transmit antennas. Orthogonal STBCs for more than two transmit antennas require a code rate that is less than unity [2, 3], which implies a reduced spectral efficiency.

When used with rotated constellations, quasi-orthogonal (QO) [4–6] and coordinate-interleaved [7] STBC can provide full diversity at full rate but require more complex decoding than the decoupled decoding of each real-valued symbol that is possible with orthogonal STBCs. In this paper, rate-1 linear dispersion codes [8] that are nearly orthogonal STBCs are generated by a genetic algorithm. Since the evolved codes are nearly orthogonal, simple suboptimal decoupled decoding that would be optimal for orthogonal codes causes only a minor reduction of diversity gain and, hence, a performance loss only at high signal-to-noise ratios. The minimum-decoding-complexity quasi-orthogonal (MDC-QO) STBCs [5] offer full diversity and a maximum-likelihood decoder that can decouple complex symbols. Compared with these codes, the evolved codes have a simplified implementation in both the decoder (decoupling of real symbols) and the modulator (no constellation rotation). As shown subsequently, when the spectral efficiency is specified, an efficient outer code is used, and the fading is severe, the evolved codes provide better performance than both rate-1 MDC-QO and orthogonal space-time block codes.

II. STBC Requirements

Let N_T denote the number of transmit antennas, N denote the number of distinct transmitted constellation symbols, and L denote the length of a space-time codeword. The space-time code rate is N/L, the number of information symbols conveyed per signaling interval. A transmitted symbol that belongs to a complex signal constellation is represented as

$$x_n = x_n^r + jx_n^i, \ 1 \le n \le N \tag{1}$$

where x_n^r and x_n^i are the real and imaginary components of the symbol, respectively, and $j = \sqrt{-1}$. Assuming that the original constellation may be transformed, the $L \times N_T$ generator matrix representing a transmitted codeword of a linear dispersive STBC may be expanded as

$$\mathbf{G} = \sum_{n=1}^{N} a_n \mathbf{A}_n + j \sum_{n=1}^{N} b_n \mathbf{B}_n \tag{2}$$

where each a_n and b_n is a real or imaginary component of a symbol and the *dispersion matrices* $\mathbf{A}_n$ and $\mathbf{B}_n$ have real-valued elements. Each space-time encoder linearly combines weighted real and imaginary components of N symbols and produces a sequence of L complex numbers. Each real and imaginary component of each of these complex numbers becomes a separate in-phase and quadrature component, respectively, of a modulated carrier.

Example 1: Consider the MDC-QO rate-1 STBC that maps $N = 4$ information symbols into $N_T = 4$ antenna outputs over $L = 4$ signaling intervals [5]. The coefficients in (2) are $a_1 = x_1^r$, $a_2 = x_2^r$, $a_3 = x_1^i$, $a_4 = x_2^i$, $b_1 = x_3^r$, $b_2 = x_4^r$, $b_3 = x_3^i$, and $b_4 = x_4^i$. The generator matrix is

$$\mathbf{G} = \begin{bmatrix} x_1^r + jx_3^r & x_2^r + jx_4^r & -x_1^i + jx_3^i & -x_2^i + jx_4^i \\ -x_2^r + jx_4^r & x_1^r - jx_3^r & x_2^i + jx_4^i & -x_1^i - jx_3^i \\ -x_1^i + jx_3^i & -x_2^i + jx_4^i & x_1^r + jx_3^r & x_2^r + jx_4^r \\ x_2^i + jx_4^i & -x_1^i - jx_3^i & -x_2^r + jx_4^r & x_1^r - jx_3^r \end{bmatrix}.$$

The dispersion matrices are

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{A}_3 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \mathbf{B}_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{A}_4 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{B}_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \quad \blacksquare$$

Only one receive antenna is assumed for simplicity, and the extension of the analysis to multiple receive antennas is straightforward. The complex channel response to transmit antenna ℓ at the sampled demodulator output is

$$h_\ell = h_\ell^r + jh_\ell^i, \ 1 \le \ell \le N_T \tag{3}$$

where h_ℓ^r and h_ℓ^i are the real and imaginary components of h_ℓ, respectively. The channel responses are assumed to be known at the receiver and constant for L symbol periods. Let $\mathbf{z}_n$ denote the $2L \times 1$ vector with its first L elements equal to the real–parts of the responses to $a_n, 1 \le n \le N$, or b_{n-N}, $N + 1 \le n \le 2N$, and its second L elements equal to the imaginary-parts of the responses to $a_n, 1 \le n \le N$, or b_{n-N}, $N + 1 \le n \le 2N$. Thus,

$$\mathbf{z}_n = a_n \mathcal{A}_n \mathbf{h}, \ 1 \le n \le N \tag{4}$$

$$\mathbf{z}_n = b_{n-N} \mathcal{B}_{n-N} \mathbf{h}, \ N + 1 \le n \le 2N \tag{5}$$

where $\mathbf{h} = [h_1^r \ h_2^r \cdots h_{N_T}^r \ h_1^i \ h_2^i \cdots h_{N_T}^i]^T$ and $\mathcal{A}_n$ and $\mathcal{B}_n$ are $2L \times 2N_T$ matrices. They have the forms

$$\mathcal{A}_n = \begin{bmatrix} \mathbf{A}_n & \mathbf{0}_{L \times N_T} \\ \mathbf{0}_{L \times N_T} & \mathbf{A}_n \end{bmatrix}, \quad \mathcal{B}_n = \begin{bmatrix} \mathbf{0}_{L \times N_T} & -\mathbf{B}_n \\ \mathbf{B}_n & \mathbf{0}_{L \times N_T} \end{bmatrix}. \tag{6}$$

Let $\mathbf{d} = [a_1 \ a_2 \cdots a_N \ b_1 \ b_2 \ldots b_N]^T$. The $2L \times 1$ vector of received symbols in the absence of noise is

$$\mathbf{y}_s = \sum_{n=1}^{2N} \mathbf{z}_n = \mathbf{H}\mathbf{d} \tag{7}$$

where

$$\mathbf{H} = [\mathcal{A}_1 \mathbf{h} \ \mathcal{A}_2 \mathbf{h} \cdots \mathcal{A}_N \mathbf{h} \ \mathcal{B}_1 \mathbf{h} \ \mathcal{B}_2 \mathbf{h} \cdots \mathcal{B}_N \mathbf{h}] \tag{8}$$

is the $2L \times 2N$ channel matrix.

