EURASIP Journal on Applied Signal Processing 2002:8, 854–864

Spatial Block Codes Based on Unitary Transformations Derived from Orthonormal Polynomial Sets

Giridhar D. Mandyam
Nokia Research Center, 6000 Connection Drive, Irving, TX 75039, USA
Email: giridhar.mandyam@nokia.com

Received 1 September 2001 and in revised form 15 March 2002

Recent work in the development of diversity transformations for wireless systems has produced a theoretical framework for space-time block codes. Such codes are beneficial in that they may be easily concatenated with interleaved d trellis codes and yet still may be decoded separately. In this paper, a theoretical framework is provided for the generation of spatial block codes of arbitrary dimensionality through the use of orthonormal polynomial sets. While these codes cannot maximize theoretical diversity performance for given dimensionality, they still provide performance improvements over the single-antenna case. In particular, their application to closed-loop transmit diversity systems is proposed, as the bandwidth necessary for feedback using these types of codes is fixed regardless of the number of antennas used. Simulation data is provided demonstrating these types of codes' performance under this implementation as compared not only to the single-antenna case but also to the two-antenna code derived from the Radon-Hurwitz construction.

Keywords and phrases: spatial block codes, closed-loop transmit diversity, space-time codes.

1. INTRODUCTION

In wireless communications systems, *fading* transmission channels are problematic due to the fact that fading channels are nonstationary, and therefore the design of effective channel codes based on assumed channel statistics becomes difficult. As a result, *diversity* is essential for addressing the problem of fading in wireless channels. Diversity essentially entails receiving several replicas of the same signal over independently fading channels [1]. Diversity may take many approaches. For instance, frequency diversity methods employ transmission of multiple symbol replicas over multiple carriers, each of the carriers separated in frequency by a sufficiently large amount to ensure independent fading. This approach is accompanied with the additional cost of increased complexity at the transmitter and receiver, along with the fact that it may be difficult to implement in bandwidth-limited systems (such as common public wireless systems that must conform to electromagnetic compatibility requirements). Temporal diversity entails transmission of signal replicas in different time slots, each slot sufficiently spaced in time to ensure independent fading. This approach suffers from reduced throughput due to multiple transmissions of the same symbol over time. Another instance of temporal diversity may be achieved in multipath channels where the signal bandwidth is larger than the coherence time of the channel; in this case the multipaths are resolvable and may be recovered by a rake receiver.

However, flat fading channels are troublesome for bandwidth-limited systems where neither frequency nor temporal diversity is possible. In such conditions, *antenna diversity* is a concept that has gained much interest. Transmission of signal replicas over multiple antennas using separable waveforms essentially results in a received signal that may be demodulated with a rake receiver. Usually, to achieve such diversity, a spatial separation of at least ten wavelengths between antennas is required to ensure independent-fading conditions for signals associated with each antenna.

While antenna diversity is a desirable alternative for public wireless systems, the actual requirements for achieving optimal diversity over such systems have recently been the subject of several studies. The two questions at hand are

(1) can coding over multiple antennas have benefits over *simple diversity* schemes, where multiple copies of the same signal are transmitted over multiple antennas at discrete instances in time? In other words, can a *space-time code* be designed?
(2) If so, how do we optimally code to achieve the full benefits of diversity in these systems?

There have been several approaches to answer both of these questions. For instance, in [2] (which is an extension of the authors' earlier work in [3]), the authors provide a criterion for block code design for transmitter diversity systems and demonstrate their benefits when proper channel estimation is possible through the use of pilot-symbol assisted modulation. This work has been addressed from a slightly different viewpoint in [4], wherein the authors construct generalized orthogonal space-time block codes based on the Radon-Hurwitz construction for unitary matrices of indeterminates. The performance data for these codes are given in [5].

Other approaches have been taken with respect to transmitter code design where trellis coding is incorporated. For instance, in [6] the authors derive several space-time trellis codes, which were found with respect to the *product* criterion, wherein the minimum of the product of distances between all distinct code word pairs is maximized assuming that the rank of the code word difference matrices are maximized. In [7], the authors provide a criterion for the design of space-time trellis codes by forming a search criterion different from [6] based on the assumption that optimal codes will satisfy the same criterion for their distance spectra as traditional trellis codes used in the single antenna case.

In [8], the author derives a criterion for space-time code design based on the Euclidean distance between all possible code word pairs. This criterion is different from the product distance used in [6, 7], but is shown to be a true metric.

In this paper, a general design methodology is presented for spatial block codes based on orthogonal designs. The reason why these codes are referred to as *spatial* block codes rather than *space-time* block codes is that, as will be shown, these codes primarily involve spatial processing but not temporal processing. Although the new design methodology does not satisfy design criteria for diversity maximization, they can be shown to be useful in closed-loop transmit diversity application. Simulation results are provided to verify the benefits of these codes in closed-loop scenarios.

This paper is organized as follows. Section 2 provides an overview of the design criteria for space-time block codes. Section 3 presents a general framework for construction of spatial block codes from unitary transform matrices and introduces an application of these codes to closed-loop transmit diversity systems. Section 4 provides simulation results using the proposed codes. Section 5 includes a discussion on the significance of the results and directions for future work.

2. SPACE-TIME BLOCK CODES: DESIGN CRITERIA

In this section, the criterion for optimal space-time block codes are derived and presented. This criterion has been derived and presented in previous work (e.g., [2, 6]). Given a space-time block code designed for L antennas for duration of K epochs, the transmitted code words may be defined by a $K \times L$ matrix $\mathbf{D}$

$$\mathbf{D}(t) = \begin{bmatrix} d_t^1 & d_t^2 & \cdots & d_t^L \\ d_{t+1}^1 & d_{t+1}^2 & \cdots & d_{t+1}^L \\ \vdots & \vdots & \ddots & \vdots \\ d_{t+K-1}^1 & d_{t+K-1}^2 & \cdots & d_{t+K-1}^L \end{bmatrix}, \tag{1}$$

where the matrix entries d_t^i represent the modulation symbol transmitted over the ith antenna at time t (t being in multiples of the symbol duration). Given a single-antenna receiver, the received signal may be represented as

$$x(t) = \sum_{i=1}^{L} d_t^i c^i(t) + n(t), \tag{2}$$

where $c^i(t)$ is the complex channel gain at time t of the signal transmitted from the ith antenna and $n_i(t)$ is the associated Gaussian noise. If it is assumed that the channel estimate, corresponding to the channel as seen from each antenna, is separable at the receiver (by means of orthogonal waveform coding, for instance) and that the complex channel gain and noise for each antenna remain constant over K epochs, then the signal corresponding to the entire code matrix received over the K-epoch duration of the space-time code may be represented as

$$\mathbf{x}(t) = \mathbf{D}(t)\mathbf{c}(t) + \mathbf{n}(t), \tag{3}$$

where the vector $\mathbf{x}(t)$ is a $K \times 1$ observation vector, $\mathbf{D}(t)$ is the $K \times L$ code word matrix, and $\mathbf{c}(t)$ is the $L \times 1$ channel gain vector defined as