For full-diversity reception and simple (decoupled) maximum-likelihood processing in the receiver, an *orthogonality condition* must be satisfied:

$$\mathbf{H}^T \mathbf{H} = ||\mathbf{h}||^2 \mathbf{I}_{2N \times 2N} \tag{9}$$

for any vector $\mathbf{h}$ with real-valued components, where $||\mathbf{h}||$ denotes the Euclidean norm of $\mathbf{h}$. The orthogonality condition is satisfied if and only if

$$\mathbf{A}_n^T \mathbf{A}_n = \mathbf{B}_n^T \mathbf{B}_n = \mathbf{I}_{N_T \times N_T}, 1 \le n \le N \tag{10}$$

$$\mathbf{A}_n^T \mathbf{A}_\ell + \mathbf{A}_\ell^T \mathbf{A}_n = \mathbf{B}_n^T \mathbf{B}_\ell + \mathbf{B}_\ell^T \mathbf{B}_n = \mathbf{0}_{N_T \times N_T},$$
$$n \neq \ell, \ 1 \le n, \ell \le N \tag{11}$$

$$\mathbf{A}_n^T \mathbf{B}_\ell = \mathbf{B}_\ell^T \mathbf{A}_n, \ n, \ell \le N. \tag{12}$$

Appendix A gives a proof using linear algebra that $(10) - (12)$ are necessary and sufficient for the orthogonality condition. An alternative, more complicated proof, may be derived from the theory of amicable orthogonal designs [Appendix B of 2]. Equation (10) also ensures that the total power transmitted for symbol-component a_n or b_n by each transmit antenna is equal to a_n^2 or b_n^2, respectively. Thus, the total power transmitted by each antenna by each transmit antenna is equal. The dispersion matrices of Example 1 do not satisfy (11) and are therefore not orthogonal.

Complex orthogonal designs that restrict the matrix elements to ± 1 only exist for $N_T = L = 2$ or 4. Generalized complex orthogonal designs that restrict the elements of the generator matrix to the alphabets $\{\pm 1, 0\}$ or $\{\pm 1, \pm 1/\sqrt{2}\}$ are known, but only for code rates less than unity. A benefit of a genetic approach to design is that the element-value set does not need to be constrained to a finite alphabet. For instance, the elements could be drawn from a continuous distribution. The advantage of an expanded element-value set, which is used by the genetic algorithm, is that a good STBC may be found that is close to being orthogonal. The approximate orthogonality may allow the low-complexity decoupled detector for an orthogonal design to be retained while still obtaining a satisfactory performance. In contrast to most orthogonal STBCs, linear dispersion codes are not restricted to complex symbols and their complex conjugates with the same magnitudes.

Let $\mathcal{E}_s$ denote the average energy per symbol transmitted by all the transmit antennas. In the presence of white Gaussian noise with spectral-density N_0, the vector of received symbols is

$$\mathbf{y}_r = \mathbf{H}\mathbf{d} + \mathbf{n} \tag{13}$$

where $\mathbf{n}$ is the zero-mean noise with covariance matrix $E[\mathbf{n}\mathbf{n}^T] = N_0 \mathbf{I}$ and $E[|d_n|^2] = \mathcal{E}_s/N_T$ for each component d_n of $\mathbf{d}$. If the columns of $\mathbf{H}$ are orthogonal, then the receiver computes the $2N \times 1$ vector

$$\mathbf{y} = ||\mathbf{h}||^{-2} \mathbf{H}^T \mathbf{y}_r = \mathbf{d} + \mathbf{n}_1 \tag{14}$$

where $E[\mathbf{n}_1 \mathbf{n}_1^T] = N_0 ||\mathbf{h}||^{-2} \mathbf{I}$. The maximum-likelihood decision for d_n is separately obtained by minimizing $|y_n - d_n|$, $n = 1, 2, \ldots, 2N$, and hence is decoupled from the other component decisions. If the columns of $\mathbf{H}$ are not orthogonal, then the computation of $||\mathbf{h}||^{-2} \mathbf{H}^T \mathbf{y}_r$ is suboptimal and produces $\mathbf{y} = \mathbf{d} + \mathbf{z} + \mathbf{n}_1$, where $\mathbf{z}$ is inter-symbol and same-symbol

interference that reduces the diversity gain. This interference diminishes as $\mathbf{H}$ approaches orthogonality.

The evolved codes are decoded using the same decoupled decoder used with orthogonal codes, and therefore enjoy the same low complexity. As an example, consider a rate-1 code for $N_T = 4$ antennas using square quadrature amplitude modulation (QAM) with M symbols. The decoupled decoder independently detects each of the $2L = 8$ real-valued symbols, and each decision requires the comparison of a real value against $\sqrt{M}$ hypotheses. The decoder for a quasi-orthogonal code [6] independently detects two pairs of complex symbols, and each decision requires the comparison of a pair of complex values against M^2 hypothesis. Thus, quasi-orthogonal codes requires a factor of $M\sqrt{M}$ more real-valued comparisons than the proposed evolved codes. The MDC-QO codes [4, 5] allow the $L = 4$ complex symbols to be independently detected, and each decision requires the comparison of a complex value against M hypothesis. Thus, the MDC-QO code requires a factor of $\sqrt{M}$ more real-valued comparisons than the evolved codes.