$$\mathbf{c}(t) = [c_1(t), c_2(t), \ldots, c_L(t)], \tag{4}$$

and $\mathbf{n}(t)$ is the $K \times 1$ noise vector defined as

$$\mathbf{n}(t) = [n(t), n(t+1), \ldots, n(t+K-1)]. \tag{5}$$

Given the received signal vector $\mathbf{x}(t)$ and assuming perfect channel estimation at the receiver, the maximum a posteriori detector is given as

$$\hat{\mathbf{D}}(t) = \max_{\mathbf{D}_\gamma \in \mathbf{S}} p(\mathbf{D}_\gamma \mid \mathbf{c}(t), \mathbf{x}(t)), \tag{6}$$

where $\mathbf{S}$ is the set of all possible codematrices, $\hat{\mathbf{D}}(t)$ is the detected codematrix, and $p(\text{arg})$ is the probability density function of arg. If it is assumed that each Gaussian noise sample is independent, zero-mean with variance σ^2, then the pdf used in (6) may be found as

$$\begin{aligned} p(\mathbf{D}_\gamma \mid \mathbf{c}(t), \mathbf{x}(t)) \\ = (2\pi\sigma^2)^{-K/2} e^{-(1/2\sigma^2)(\mathbf{x}(t)-\mathbf{D}_\gamma\mathbf{c}(t))^H(\mathbf{x}(t)-\mathbf{D}_\gamma\mathbf{c}(t))}. \end{aligned} \tag{7}$$

Using (7), the probability of decoding error may be found as

$$\begin{aligned} P_{\text{error}} &= P(\hat{\mathbf{D}}(t) = \mathbf{D}_\alpha \mid \mathbf{D}(t) = \mathbf{D}_\beta) \\ &= P\{p(\mathbf{D}_\alpha \mid \mathbf{c}(t), \mathbf{x}(t)) > p(\mathbf{D}_\beta \mid \mathbf{c}(t), \mathbf{x}(t))\} \\ &= P\{\ln(p(\mathbf{D}_\alpha \mid \mathbf{c}(t), \mathbf{x}(t))) < \ln(p(\mathbf{D}_\beta \mid \mathbf{c}(t), \mathbf{x}(t)))\}. \end{aligned} \tag{8}$$

The relationship of (8) may be simplified to

$$P_{\text{error}} = P\Big\{\big(\mathbf{x}(t) - \mathbf{D}_\alpha \mathbf{c}(t)\big)^H \big(\mathbf{x}(t) - \mathbf{D}_\alpha \mathbf{c}(t)\big) < \big(\mathbf{x}(t) - \mathbf{D}_\beta \mathbf{c}(t)\big)^H \big(\mathbf{x}(t) - \mathbf{D}_\beta \mathbf{c}(t)\big)\Big\}. \tag{9}$$

Noting that in the original expression in (7), it was assumed that if $\mathbf{D}(t) = \mathbf{D}_\beta$, then $\mathbf{x}(t) = \mathbf{D}_\beta \mathbf{c}(t) + \mathbf{n}(t)$. Therefore, the error probability may be simplified to

$$P_{\text{error}} = P\Big\{2\,\text{Re}\left[\mathbf{n}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t)\right] > \mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t)\Big\}. \tag{10}$$

Observing (10), it is clear that the probability of error decreases as the term on the right-hand side of the inequality increases. Noting that $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ is a $K \times L$ matrix, it may be decomposed using singular value decomposition. As a result, $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ is equivalent to $\mathbf{V}^H \Sigma \mathbf{W}$, where $\mathbf{V}$ is a $K \times K$ unitary matrix, W is an $L \times L$ unitary matrix, and Σ is a $K \times L$ matrix whose diagonal entries are the singular values in order of value of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ (i.e., the eigenvalues of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$). Therefore, the following equation may be derived:

$$\begin{aligned}\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t) &= \mathbf{c}^H(t)\mathbf{V}^H\Sigma\mathbf{W}\mathbf{W}^H\Sigma^H\mathbf{V}\mathbf{c}(t) \\ &= \mathbf{c}^H(t)\mathbf{V}^H\Sigma\Sigma^H\mathbf{V}\mathbf{c}(t).\end{aligned} \tag{11}$$

If we assume that Σ has the structure $\text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots]$, where λ_k denotes the kth nonzero eigenvalue of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$, then taking into account the unitarity of $\mathbf{V}$, then (11) may be further simplified as

$$\begin{aligned}\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t) &= \mathbf{c}^H(t)\mathbf{V}^H\Sigma\Sigma^H\mathbf{V}\mathbf{c}(t) \\ &= \mathbf{c}^H(t)\Sigma\Sigma^H\mathbf{c}(t) \\ &= \sum_{i=1}^{r}\lambda_i^2 c_i^2(t).\end{aligned} \tag{12}$$

Quite clearly, the larger the rank of the $L \times L$ matrix $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$, the lower the decision error probability. If this matrix is full rank, then the maximum gains from diversity are achieved. However, this criterion is general, and it would be of interest for code design to find a narrower criterion. This may be accomplished by examining of $\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t)$ in the mean sense. Firstly, it is assumed that the transmitted symbol energy from each antenna is E_s. Since the channel itself does not create or destroy energy, the mean energy from the complex channel gain coefficients as seen at the receiver should be $E\{c_i^2(t)\} = E_s$. Therefore, the following equations may be derived:

$$\begin{aligned}E\Big\{\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t) \mid \mathbf{D}_\beta, \mathbf{D}_\alpha\Big\} &= E\left\{\sum_{i=1}^{r}\lambda_i^2 c_i^2(t)\right\} \\ &= \sum_{i=1}^{r}E\{\lambda_i^2 c_i^2(t)\} \\ &= \sum_{i=1}^{r}E\{\lambda_i^2\}E\{c_i^2(t)\} \\ &= \sum_{i=1}^{r}E\{\lambda_i^2\}E_s.\end{aligned} \tag{13}$$

In [8], the author proposes that $E\{\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t) \mid \mathbf{D}_\beta, \mathbf{D}_\alpha\}$ may be bounded using the Cauchy-Schwartz inequality *assuming that the singular values of* $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ *are deterministic*. As a result, the relationship in (13) may be bounded as

$$\begin{aligned}E\Big\{\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t) \mid \mathbf{D}_\beta, \mathbf{D}_\alpha\Big\} &= \sum_{i=1}^{r}E\{\lambda_i^2\}E_s \\ &= \sum_{i=1}^{r}\lambda_i^2 E_s \\ &\le \sqrt{\left(\sum_{i=1}^{r}\lambda_i^4\right)}\sqrt{rE_s^2}.\end{aligned} \tag{14}$$

Therefore, if $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ is a diagonal matrix with all entries of the diagonal being equal, the bound of (14) becomes tight. However, even the singular values of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ are in fact not deterministic in the mean-sense, due to the fact that for all given code words these values are functions of the mean code word differences. Therefore, given a set of code word symbols which may be transmitted, one may use the distribution of all possible code word symbol differences to form an expression for $E\{\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t)\}$. It is clear that if $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ is diagonal and all entries along the diagonal are nonzero, then the maximum gain from diversity is achieved. In this case, the singular values of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ are functions of the symbol differences between the code words.

3. SPATIAL BLOCK CODE DESIGN FROM UNITARY TRANSFORM MATRICES: A GENERAL DESIGN FRAMEWORK

As established in Section 2, the desired criterion for space-time block code design is to find codes whose difference matrices satisfy the condition that there exists the maximum number of singular values associated with these matrices. One such construct, as discussed in [4], is the Radon-Hurwitz unitary matrix construction. A set of k unitary matrices $\{B_I\}$ of size $L \times L$ is part of the Radon-Hurwitz family if the following three rules hold:

$$B_i^T B_i = I,$$
$$B_i^T = -B_i, \quad 1 \le i \le k, \tag{15}$$
$$B_i B_j = -B_j B_i, \quad 1 \le i, j \le k.$$

In [9], the author presented a code that corresponded to a special case of the Radon-Hurwitz family

$$\mathbf{D}(t) = \begin{bmatrix} s_1 & s_2 \\ -(s_2)^* & (s_1)^* \end{bmatrix}, \tag{16}$$

where $\{s_i\}$ are the set of symbols to be transmitted over the K time epochs of the code (in this case, $K = 2$). The code in (16) is an example of a rate 1 code, where the number of symbols transmitted is equal to the number of time epochs required for the code. Such codes are desirable for bandwidth-limited systems. In [4], the authors show that for the Radon-Hurwitz family of code constructs, rate 1 designs exist for real constellations only for $L = K = 2$, 4, and 8. The authors conclude that real orthogonal designs therefore exist only for these dimensions. The singular values for the code word difference matrix $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ can be shown to be (see [4])

$$\lambda_i^2 = \sum_{l=1}^{L} |s_{\alpha l} - s_{\beta l}|^2 \quad \forall i. \tag{17}$$

In (17), $[s_{\alpha 1} \;\; s_{\alpha 2} \;\; \cdots \;\; s_{\alpha L}]$ is the first row of $\mathbf{D}_\alpha$ and $[s_{\beta 1} \;\; s_{\beta 2} \;\; \cdots \;\; s_{\beta L}]$ is the first row of $\mathbf{D}_\beta$. Since it is assumed that $\mathbf{D}_\alpha$ differs from $\mathbf{D}_\beta$ in at least one position, all the singular values of the code word difference matrix are nonzero, that is, $r = L$ in (13). As a result, the relationship of (13) may be found as

$$\begin{aligned} &E\Big\{\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t) \mid \mathbf{D}_\beta, \mathbf{D}_\alpha\Big\} \\ &\quad = \sum_{i=1}^{r} E\{\lambda_i^2\}E_s \\ &\quad = E_s L \sum_{l=1}^{L} |s_{\alpha l} - s_{\beta l}|^2. \end{aligned} \tag{18}$$

However, given an $L \times L$ unitary matrix $\mathbf{U}$ whose elements are denoted by $\mathbf{U}_{ij}$, then by forming a diagonal matrix $\mathbf{G} = \text{diag}[s_1, s_2, \ldots, s_L]$, a code word matrix may be formed as $\mathbf{D}(t) = \mathbf{GU}$. This matrix will be unitary *assuming that the symbol constellation points have equal magnitude*, and under this assumption the singular values of the code word difference matrix $(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ are simply $[|s_{\alpha 1} - s_{\beta 1}|^2, |s_{\alpha 2} - s_{\beta 2}|^2, \ldots, |s_{\alpha K} - s_{\beta K}|^2]$. This type of design was demonstrated in [2] for a specific code, but taking into account the fact that unitary matrices may be constructed from sets of orthonormal polynomials, a general method for designing block codes based on unitary matrices may be specified. As a result, the relationship in (13) for this type of code, herein denoted as a *simple orthogonal code*, becomes

$$\begin{aligned} &E\Big\{\mathbf{c}^H(t)(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)\mathbf{c}(t) \mid \mathbf{D}_\beta, \mathbf{D}_\alpha\Big\} \\ &\quad = \sum_{i=1}^{L} E\{\lambda_i^2\}E_s \\ &\quad = E_s \sum_{i=1}^{r} |s_{\alpha i} - s_{\beta i}|^2. \end{aligned} \tag{19}$$

Comparing (18) to (19), it can be seen that at best, the performance of the simple orthogonal code can match that of the Radon-Hurwitz code for any given code word pair. This is due to the fact that although the first column of $\mathbf{D}_\alpha$ and that of $\mathbf{D}_\beta$ are distinct, they may differ in at least one position. Therefore, to analyze the diversity gain of a simple orthogonal code, we must analyze it in the mean-sense. This would mean that we should look at the average rank of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ rather than the rank of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ for a particular code word pair. This would be determined by the average number of positions in which the first column of $\mathbf{D}_\alpha$ and that of $\mathbf{D}_\beta$ differ for all possible distinct code word pairs $(\mathbf{D}_\alpha, \mathbf{D}_\beta)$.

Given a symbol alphabet of dimensionality M, the set of all possible code words that make up the first column of the code word matrix derived from a simple orthogonal code of dimension $L \times L$ is M^L. Given that there are $\binom{M^L}{2}$ distinct code word pairs, the average rank (i.e., $E\{r\}$, where r is the number of singular values of $(\mathbf{D}_\beta - \mathbf{D}_\alpha)^H(\mathbf{D}_\beta - \mathbf{D}_\alpha)$ for any code word pair $(\mathbf{D}_\alpha, \mathbf{D}_\beta)$) is

$$E\{r\} = \sum_{i=1}^{L} i p_i = \frac{\sum_{i=1}^{L} i A_i}{\binom{M^L}{2}}, \tag{20}$$

where A_i is the number of code word pairs differing in i positions and p_i is the probability that any two code words differ in i positions,

$$p_i = \frac{A_i}{\binom{M^L}{2}}. \tag{21}$$

Finding the general form for A_i is cumbersome; however, we may derive an upper bound for $E\{r\}$ based on the value of A_L. This value may be shown to be

$$A_L = \sum_{i=0}^{M-2} M^{L-1}\big[(M-1)^L - i(M-1)^{L-1}\big]. \tag{22}$$

Given this value, we can find the value of p_L. In addition, since $p_i \ge 0$,

$$\sum_{i=1}^{L} p_i = 1 \Longrightarrow p_L + \sum_{i=1}^{L-1} p_i = 1 \Longrightarrow p_i \le (1 - p_L) \quad \forall 1 \le i < L. \tag{23}$$

Therefore, the maximum value of $E\{r\}$ based on the value of A_L may be derived as

$$E\{r\} = \sum_{i=1}^{L} i p_i = L p_L + \sum_{i=1}^{L-1} i p_i \le L p_L + \sum_{i=1}^{L-1} i(1-p_L) \le L p_L + (L-1)(1-p_L). \quad (24)$$

Substituting in for p_L, (24) may be expressed as

$$E\{r\} \le \frac{L A_L}{\binom{M^L}{2}} + (L-1)\left(1 - \frac{A_L}{\binom{M^L}{2}}\right), \quad (25)$$

where the first term on the right-hand side of the inequality represents the likelihood that a code word pair differs in all L positions, while the second term represents the likelihood that a code word pair differs in at most $(L-1)$ positions.