III. DESCRIPTION OF THE GENETIC ALGORITHM

One of the most critical factors in designing a genetic algorithm [9] is the choice of the cost or fitness function. To optimize the set of dispersion matrices $\{\mathbf{A}_n, \mathbf{B}_n\}$, the cost function was defined as

$$
C = c_1 \sum_{n=1}^{N} \left[\left\| \mathbf{A}_n^T \mathbf{A}_n - I \right\|^2 + \left\| \mathbf{B}_n^T \mathbf{B}_n - I \right\|^2 \right]
$$
$$
+ c_2 \sum_{n=1}^{N} \sum_{\substack{\ell=1 \\ \ell \neq n}}^{N} \left[\left\| \mathbf{A}_n^T \mathbf{A}_\ell + \mathbf{A}_\ell^T \mathbf{A}_n \right\|^2 + \left\| \mathbf{B}_n^T \mathbf{B}_\ell + \mathbf{B}_\ell^T \mathbf{B}_n \right\|^2 \right]
$$
$$
+ c_3 \sum_{n=1}^{N} \sum_{\ell=1}^{N} \left\| \mathbf{A}_n^T \mathbf{B}_\ell - \mathbf{B}_\ell^T \mathbf{A}_n \right\|^2 \tag{15}
$$

where $\| \cdot \|$ denotes the Frobenius norm, and c_1, c_2, and c_3 are constants. This cost function penalizes STBC designs with matrices that do not satisfy the orthogonality conditions of $(10) - (12)$.

The algorithm is initialized with a random population of K individuals $\mathcal{X} = \{\mathbf{x}_1, ..., \mathbf{x}_K\}$. Each $\mathbf{x}_k$ is a string of *genes* that specifies the entries of the $\{\mathbf{A}_n, \mathbf{B}_n\}$ matrices of a particular STBC. The entries of $\{\mathbf{A}_n, \mathbf{B}_n\}$ may be drawn from a continuous alphabet or a discrete alphabet. If drawn from a continuous alphabet, then each gene of $\mathbf{x}_k$ corresponds to one entry of a $\mathbf{A}_n$ or $\mathbf{B}_n$ matrix, and therefore there are $2NN_TL$ genes in a string. If drawn from a discrete alphabet, then each gene of $\mathbf{x}_k$ corresponds to an entire column of a $\mathbf{A}_n$ or $\mathbf{B}_n$ matrix, and therefore there are only $2NN_T$ genes in a discrete-alphabet string.

Strings are normalized to force each column of the $\{\mathbf{A}_n, \mathbf{B}_n\}$ matrices have unit norm. For continuous-alphabets, this normalization is performed by multiplying each element of a column by an appropriately chosen scaling factor. The normalization occurs when the string is first randomly generated or whenever it evolves. For discrete-alphabets, normalization

is accomplished by constraining the set of allowable genes to only those columns that have unit norm.

In the continuous-alphabet case, each $\mathbf{x}_k$ is generated such that its entries are independent and uniformly distributed over $(-1, +1)$ and then the columns are normalized. In the discrete-alphabet case, the $\mathbf{x}_k$ are generated by selecting each column from the set of unit-norm columns with equal probability. The cost of each member of the initial population is determined according to (15), and the population is sorted in ascending order of cost.

During each generation, two parents $\mathbf{x}_j$ and $\mathbf{x}_k$, $k > j$, are selected. Four parent selection strategies are considered: (1) *random selection*, (2) *preferred parenting*, (3) *eugenic selection*, and (4) *alpha-male selection*. With random selection, two distinct individuals are picked at random from the population. With preferred parenting, $\mathbf{x}_k$ is picked at random from the entire population except for the best (i.e., $2 \leq k \leq K$), while $\mathbf{x}_j$ is selected at random from those mappings that are better than the first (i.e., $1 \leq j \leq k$). With eugenic selection, the best two individuals are selected: $\mathbf{x}_j = \mathbf{x}_1$ and $\mathbf{x}_k = \mathbf{x}_2$. With alpha-male selection, the first parent is the best individual $\mathbf{x}_j = \mathbf{x}_1$, while the second parent $\mathbf{x}_k$ is selected at random from the rest of the population (i.e., $2 \leq k \leq K$).

The two selected parents *breed* a pair of children $\{\tilde{\mathbf{x}}_j, \tilde{\mathbf{x}}_k\}$. Each child has a *direct* ancestor, which is the parent with the same index, and an *indirect* ancestor, which is the other parent. The genes of each child are identical to those of its direct ancestor except at randomly chosen *crossover* positions, where the genes are equal to those of its indirect ancestor. The crossover positions are chosen at random such that crossovers are independent and occur with probability p_x. In the continuous-alphabet case, columns are normalized after breeding.

After breeding, *Q mutants* are created for each child. The genes of a mutant are identical to that of the child from which it is derived except at randomly chosen *mutation* positions, where the gene of the child is discarded and replaced with a randomly generated gene. The mutation positions are independent and occur with probability p_m. The random genes are uniformly distributed over $(-1, 1)$ in the continuous-alphabet case or over the set of unit-norm columns in the discrete-alphabet case. In the continuous-alphabet case, columns are normalized after mutating.

Next, the algorithm randomly chooses between two replacement strategies: *normal replacement* and *culling*. Each time an individual is replaced, the population is re-sorted in ascending order of cost. With normal replacement, each of the two parents is replaced by its best offspring if that offspring has a lower cost than its parent. When culling, the best offspring of each parent is compared with the cost of $\mathbf{x}_K$, the worst of the population. If the offspring is better than $\mathbf{x}_K$, then the latter is replaced. Culling provides an opportunity to periodically eliminate poor designs and replace them with designs that are more fit. In the following discussion, p_c is defined to be the probability that culling is used during a particular generation.

IV. OPTIMIZATION RESULTS

The genetic algorithm was used to produce rate-1 designs for $N_T = 3$ and 4 antennas with several values of $N = L$ and both continuous and discrete alphabets. In each case, the genetic algorithm was run with cost vector $\mathbf{c} = [c_1, c_2, c_3] = [1, 1, 1]$ and $Q = 2$ mutations per child. When a discrete alphabet was used, the alphabet was $\{0, \pm 1, \pm\sqrt{3}/2, \pm\sqrt{2}/2, \pm 1/2\}$, and the algorithm was run with crossover probability $p_x = 1/(2NN_T)$, mutation probability $p_m = 1/(2NN_T)$, culling probability $p_c = 0.01$, population size $K = 400$, and preferred parenting. When a continuous alphabet was used, the algorithm was run with $p_x = 0.05$, $p_m = 1/(2NN_T L)$, $p_c = 0.1$, $K = 200$, and alpha-male selection. These parameters were selected after exploring many other alternatives, and concluding that these were the most effective parameters of those that were evaluated. Furthermore, in the continuous-alphabet case, the first row of $\mathbf{A}_1$ and $\mathbf{B}_1$ were constrained such that only the first entry of the row could be nonzero while the remaining $N_T - 1$ entries were forced to zero. This constraint is motivated by observing that multiplication of a channel matrix $\mathbf{H}$ by an orthogonal matrix gives a new channel matrix with the same product $\mathbf{H}^T\mathbf{H}$, and hence degree of orthogonality, as the original one. Thus, there are many STBCs with the lowest cost, and constraining either a pair of rows or a pair of columns will not eliminate all of the lowest-cost STBCs while significantly reducing the search space of the genetic algorithm.