Similarly, a lower bound may be derived for $E\{r\}$. We may first derive the number of code word pairs that differ in only one position

$$A_1 = L M^{L-1} \sum_{i=0}^{M-2} (M-1) - i. \quad (26)$$

Therefore, the lower bound is

$$E\{r\} \ge p_1 + 2(1-p_1) \implies E\{r\} \ge \frac{A_1}{\binom{M^L}{2}} + 2\left(1 - \frac{A_1}{\binom{M^L}{2}}\right). \quad (27)$$

In (27), the first term on the right-hand side of the inequality represents the probability of a code word pair differing in only one position, and the second term represented the probability of a code word pair differing in at least two positions. The derivations of A_1 and A_L are explained further in the appendix.

As an example, consider a 3-antenna code using QPSK symbols. In this case, $M = 4$ and $L = 3$. Using (25) and (27), the following bounds are derived:

$$1.86 \le E\{r\} \le 2.43. \quad (28)$$

To see how meaningful these bounds are, a simulation was run under one-path Rayleigh fading conditions at a velocity of 1 km/h using a 19.2 kbit/s transmission rate (2 bits/QPSK, symbol) and a carrier frequency of 1960 MHz. The demodulated bit error rate (BER) as a function of QPSK symbol, SNR was compared between three-transmission methods: no diversity, use of the 2×2 Radon-Hurwitz code as in (16), and a 3-antenna orthogonal code based on the discrete Fourier transform (DFT) matrix (see (45)). The results are depicted in Figure 1. The simple orthogonal code provided nearly the same performance as the 2-antenna Radon-Hurwitz code, which has diversity order 2. The bounds given in (28) predict a mean diversity order near 2 as well.

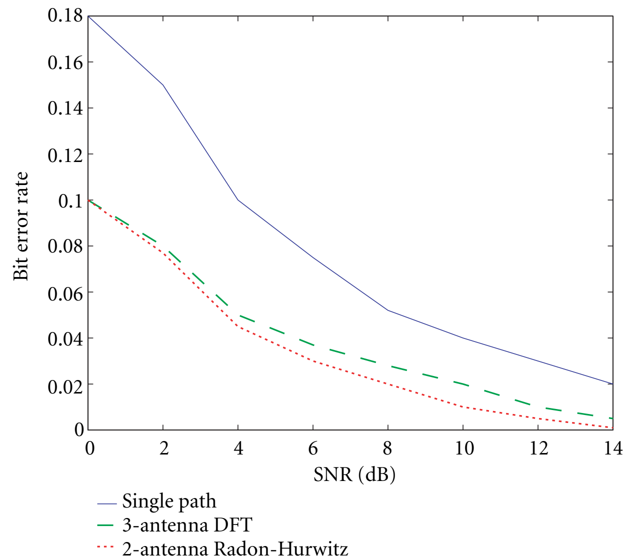

Figure 1: Performance of simple orthogonal code.

Now that a general description of simple orthogonal spatial block codes has been presented, a general method for deriving these codes may be formulated starting with a general method for deriving **U**, the unitary transformation matrix.

3.1. The Gauss-Jacobi procedure for unitary transform matrix derivation

A set of real polynomials $\{P_k(x)\}$, each of degree k, is said to be orthonormal with respect to the weighting function $\{p(x)\}$ over the support space Ω, if

$$\int_{\Omega} P_i(x) P_j(x) p(x) dx = \delta_{ij}, \quad (29)$$

where δ is the Dirac delta operator. In order to ensure orthogonality, the polynomials $P_k(x)$ of degree greater than zero must satisfy [10]

$$\int_{\Omega} P_k(x) x^m p(x) dx = 0, \quad 0 \le m < k. \quad (30)$$

The Lagrange interpolating polynomial for a set of n discrete sample points $f(x_k)$ of a function $f(x)$ is defined as

$$F_n(x) = \sum_{k=1}^{n} \frac{\omega(x)}{(x-x_k)\omega'(x_k)} f(x_k), \quad (31)$$

where $\omega(x) = (x-x_1)(x-x_2)\cdots(x-x_n)$ and $\omega'(x_k)$ is the polynomial given by $(x_k-x_1)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)$. If $f(x)$ is a polynomial of degree less than n, then $F_n(x) = f(x)$. Otherwise, we can form the interpolatory polynomial $F_n(x)$ and represent $f(x)$ as

$$f(x) = F_n(x) + r(x), \quad (32)$$

where $r(x)$ is a remainder polynomial. Integrating $f(x)$ over Ω with respect to $p(x)$, and assuming that the remainder polynomial $r(x)$ is negligible, we obtain

$$\int_{\Omega} f(x)p(x)dx \approx \sum_{k=1}^{n} A_k f(x_k), \tag{33}$$

where

$$A_k = \int_{\Omega} \frac{\omega(x)}{(x - x_k)\omega'(x_k)} p(x)dx. \tag{34}$$

The right-hand side of (33) is commonly referred to as a *quadrature formula*, and leads to two known theorems (the proofs may be found in [11]). The first theorem is that the quadrature formula in (33) is interpolatory if and only if it is exact for all possible polynomials $\{f(x)\}$ of degree less than or equal to $n - 1$. The second theorem is that the quadrature formula in (33) is exact for all polynomials of degree less than or equal to $2n - 1$, if and only if (i) the quadrature formula in (33) is interpolatory and (ii) for all polynomials $Q(x)$ of degree less than n,

$$\int_{\Omega} \omega(x)Q(x)p(x)dx = 0. \tag{35}$$

Assuming a set of orthonormal polynomials $P_k(x)$ over support space Ω, a discrete unitary transform matrix can now be constructed. Assume that the polynomials $P_k(x)$ are arranged in order of increasing degree, that is, $\deg(P_1(x)) \leq \deg(P_2(x)) \leq \cdots \leq \deg(P_k(x)) \leq \cdots$. If an $N \times N$ unitary matrix is desired, it can be generated by first taking the discretization points x_k as the roots of $P_{N+1}(x)$. If we form $\omega(x)$ from these points, we know that $P_{N+1}(x)$ is directly proportional to $\omega(x)$ and therefore, any polynomials orthogonal to $P_{N+1}(x)$ will also be orthogonal to $\omega(x)$. We also know that if we define $f(x)$ in (33) as the product of $P_{N+1}(x)$ and the arbitrary polynomial $d(x)$ of degree less than N, then $f(x)$ is a polynomial of degree less than or equal to $(2N - 1)$ and

$$\int_{\Omega} P_{N+1}(x)d(x)p(x)dx = 0 = \sum_{k=1}^{N} A_k P_{N+1}(x_k)d(x_k), \tag{36}$$

since $P_{N+1}(x_k) = 0$ for all k and since we can always find A_k such that (33) is exact. Therefore, all polynomials of degree less than n are orthogonal to $\omega(x)$ and thus by the second theorem, previously mentioned, (33) is exact for all polynomials of degree less than $2N - 1$. Thus, by the orthonormality condition of the polynomials $P_k(x)$, we conclude that

$$\int_{\Omega} P_i(x)P_j(x)p(x)dx = \delta_{ij} = \sum_{k=1}^{N} A_k P_k(x_k)P_j(x_k) \tag{37}$$

for i, j less than $(N + 1)$. Therefore, if the (i, k) entry of an $N \times N$ matrix is formed by the value $P_i(x_k)$, we can find A_k such that this matrix is unitary, that is, all the row vectors are mutually orthonormal.

To that end, we first consider the *Christoffel-Darboux* identity [11], which is defined as follows: given a set of orthonormal polynomials $\{P_n(x)\}$, each of order n, with the nth-order term $\{P_n(x)\}$ in each being of the form $a_n x^n$, it can be shown that

$$(x - t)\sum_{s=0}^{n} P_s(x)P_s(t) = -\frac{a_n}{a_{n+1}}\left[P_{n+1}(x)P_n(t) - P_n(x)P_{n+1}(t)\right]. \tag{38}$$

In (38), if we set t to be the roots of $P_n(x)$, that is, x_k, then it can be shown that

$$\sum_{s=0}^{n-1} P_s(x)P_s(x_k) = -\frac{a_n}{a_{n+1}} \frac{P_n(x)P_{n+1}(x_k)}{x - x_k}. \tag{39}$$

Multiplying both sides by $p(x)$ and integrating over Ω, we get

$$1 = -\frac{a_n}{a_{n+1}} P_{n+1}(x_k) \int_{\Omega} p(x)\frac{P_n(x)}{x - x_k}dx. \tag{40}$$

The result in (40) follows from the fact that the quantity $P_s(x_k)\int_{\Omega} P_s(x)p(x)dx$ equals 0 for $s > 0$, as a result of the orthogonality condition in (30), and equals 1 for $s = 0$, as a result of the orthonormality of $P_0(x)$. We note that the integral on the right-hand side of (40) is similar to the definition of A_k in (34), from which it follows that

$$A_k = -\frac{a_n}{a_{n+1}} \frac{1}{P_n'(x_k)P_{n+1}(x_k)}. \tag{41}$$

If we refer to the desired $N \times N$ unitary matrix as $\mathbf{U}$, we can define the elements of $\mathbf{U}$ as

$$\mathbf{U}_{ij} = \sqrt{A_j}P_i(x_j). \tag{42}$$

It should be noted that the theory presented here is not directly applicable to complex orthonormal polynomials. This will be discussed in more detail in Section 3.2.

3.2. Sample orthogonal designs

Returning to the code construct $\mathbf{D}(t) = \mathbf{GU}$, where $\mathbf{G} = \mathrm{diag}[s_1, s_2, \ldots, s_L]$ and $\mathbf{U}$ is a unitary matrix, several codes may be derived, which satisfy the rank criterion for the code word difference matrix. We primarily concentrate on the 3×3 case, as this is the lowest order where Radon-Hurwitz codes do not exist. For instance, the discrete Fourier transform matrix of dimension $L \times L$ is derived from the rule

$$\mathbf{F}_{lm} = \frac{e^{-j2\pi(l-1)(m-1)/L}}{\sqrt{L}}, \tag{43}$$

where l is the row index ranging from 0 to $(L - 1)$ and m is the column index also ranging from 0 to $(L - 1)$. Thus, returning to the terminology presented in Section 3.1, the set of orthonormal polynomials for the DFT matrix are simply described by $\{P_n(x = e^{-j2\pi i/L})\} = \{x^n\} = \{e^{-j2\pi in/L}\}$, where n is the order of the polynomial, and the normalization factors are simply $A_j = 1/L$. Clearly, since the polynomial set is fully described by a complex exponential, roots of zero do not exist for any of these polynomials in the conventional case. This means that much of the analysis presented in Section 3.1 is not directly applicable to the DFT. However, we may find the so-called *roots of unity* for these polynomials; it can be shown that for the Lth degree complex polynomial x^L that there are exactly L roots of unity for these complex exponentials [12]. These roots of unity may be found at $i = \{0, 1, \ldots, L-1\}$. These values will satisfy, for any $k, l < L$,

$$\sum_{r=0}^{L-1} P_k(x_k)\left(P_l(x_k)\right)^* = \sum_{r=0}^{L-1} e^{-j2\pi(l-m)r/L} = \delta(l-m). \quad (44)$$

In the 3×3 case, this relationship generates the transform matrix

$$\mathbf{F}_{3\times 3} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{e^{-j2\pi/3}}{\sqrt{3}} & \frac{e^{-j4\pi/3}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{e^{-j4\pi/3}}{\sqrt{3}} & \frac{e^{-j8\pi/3}}{\sqrt{3}} \end{bmatrix}. \quad (45)$$

This matrix is a transpose of the one presented in [2, Section IV.A]. This type of simple orthogonal code will be denoted as *distance preserving*. This means that at any given instant in time, for two distinct code words $\mathbf{D}_\alpha = \{s_{\alpha 1}, \ldots, s_{\alpha L}\}$ and $\mathbf{D}_\beta = \{s_{\beta 1}, \ldots, s_{\beta L}\}$, the expected value of $|s_{\alpha i} - s_{\beta i}|^2$ does not change for $1 \le i \le L$. Since the DFT-derived simple orthogonal code involves only phase shifts and any symbol transmitted at any instant in time over any antenna has a constant magnitude, this code is in fact distance preserving.

Although the DFT matrix is well known, the fact that a complex phase shift needs to be performed may not be desirable. As a result, real-number transformations may be used. For instance, the discrete cosine transform, which is derived from the roots of the Tchebychev polynomials $\{P_n(x)\} = \cos(n\cos^{-1}(x))$, yields only real matrix entries. The general form for the discrete cosine transform (DCT) matrix is

$$\mathbf{C}_{lm} = \begin{cases} \frac{1}{\sqrt{L}}, & l = 0, \\ \sqrt{\frac{2}{L}} \cos\frac{\pi(2m+1)l}{2L}, & l > 0. \end{cases} \quad (46)$$

The 3×3 matrix associated with the DCT is

$$C_{3\times 3} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}. \quad (47)$$

This matrix is not distance preserving, and as a result, the instantaneous code word symbol differences will be different from codes derived from the DFT. Assuming a QPSK constellation, an average code symbol difference may be derived, assuming perfect synchronization, however. Due to the fact that the matrix is unitary, the mean code word difference will be equivalent to the block code derived from the DFT. However, this does not imply how diversity will affect the performance of this code when other elements of a typical digital communications system are considered (e.g., trellis coding, interleaving). This particular matrix may also not be desirable, due to the fact that one of the entries is zero; this results in large peak-to-average ratios for the transmitted data. Another possible transform matrix is based on the discrete Laguerre transform [10], which is based on the Laguerre polynomials. Due to the fact that this transform is not a sinusoidal transform, no general closed form solution exists for this transform. The 3×3 discrete Laguerre transform is given below (rounded to 4 digits):

$$L_{3\times 3} = \begin{bmatrix} 0.8433 & 0.5277 & 0.1019 \\ -0.4927 & 0.6831 & 0.5392 \\ 0.2149 & -0.5049 & 0.8360 \end{bmatrix}. \quad (48)$$

This matrix, which avoids the complex phase rotation of the DFT matrix, yet does not suffer from the same power-balancing problems from which DCT-derived matrix does. However, both of these codes are not distance preserving.