Let (a, b, c) denote a code with the parameter values $N = a$, $N_T = b$, and $L = c$. Fig. 1 shows the cost of the best design per generation for a $(6, 3, 6)$ code and a $(4, 3, 4)$ code. The $(4, 3, 4)$ code was evolved using both continuous and discrete alphabets, while the $(6, 3, 6)$ code was evolved only using a continuous alphabet because a good design with a discrete alphabet could not be found. After 1 million generations, the cost of the $(4, 3, 4)$ code converged to 12.01 with the continuous alphabet and 12.00 with the discrete alphabet, and the cost of the $(6, 3, 6)$ code converged to 18.00. The dispersion matrices for the discrete-alphabet $(4, 3, 4)$ code and the $(6, 3, 6)$ code after 1 million generations of evolution are given in Appendix B. Although the discrete-alphabet used by the genetic algorithm in generating the $(4, 3, 4)$ code has 9 symbols, only 7 of these symbols appear in the dispersion matrices of Appendix B. Despite not appearing in the final design, the 2 missing symbols appeared in designs as late as 400,000 generations.

Fig. 2 shows the bit error rate (BER) performance for the discrete-alphabet $(4, 3, 4)$ code by simulating it over a Rayleigh fading channel with quadriphase shift keying (QPSK). The simulated fading coefficients are constant for blocks of L symbols, but independent from block to block and have zero-mean, unit-variance, complex Gaussian distributions. The figure shows the BER performance of both maximum-likelihood (ML) and decoupled decoding for the best designs obtained after 100 thousand and 1 million generations.
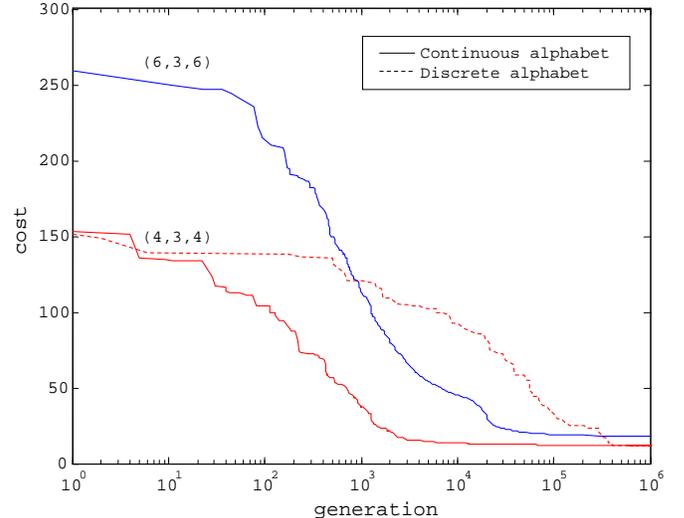
Fig. 1. Cost as a function of generation of evolved $(4, 3, 4)$ and $(6, 3, 6)$ codes.

Also shown in Fig. 2 is the performance of the $(4, 3, 4)$ linear dispersion code presented in [8]. As can be seen, the reference code outperforms the evolved code when ML decoding is used. The reason is that the cost function of (15) is not designed to optimize ML performance. However, when the reference code is detected with the simple decoupled decoder, performance is poor and exhibits a relatively high error floor. The reason is that the cost of the reference code is $C = 64c_2$, which is significantly higher than the costs of the evolved designs. The designs obtained during early generations of the genetic algorithm also exhibit an error floor with decoupled decoding. However, as the design becomes more highly evolved, the decoupled-decoding error floor is lowered. After 1 million generations, the performances of ML and decoupled decoding are identical down to at least a BER of 10^{-6}. While the decoupled performance improves as the design becomes more highly evolved, the ML performance actually degrades as the design evolves.

Fig. 3 shows the BER of the continuous-alphabet $(6, 3, 6)$ code with QPSK over a Rayleigh fading channel after 10 thousand, 100 thousand, and 1 million generations. Also shown is the $(6, 3, 6)$ reference linear dispersion code from [8]. Similar to the $(4, 3, 4)$ case, the decoupled-performance of the $(6, 3, 6)$ code improves as the code evolves. Meanwhile, the ML performance degrades as it evolves. After 1 million generations, the ML and decoupled performance are identical for the range of BER that was simulated.