Many other codes based on unitary transform matrices exist. Considering that all these codes have identical performance in terms of diversity, the code chosen would be based on not just raw symbol error rate but other criteria as well.

3.3. Enhancing diversity of simple orthogonal codes

Although (24) places an upper bound on the maximum diversity achievable by simple orthogonal codes as defined in this section, we may enhance the diversity performance of the code by implementing the code in a *closed-loop* transmit diversity method. Closed-loop transmit diversity methods are methods that rely on feedback so that using a complex weighting of each of the symbols to be transmitted from each antenna, a coherent combination is possible at the receiver. Essentially, this approach is used to pre-equalize the channel prior to transmission. In contrast, transmit diversity methods such as the Radon-Hurwitz space-time block code of (16), which do not require receiver feedback, are also classified as *open loop*.

One of the first approaches to this problem was provided in [13], where the authors proposed transmitting training sequences to several users in the network. These sequences are transmitted over L antennas. If we assume that the channel, as seen by a single user k with respect to L antenna elements at time t, may be represented by the channel vector $\mathbf{a}_k(t) = [a_{k1}(t)a_{k2}(t)\cdots a_{kL}(t)]$, where $a_{ki}(t)$ is the complex channel response for antenna i with respect to user k at time t, then the transmitter can make use of this information to scale each antenna input accordingly so that a coherent combination of the signals from each antenna is possible at the receiver. Thus, if the receiver estimates the channel from each antenna as $\hat{\mathbf{a}}_k(t) = [\hat{a}_{k1}(t)\hat{a}_{k2}(t)\cdots\hat{a}_{kL}(t)]$, then these estimates may be relayed to the transmitter. Thus, if we assume that the signal $d(t)$ is transmitted from each antenna at time t, the received signal after scaling would be

$$r(t) = \hat{a}_{k1}^*(t)a_{k1}(t) + \hat{a}_{k2}^*(t)a_{k2}(t) + \cdots + \hat{a}_{kL}^*(t)a_{kL}(t) + n(t), \quad (49)$$

where $n(t)$ is an additive Gaussian noise term. Clearly, if

$\hat{a}_{ki}(t) \approx a_{ki}(t)$ then the received SNR is maximized. This approach has been narrowed to include quantized relative phase feedback in [14]. This approach has also been addressed for two antennas in CDMA systems in [15].

However, the amount of coding given to the feedback information and the latency of the feedback information become critical to performance of these systems. As a result, these systems tend to actually degrade performance with respect to space-time block coded systems such as the 2×2 Radon-Hurwitz transformation at high mobile speeds. For instance, in [16] the authors present a theoretical framework for the performance of closed-loop transmit diversity and demonstrate how the performance degrades at high Doppler with respect to the Radon-Hurwitz code as a result of feedback latency. More specifically, in [17] the authors show a severe degradation in performance of a 2-antenna closed-loop method versus a 2-antenna Radon-Hurwitz transformation at speeds of 30 km/h or greater at 2 GHz carrier frequency in a CDMA system. Moreover, with respect to one-path Rayleigh fading conditions in a CDMA system, results presented in [15] actually demonstrated worse performance for closed-loop transmit diversity methods with respect to not using any diversity methods at all at speeds of 100 km/h under certain high SNR conditions due to the additional degradation provided by fast power control.

In addition, closed-loop systems require increased bandwidth for feedback information as the number of antennas increase. Balancing this need with the need for reliability on the feedback information could result in suboptimal performance for a large number of antennas.

However, the use of simple orthogonal block codes could be used to address these problems in a closed-loop implementation. Assume that we have a block of K transformations to choose from for modulating the input data matrix into the transmit antenna array. If each of these $L \times L$ block transformations can be grouped as $\mathbf{T} = [\mathbf{T}_1 \ \mathbf{T}_2 \ \cdots \ \mathbf{T}_K]$, then knowing the channel estimates from each antenna, the receiver may make a prediction of the best available transform and feed this information back to the transmitter. Since the transforms may be generated for arbitrary dimensionality (as shown in Section 3), the feedback requires $\log_2 K$ bits for an arbitrary number of antennas. Assuming that the estimated channel vector is still $\hat{\mathbf{a}}_k(t) = [\hat{a}_{k1}(t) \ \hat{a}_{k2}(t) \ \cdots \ \hat{a}_{kL}(t)]$ and that this channel estimate remains relatively constant over the L time epochs of the block code, then the receiver transform selection $\hat{\mathbf{T}}(t)$ for feedback that maximizes SNR would be

$$\hat{\mathbf{T}}(t) = \max_{\mathbf{T}_i \in \mathbf{T}} \left\| \mathbf{T}_i \mathbf{a}_k^T(t) \right\|. \tag{50}$$

This transform selection may be sent to the transmitter for application in the ensuing data sequence.

Each data sequence to which a transform is applied should include a means of error detection, for example, a cyclic redundancy check (CRC). This is necessary due to the fact that the feedback of the transform selection may not be implemented due to feedback error. However, using an error detection mechanism such as a CRC, the receiver may decode the received data sequence using multiple hypotheses testing, with up to K hypotheses. A simple decoding algorithm at the receiver may be attained:

(1) determine appropriate transform for the next data sequence and relay selection to transmitter;
(2) for the next received data sequence, apply selected transform and decode. If CRC passes, return to step (1) for next data sequence;
(3) if CRC fails, sequentially apply each of the other $K - 1$ possible transforms to the received data and decode. If a CRC passes for a transform, return to step (1) for next data sequence;
(4) classify the received data sequence as an erasure. Return to step (1) for next data sequence.

The drawback of this type of method is that using CRCs for short data sequences could severely impact throughput. As a result, this type of feedback mechanism would in practice perform relatively slowly with respect to channel conditions. On the other hand, this method is not as sensitive to feedback errors as the method described in [15] due to the use of multiple hypotheses testing. More importantly, however, this method will still provide diversity gains at fast fading conditions, despite the fact that the feedback mechanism is highly inaccurate in these types of channel conditions. This is due to the fact that these simple orthogonal block codes provide at least the diversity order given in (27). Therefore, for instance, a 3-antenna code for a QPSK constellation will always provide mean diversity order of nearly 2, regardless of feedback error.