Fig. 4 compares the BERs over a Rayleigh fading channel of the $(4, 3, 4)$ and $(6, 3, 6)$ evolved codes listed in Appendix B against a rate-3/4 orthogonal $(3, 3, 4)$ code presented in [2] and the MDC-QO $(4, 3, 4)$ code created by deleting the last columns of the dispersion matrices given in Example 1. In order to provide a fair comparison, the spectral efficiency is maintained at 3 bits/s/Hz by using 8-phase shift keying (8-PSK) for the rate-1 codes and 16-QAM for the rate-3/4 orthogonal code. For the MDC-QO code, the 8-PSK constellation is rotated by 5.915 degrees, which minimizes
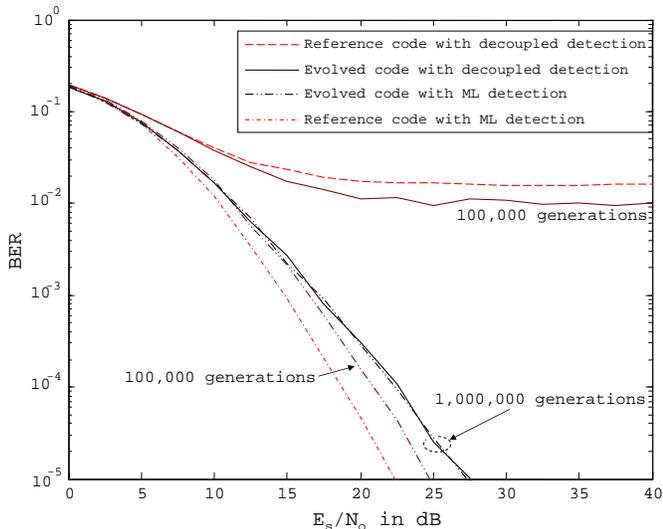
Fig. 2. BERs over a Rayleigh fading channel for ML and decoupled decoding of the $(4, 3, 4)$ evolved code after different numbers of generations. The $(4, 3, 4)$ code from [8] is the "reference code".
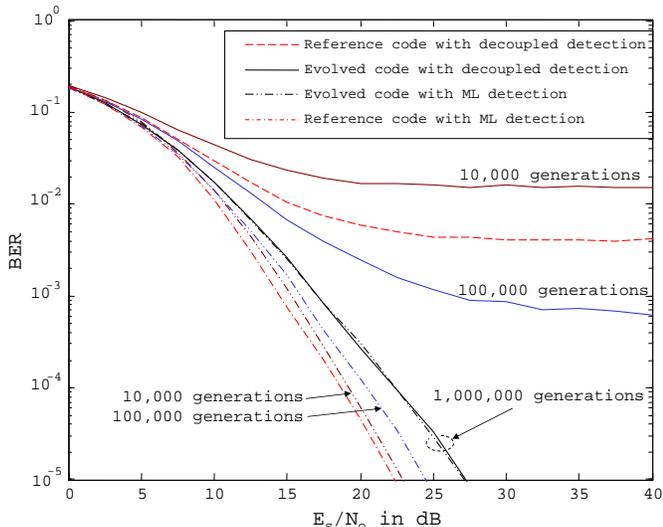
Fig. 3. BERs over a Rayleigh fading channel for ML and decoupled decoding of the $(6, 3, 6)$ evolved code after different numbers of generations. The $(6, 3, 6)$ code from [8] is the "reference code".

Fig. 4. BERs over a Rayleigh fading channel of $(4, 3, 4)$ and $(6, 3, 6)$ codes with 8-PSK and orthogonal $(3, 3, 4)$ code with 16-QAM.
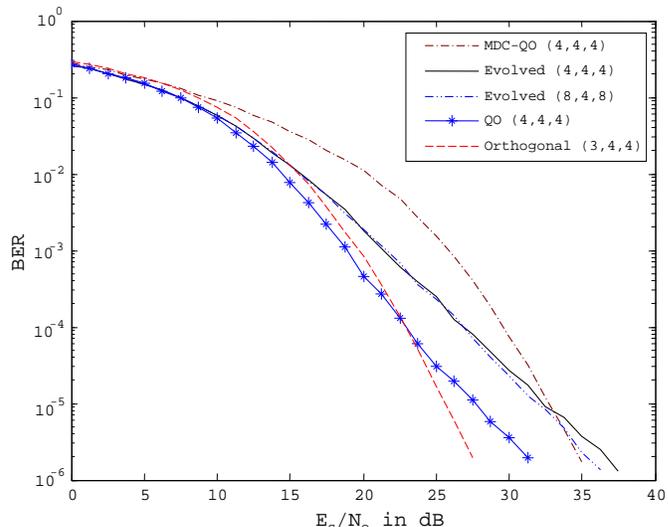
Fig. 5. BERs over a Rayleigh fading channel of evolved $(4, 4, 4)$ and $(8, 4, 8)$ codes and $(4, 4, 4)$ MDC-QO code with 8-PSK and orthogonal $(3, 4, 4)$ code with 16-QAM.

the symbol error rate [4]. As can be seen, the evolved rate-1 codes with 8-PSK are better than the orthogonal code with 16-QAM for $\mathcal{E}_s/N_0 < 18.5$ dB and $BER > 3 \times 10^{-3}$, which indicates that full diversity is advantageous primarily for high signal-to-noise ratios. The two evolved codes have nearly identical performance, suggesting that the $(4, 3, 4)$ code would be a more appropriate choice since its shorter length requires less processing, incurs less latency, and can be used over a more rapidly fading channel. Furthermore, the evolved rate-1 codes outperform the MDC-QO code for $\mathcal{E}_s/N_0 < 37$ dB and $BER > 10^{-6}$.

Space-time codes were also evolved using continuous alphabets and $N_T = 4$ transmit antennas. Fig. 5 shows the performance over a Rayleigh fading channel of $(4, 4, 4)$ and $(8, 4, 8)$ codes with 8-PSK. These codes were each evolved for 10 million generations and attained costs of 32.01 and 64.00,

respectively. Two additional $(4, 4, 4)$ codes were evolved using discrete alphabets and attained costs of 32.00. However, these codes exhibited a high BER floor when simulated.

The $(8, 4, 8)$ code evolved in parallel on a cluster of four computers. Each computer was seeded with the same initial population and was allowed to evolve using its own gene pool. Periodically, the populations were compared and if there was one individual that was superior to the best designs attained on the other computers, then that individual was copied to the other computers to become part of the other gene pools. This process of *cloning and immigration* from one gene pool to another was a key to attaining a low cost.

Fig. 5 also shows the performance of the $(4, 4, 4)$ quasi-orthogonal code of [6] with 8-PSK, the $(4, 4, 4)$ MDC-QO code from Example 1 with 5.915-degree rotated 8-PSK, and a rate-3/4 $(3, 4, 4)$ orthogonal code from [2] with 16-QAM. The evolved codes outperform the MDC-QO code for $BER >$

10^{-5} and outperform the orthogonal code for $BER > 10^{-2}$. The evolved codes show a loss relative to the quasi-orthogonal code, but are far simpler to decode since the latter requires pairs of complex symbols to be jointly decoded [6].

By comparing the performance of the $(3, 4, 4)$ orthogonal code with that of the $(3, 3, 4)$ orthogonal code shown in Fig. 4, it is apparent that the extra antenna provides more diversity in Rayleigh fading when orthogonal codes are used. However, the evolved codes do not show the increase in diversity. The performance of the full-rate system with 4 transmit antennas is nearly identical to that achieved with 3 transmit antennas and also to that achieved with 2 transmit antennas and the full-rate $(2, 2, 2)$ Alamouti orthogonal code [2].

While the evolved designs cannot exploit the added potential diversity when the fading is Rayleigh, they still outperform the orthogonal systems when the fading is more severe than Rayleigh [10] and an outer turbo code is used. Fig. 6 shows the coded performance of the evolved full-rate $(4, 4, 4)$ STBC with 8-PSK modulation, the rate-3/4 orthogonal $(3, 4, 4)$ STBC with 16-QAM, and the full-rate orthogonal Alamouti $(2, 2, 2)$ STBC with 8-PSK modulation. The fading is either Rayleigh or the more severe Nakagami-m [1] with $m = 1/2$. The outer code is a rate-1/2 turbo code specified by the UMTS third-generation cellular standard [11]. The data message is 4500 bits long, and the channel gains are assumed to be constant for the duration of one space-time codeword of L signaling intervals but independent from one codeword to the next. The turbo code is decoded using 14 iterations of the log-MAP algorithm [1]. The turbo decoder uses soft decisions produced by the decoupled STBC decoder in the form of the conditional density $p(x_n|\mathbf{y})$, where x_n and $\mathbf{y}$ are given by (1) and (14), respectively.

As seen in the figure, the full-rate codes both provide a gain of at least 1.4 dB over the rate-3/4 orthogonal codes when the fading is Rayleigh. This gain occurs because the turbo code allows operation at $\mathcal{E}_s/N_0$ lower than the 18.5 dB breakpoint where the rate-3/4 orthogonal designs become superior to the full-rate designs. Simulations of many designs generated by the genetic algorithm indicate that it is highly probable that regardless of the SNR and outer code, no rate-1 STBC with decoupled decoding exists such that its BER is lower over the Rayleigh channel than that of the Alamouti $(2, 2, 2)$ STBC. However, the figure indicates that over the Nakagami-1/2 channel, the evolved $(4, 4, 4)$ STBC outperforms the Alamouti and the rate-3/4 orthogonal STBCs by 0.7 dB and 1.5 dB, respectively.

V. CONCLUSIONS

A genetic algorithm has been designed to produce rate-1 space-time block codes optimized for three or more transmit antennas and decoupled decoding in the receiver. Although the evolved codes do not have orthogonal generator matrices and the decoding is simple, excellent performance at practical signal-to-noise ratios is obtained in a fading environment. When the spectral efficiency is specified and an efficient outer code is used, evolved codes provide improved performance
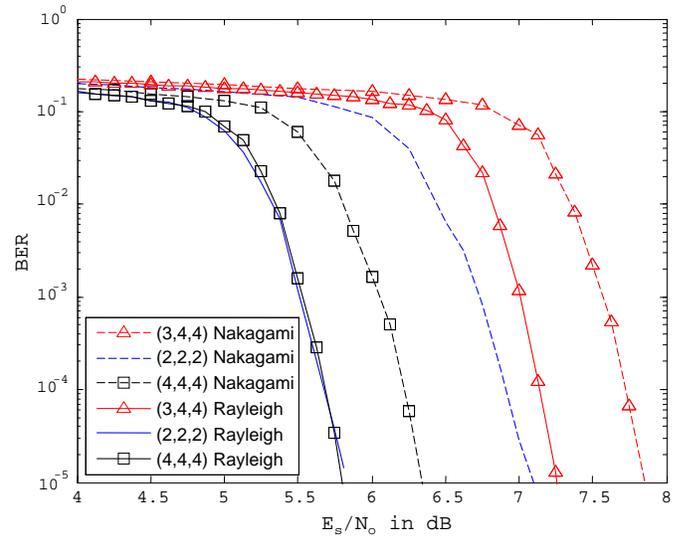
Fig. 6. BERs of turbo-coded 8-PSK-modulated $(4, 4, 4)$ and $(2, 2, 2)$ codes and 16-QAM-modulated $(3, 4, 4)$ codes. The fading is Rayleigh or Nakagami-1/2. The rate-1/2 turbo code has message length 4500.

relative to minimum-decoding-complexity quasi-orthogonal space-time codes at typical operating signal-to-noise ratios. When the fading is more severe than Rayleigh fading, evolved codes outperform orthogonal space-time block codes.

The genetic algorithm uses several novel methods that proved invaluable in finding effective space-time block codes. Both alpha-male parenting and parallel execution using cloning and immigration are likely to be useful methods in genetic algorithms for other applications.

APPENDIX A
ORTHOGONALITY CONDITION

Let $R(\mathbf{S})$ and $N(\mathbf{S})$ denote the range space and nullspace of matrix $\mathbf{S}$, respectively. Let $\mathbf{x}\perp\mathbf{y}$ denote the orthogonality of vectors $\mathbf{x}$ and $\mathbf{y}$, and $\mathcal{S}^\perp$ denote the orthogonal complement of subspace $\mathcal{S}$.

Proposition: Equations $(10) - (12)$ are necessary and sufficient for the satisfaction of the orthogonality condition defined by (9). *Proof:* Substitution of (8) into $\mathbf{H}^T\mathbf{H}$ indicates that (9) is satisfied if and only if $\mathbf{h}^T\mathcal{A}_n^T\mathcal{A}_n\mathbf{h} =||\mathbf{h}||^2$, $\mathbf{h}^T\mathcal{B}_n^T\mathcal{B}_n\mathbf{h} =||\mathbf{h}||^2$, $\mathbf{h}^T\mathcal{A}_n^T\mathcal{A}_\ell\mathbf{h} = \mathbf{0}$, $\mathbf{h}^T\mathcal{B}_n^T\mathcal{B}_\ell\mathbf{h} = \mathbf{0}$, $\mathbf{h}^T\mathcal{B}_n^T\mathcal{A}_\ell\mathbf{h} = \mathbf{0}$, $n, \ell = 1, 2, \ldots, N$, $n \neq \ell$, for any vector $2N \times 1$ $\mathbf{h}$ with real-valued components. Let $\mathbf{h} = [\mathbf{h}_1^T \ \mathbf{h}_2^T]^T$, where $\mathbf{h}_1$ and $\mathbf{h}_2$ are $N \times 1$ vectors. Substitution of (6) into these equations and the use of matrix algebra indicate that the orthogonality condition defined by (9) is satisfied if and only if