4. EXAMPLE: QPSK SYSTEM

The 3×3 block codes presented in Section 3.2 were simulated in a simple QPSK system under single-path Rayleigh fading conditions [18]. It was not merely of interest to determine the benefit of the proposed codes versus no diversity, but also to measure the difference in performance between the dual and triple antenna cases. For the closed-loop method, codes used were based on the DFT, DCT, and discrete Laguerre transform (DLT) as described in Section 3.2. In addition, the conjugate transposes of these matrices were also used for the closed-loop method. For comparison, the 2×2 Radon-Hurwitz code and 2×2 closed-loop method results were provided, in addition to ideal triple-diversity results. It should be noted that only two transforms were used for the 2-antenna closed-loop method, the DLT and the DFT. This is due to the fact that the 2×2 DCT transform is identical to the 2×2 DFT.

The system under consideration was a QPSK system that mapped two bits to each constellation point. It is assumed that pilot signals from different antennas arrive at the receiver simultaneously and are separated using orthogonal waveform modulation; for simulation purposes, however, perfect channel knowledge at the receiver was assumed. If the signals from different antennas did not arrive simultaneously, then self-interference would occur due to imperfect suppression

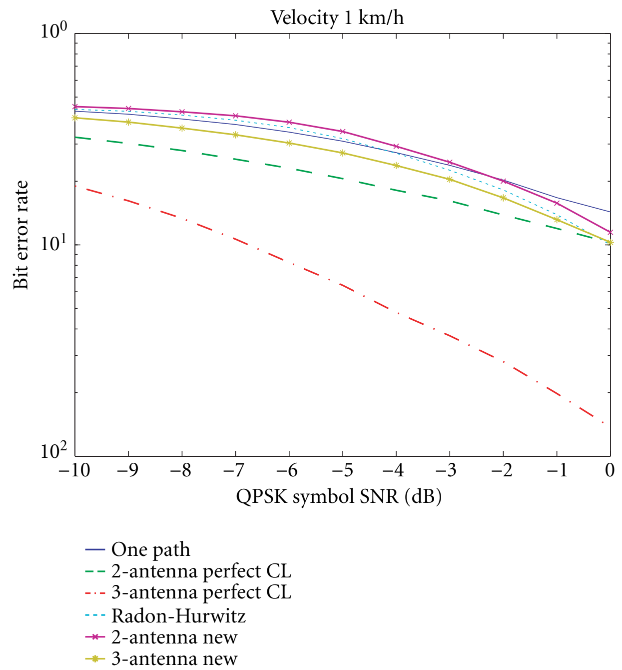

FIGURE 2: 1 km/h results.

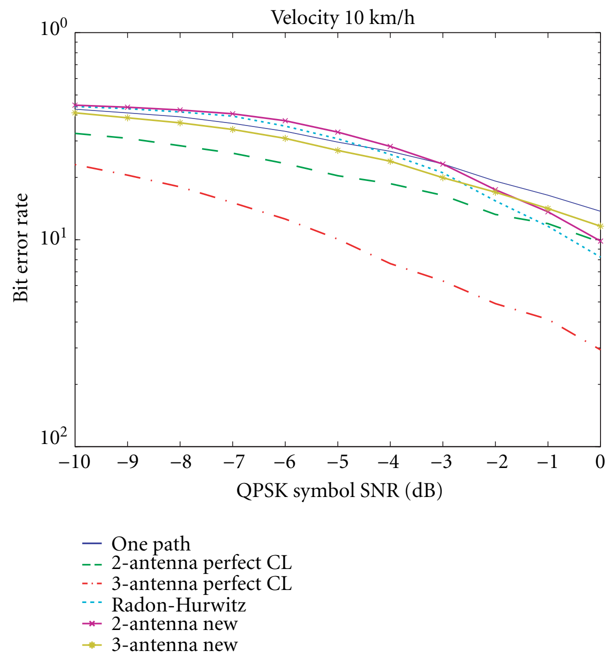

FIGURE 3: 10 km/h results.

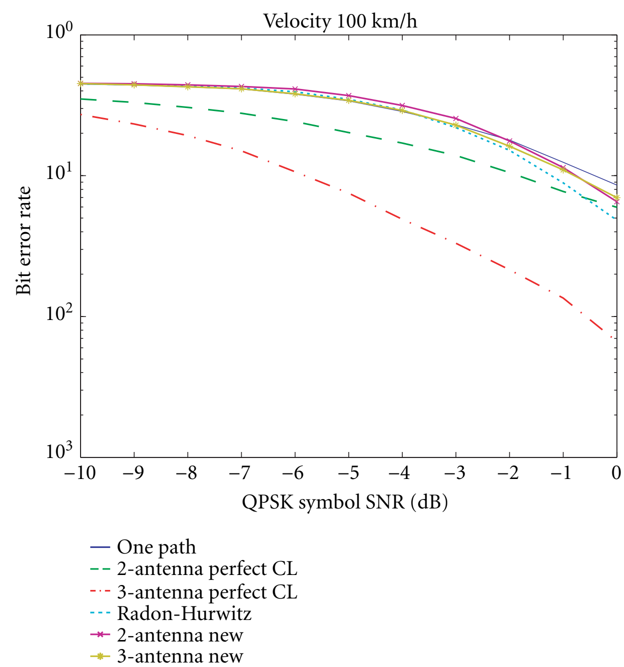

FIGURE 4: 100 km/h results.

of other antenna signals when demodulating the signal from a particular antenna.

An information source at 12.8 kbit/s was assumed. This source was passed into a rate 1/3, constraint length 9, convolutional encoder, and block interleaved. The interleaved data was then modulated using a QPSK constellation. No power control was assumed. The carrier frequency assumed was 1960 MHz. Under such conditions, the diversity performance for different space-time block coding methods may be isolated for evaluation. The metric for performance, however, was BER after decoding. 192000-bit simulations were run for each given code, velocity, and QPSK symbol SNR. The results for 1 km/h, 10 km/h, and 100 km/h are shown in Figures 2, 3, and 4. In these figures, single path results are provided, and the proposed closed-loop method results for two and three antennas are designated as *2-antenna new* and *3-antenna new.* In addition, perfect closed-loop transmit diversity results are provided that emulate closed-loop transmit diversity with no feedback delay or error. The results for 2 and 3 antennas are designated as *2-antenna perfect CL* and *3-antenna perfect CL*, respectively.

The simulation results show benefits not only to space-time block coding but also to increasing from two antennas (Radon-Hurwitz) to three (using the proposed method) in certain situations, particularly at low SNR (as much as 4 dB performance improvement at 1 km/h velocity). The new closed-loop methods did start to degrade in the 3-antenna case versus the Radon-Hurwitz block code at high velocities and high SNR, but this is most likely due to the limited set of transform choices. At high speeds, the 3-antenna code not only performed well with respect to the Radon-Hurwitz at low SNR, but also provided very little degradation (less than 0.5 dB) at high SNR. Therefore, the proposed method shows promise in increasing the *crossover Doppler frequency* [16], that is, the Doppler frequency at which an open-loop method such as the Radon-Huwitz transform outperforms a closed-loop method. This is a result of the proposed method reverting back to the performance bounds described by the mean diversity performance derived in (28). The 2-antenna closed-loop method did not provide quite the gains of the 3-antenna closed-loop method, but this was also most likely due to an even more limited transform set size than the 3-antenna case (once again, due to the fact that many of the transform kernels used provide the exact same transform matrix in the 2×2 case).