$$\mathbf{h}_1^T(\mathbf{A}_n^T\mathbf{A}_n - \mathbf{I})\mathbf{h}_1 = \mathbf{0}, \quad n = 1, 2, \ldots, N \quad (16)$$

$$\mathbf{h}_1^T(\mathbf{B}_n^T\mathbf{B}_n - \mathbf{I})\mathbf{h}_1 = \mathbf{0}, \quad n = 1, 2, \ldots, N \quad (17)$$

$$\mathbf{h}_1^T\mathbf{A}_n^T\mathbf{A}_\ell\mathbf{h}_1 = \mathbf{0}, \quad n, \ell = 1, 2, \ldots, N, \ n \neq \ell \quad (18)$$

$$\mathbf{h}_1^T\mathbf{B}_n^T\mathbf{B}_\ell\mathbf{h}_1 = \mathbf{0}, \quad n, \ell = 1, 2, \ldots, N, \ n \neq \ell \quad (19)$$

$$\mathbf{h}_1^T(\mathbf{A}_n^T\mathbf{B}_\ell - \mathbf{B}_\ell^T\mathbf{A}_n)\mathbf{h}_2 = \mathbf{0}, \quad n, \ell = 1, 2, \ldots, N, \ n \neq \ell \quad (20)$$

$$\mathbf{A}_1 = \frac{1}{2}\begin{bmatrix} -1 & -1 & 2 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix} \quad \mathbf{A}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \mathbf{A}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \quad \mathbf{A}_4 = \frac{1}{2}\begin{bmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix}$$

$$\mathbf{B}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \quad \mathbf{B}_2 = \frac{1}{2}\begin{bmatrix} 0 & -1 & 1 \\ -2 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \quad \mathbf{B}_3 = \frac{1}{2}\begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ -2 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \quad \mathbf{B}_4 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

Fig. 7. Dispersion matrices for the evolved $(4,3,4)$ code drawn from a discrete alphabet.

$$\mathbf{A}_1 = \begin{bmatrix} +0.2628 & -0.1397 & +0.1164 \\ +0.2237 & -0.0495 & +0.1744 \\ +0.6018 & -0.7709 & -0.1711 \\ +0.1061 & -0.5859 & -0.4770 \\ -0.6665 & -0.0380 & -0.7050 \\ +0.2514 & +0.1977 & +0.4497 \end{bmatrix} \quad \mathbf{A}_2 = \begin{bmatrix} -0.6671 & +0.5250 & -0.1421 \\ -0.0218 & -0.1765 & -0.1992 \\ -0.0524 & +0.3647 & +0.3075 \\ +0.4333 & -0.5554 & -0.1249 \\ -0.0211 & +0.4631 & +0.4445 \\ +0.6029 & +0.1933 & +0.7952 \end{bmatrix} \quad \mathbf{A}_3 = \begin{bmatrix} +0.4024 & +0.4910 & +0.8934 \\ -0.5319 & +0.7043 & +0.1722 \\ +0.3950 & -0.0106 & +0.3836 \\ +0.3414 & -0.1950 & +0.1432 \\ +0.4783 & -0.4387 & +0.0366 \\ +0.2317 & -0.1796 & +0.0568 \end{bmatrix}$$

$$\mathbf{A}_4 = \begin{bmatrix} +0.2400 & -0.3611 & -0.1223 \\ -0.1040 & +0.5662 & +0.4619 \\ -0.4319 & +0.1831 & -0.2463 \\ +0.7325 & +0.0151 & +0.7482 \\ -0.3318 & +0.2130 & -0.1156 \\ -0.3136 & +0.6855 & +0.3713 \end{bmatrix} \quad \mathbf{A}_5 = \begin{bmatrix} -0.0529 & +0.4248 & +0.3716 \\ -0.7724 & +0.1832 & -0.5864 \\ -0.1966 & -0.4856 & -0.6840 \\ -0.3587 & +0.4929 & +0.1345 \\ -0.4629 & +0.5540 & +0.0900 \\ +0.1378 & +0.0193 & +0.1547 \end{bmatrix} \quad \mathbf{A}_6 = \begin{bmatrix} -0.5133 & +0.3909 & -0.1229 \\ -0.2431 & -0.3410 & -0.5857 \\ +0.5039 & -0.0559 & +0.4479 \\ +0.1384 & +0.2594 & +0.3980 \\ -0.0460 & -0.4891 & -0.5316 \\ -0.6343 & +0.6491 & +0.0176 \end{bmatrix}$$

$$\mathbf{B}_1 = \begin{bmatrix} +0.2434 & -0.1222 & +0.1157 \\ -0.3793 & +0.2355 & -0.1451 \\ -0.0638 & +0.1675 & +0.1073 \\ -0.3258 & -0.5548 & -0.8808 \\ +0.1321 & +0.2799 & +0.4102 \\ -0.8180 & +0.7179 & -0.0994 \end{bmatrix} \quad \mathbf{B}_2 = \begin{bmatrix} +0.0397 & +0.2093 & +0.2505 \\ -0.0515 & -0.3432 & -0.3925 \\ -0.9318 & +0.2431 & -0.6877 \\ +0.2863 & -0.5082 & -0.2238 \\ +0.2113 & -0.7215 & -0.5101 \\ +0.0285 & -0.0198 & +0.0039 \end{bmatrix} \quad \mathbf{B}_3 = \begin{bmatrix} +0.3109 & +0.4283 & +0.7393 \\ +0.6497 & -0.2846 & +0.3615 \\ -0.2556 & +0.4719 & +0.2192 \\ -0.6129 & +0.5428 & -0.0724 \\ -0.1991 & +0.0232 & -0.1748 \\ +0.0230 & +0.4666 & +0.4889 \end{bmatrix}$$

$$\mathbf{B}_4 = \begin{bmatrix} -0.3271 & +0.7113 & +0.3825 \\ +0.3398 & -0.0157 & +0.3284 \\ -0.0627 & +0.0700 & +0.0095 \\ +0.3964 & -0.3490 & +0.0455 \\ -0.6012 & +0.4930 & -0.1048 \\ -0.5049 & -0.3523 & -0.8560 \end{bmatrix} \quad \mathbf{B}_5 = \begin{bmatrix} -0.6338 & +0.1644 & -0.4700 \\ -0.3817 & +0.8396 & +0.4592 \\ -0.2334 & +0.3890 & +0.1551 \\ -0.5177 & +0.1069 & -0.4074 \\ -0.3230 & -0.2919 & -0.6146 \\ +0.1607 & -0.1417 & +0.0213 \end{bmatrix} \quad \mathbf{B}_6 = \begin{bmatrix} +0.5778 & -0.4744 & +0.1010 \\ -0.4125 & -0.2013 & -0.6126 \\ -0.0630 & +0.7305 & +0.6660 \\ +0.1039 & -0.0776 & +0.0286 \\ -0.6578 & +0.2688 & -0.3903 \\ +0.2204 & -0.3502 & -0.1332 \end{bmatrix}$$

Fig. 8. Dispersion matrices for the evolved $(6,3,6)$ code drawn from a continuous alphabet.

for any $N \times 1$ vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ with real-valued components. By direct substitution, it is verified that (10) and (12) imply (16), (17), and (20). Substitution of (11) indicates that $0 = \mathbf{h}_1^T \mathbf{A}_n^T \mathbf{A}_\ell \mathbf{h}_1 - \left(\mathbf{h}_1^T \mathbf{A}_n^T \mathbf{A}_\ell \mathbf{h}_1\right)^T = 2\mathbf{h}_1^T \mathbf{A}_n^T \mathbf{A}_\ell \mathbf{h}_1$, which implies (18). Similarly, (11) implies (19). Thus, (10) − (12) are sufficient for the satisfaction of the orthogonality condition. For the necessity, observe that (18) and (19) require that $\mathbf{h}_1^T (\mathbf{A}_n^T \mathbf{A}_\ell + \mathbf{A}_\ell^T \mathbf{A}_n)\mathbf{h}_1 = \mathbf{0}$ and $\mathbf{h}_1^T (\mathbf{B}_n^T \mathbf{B}_\ell + \mathbf{B}_\ell^T \mathbf{B}_n)\mathbf{h}_1 = \mathbf{0}$, $n, \ell = 1, 2, \ldots, N$, $n \neq \ell$. Thus, (16) − (19) each requires that $\mathbf{h}_1^T \mathbf{S}\mathbf{h}_1 = \mathbf{0}$ for a symmetric matrix $\mathbf{S}$ and every $N \times 1$ vectors $\mathbf{h}_1$. If $\mathbf{S} \neq \mathbf{0}$, then there exists some $\mathbf{h}_1 \notin N(\mathbf{S})$, and $\mathbf{S}\mathbf{h}_1 \in R(\mathbf{S})$. Since $\mathbf{h}_1 \perp \mathbf{S}\mathbf{h}_1$ and $R(\mathbf{S})^\perp = N(\mathbf{S}^T) = N(\mathbf{S})$, $\mathbf{h}_1 \in N(\mathbf{S})$. This contradiction proves that $\mathbf{S} = \mathbf{0}$. Thus, the orthogonality condition implies the necessity of (10) and (11). Equation (20) has the form $\mathbf{h}_1^T \mathbf{S}\mathbf{h}_2 = \mathbf{0}$ for a matrix $\mathbf{S}$ and every $\mathbf{h}_2$, which implies that every $\mathbf{h}_1 \perp R(\mathbf{S})$. Thus, every $\mathbf{h}_1 \in R(\mathbf{S})^\perp = N(\mathbf{S}^T)$, which implies that $\mathbf{S} = \mathbf{0}$, and hence (12) is necessary. ∎

APPENDIX B
DISPERSION MATRICES

The dispersion matrices for the $(4,3,4)$ code that evolved after 1 million generation are given in Fig. 7, and the dispersion matrices for the $(6,3,6)$ code that evolved after 1 million generation are given in Fig. 8.

REFERENCES

[1] J. G. Proakis and M. Salehi, *Digital Communications, fifth ed.* New York: McGraw-Hill, 2008.

[2] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications.* New York: Cambridge Univ. Press, 2003.

[3] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Info. Theory*, vol. 45, pp. 1456-1467, July 1999.

[4] D. N. Dao and C. Tellambura, "Quasi-Orthogonal STBC with Minimum Decoding Complexity: Performance Analysis, Optimal Signal Transformations, and Antenna Selection Diversity," *IEEE Trans. Commun.*, vol. 56, pp. 849-853, June 2008.

[5] C. Yuen, Y. L. Guan, and T. T. Tjhung, "Quasi-orthogonal STBC with minimum decoding complexity," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2089-2094, Sept. 2005.

[6] H. Jafarkhani, "A quasi-orthogonal space–time block code," *IEEE Trans. Commun.*, vol. 49, pp. 1-4, Jan. 2001.

[7] D. N. Dao and C. Tellambura, "Decoding, Performance Analysis, and Optimal Signal Designs for Coordinate Interleaved Orthogonal Designs," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 48-53, Jan. 2008.

[8] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Info. Theory*, vol. 48, pp. 1804-1824, July 2002.

[9] K. A. DeJong, *Evolutionary Computation: A Unified Approach* Cambridge, MA: MIT Press, 2006.

[10] I. Sen, D. W. Matolak, W. Xiong, "Wireless channels that exhibit 'worse than Rayleigh' fading: Analytical and measurement results," *IEEE Military Commun. Conf.*, 2006.

[11] European Telecommunications Standards Institute, "Universal mobile telecommunications system (UMTS): Multiplexing and channel coding (FDD)," *ETSI TS 25.212 version 6.6.0*, Sept. 2005.