5. CONCLUSIONS

A general framework for deriving space-time block codes was presented. This framework involves starting with sets of orthonormal polynomials and deriving unitary transform matrices from these sets. These transform matrices may in turn be used to generate orthogonal spatial block codes. Simulation results in a closed-loop deployment show benefit for this approach to code generation as opposed to the approach presented in [4] under certain scenarios, as these codes may be defined for arbitrary dimensions and their usage in the proposed closed-loop framework did not result in a significant degradation in performance at high velocities. However, since these codes do not maximize diversity in the mean-sense for a given dimensionality, further analysis should be performed on methods for increasing the diversity of these codes in typical wireless environments.

APPENDIX

CODE WORD PAIR DIFFERENCE PROBABILITY

Assume a diversity transformation of rate 1 using L antennas, and a symbol constellation set of cardinality M. Each possible code word may be represented as a base-M number consisting of L digits. All possible code words may be listed as follows:

$$\begin{array}{ccc}
0 & \overbrace{0 \cdots 0}^{L-2} & 0 \\
0 & \overbrace{0 \cdots 0}^{L-2} & 1 \\
\vdots & \vdots & \vdots \\
0 & \overbrace{0 \cdots 0}^{L-2} & (M-1) \\
0 & \overbrace{0 \cdots 1}^{L-2} & 0 \\
\vdots & \vdots & \vdots \\
1 & \overbrace{0 \cdots 0}^{L-2} & 0 \\
\vdots & \vdots & \vdots \\
1 & \overbrace{0 \cdots 0}^{L-2} & (M-1) \\
1 & \overbrace{0 \cdots 1}^{L-2} & 0 \\
\vdots & \vdots & \vdots \\
(M-1) & \overbrace{(M-1) \cdots (M-1)}^{L-2} & (M-1).
\end{array} \tag{A.1}$$

If we examine only the code words designated by the numerals

$$0 \ \overbrace{0 \ \cdots \ 0}^{L-2} \ 0 \text{ through } 0 \ \overbrace{0 \ \cdots \ 0}^{L-2} \ (M-1),$$

it can be seen that there are $\sum_{i=0}^{M-2}(M-1)-i$ code word pairs that differ in only one position. Similarly, if we examine only the code words designated by the numerals

$$0 \ \overbrace{0 \ \cdots \ 1}^{L-2} \ 0 \text{ through } 0 \ \overbrace{0 \ \cdots \ 1}^{L-2} \ (M-1),$$

it can be seen that there are still $\sum_{i=0}^{M-2}(M-1)-i$ code word pairs that differ in only one position. In fact, if we examine the code words designated by the numerals

$$s_0 \ \overbrace{s_1 \ \cdots \ s_{L-2}}^{L-2} \ 0 \text{ through } s_0 \ \overbrace{s_1 \ \cdots \ s_{L-2}}^{L-2} \ (M-1)$$

for arbitrary symbols $s_0, s_1, \ldots, s_{L-2}$, then the same number of code word pairs differing in one position remains as $\sum_{i=0}^{M-2}(M-1)-i$. As a result, there are $M^{L-1}\sum_{i=0}^{M-2}(M-1)-i$ code word pairs that only differ in the last position. This relationship also holds true for code word pairs differing only in the second-to-last position, and so on for all remaining $L-2$ positions. As a result, the total number of code word pairs differing in only one position is $A_1 = LM^{L-1}\sum_{i=0}^{M-2}(M-1)-i$.

The next case to be examined is the number of code word pairs that differ in all L positions. For instance, take the code words defined by $\{0 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$, that is, code words which have 0 for the first digit. For any given value of $\{0 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$, there exist $(M-1)^L$ code word pairs which differ from $\{0 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$ in L positions. Since there are M^{L-1} possible values for $\{0 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$, there exist $M^{L-1}(M-1)^L$ code word pairs that differ in L positions for all possible values of $\{0 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$. Now examine the code words defined by $\{1 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$, that is, code words which have 1 for the first digit. For any given value of $\{1 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$, there exist $(M-1)^L$ code word pairs that differ in all L positions. Among these code words, $(M-1)^{L-1}$ have 0 as the first digit. If we assume that these code word pairs were already accounted for when we examined code words of the structure $\{0 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$, then for all possible values of $\{1 \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$ there exist $M^{L-1}\lfloor(M-1)^L-(M-1)^{L-1}\rfloor$ additional code word pairs that differ in all L positions. Using this reasoning, we can say that, for a given i $(0 \le i < M)$, for all possible values of $\{i \ s_1 \ \cdots \ s_{L-2} \ s_{L-1}\}$, there exist $M^{L-1}\lfloor(M-1)^L-i(M-1)^{L-1}\rfloor$. As a result, the total number of code word pairs differing in L positions is

$$A_L = \sum_{i=0}^{M-2} M^{L-1}[(M-1)^L - i(M-1)^{L-1}]. \tag{A.2}$$

REFERENCES

[1] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, USA, 3rd edition, 1995.

[2] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Communications*, vol. 47, no. 4, pp. 527–537, 1999.

[3] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE Vehicular Technical Conference*, pp. 136–140, Atlanta, Ga, USA, April 1996.

[4] 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.

[5] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 451–460, 1999.

[6] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[7] J. Grimm, M. P. Fitz, and J. V. Krogmeier, "Further results on space-time coding for Rayleigh fading," in *Allerton Conference on Communication, Control, and Computing*, pp. 391–400, Monticello, Ill, USA, September 1998.

[8] D. M. Ionescu, "New results on space-time code design criteria," in *Proc. IEEE Wireless Communications and Networking Conference*, New Orleans, La, USA, September 1999.

[9] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[10] G. Mandyam and N. Ahmed, "The discrete Laguerre transform: derivation and applications," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 2925–2931, 1996.

[11] V. I. Krylov, *Approximate Calculation of Integrals*, MacMillan, New York, USA, 1962.

[12] R. Reinhold, *Theory of Complex Functions*, Springer-Verlag, New York, NY, USA, 1991.

[13] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna arrays with feedback," *IEEE Signal Processing Letters*, vol. 1, no. 10, pp. 150–152, 1994.

[14] R. W. Heath and A. J. Paulraj, "A simple scheme for transmit diversity using partial channel feedback," in *Proc. 32nd Annual Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 1073–1078, Pacific Grove, Calif, USA, November 1998.

[15] K. Rohani, M. Harrison, and K. Kuchi, "A comparison of base station transmit diversity methods for third generation cellular standards," in *Proc. IEEE Vehicular Technical Conference*, pp. 939 –943, Houston, Tex, USA, May 1999.

[16] E. N. Onggosanusi, A. Gatherer, A. G. Dabak, and S. Hosur, "Performance analysis of closed-loop transmit diversity in the presence of feedback delay," *IEEE Trans. Communications*, vol. 49, no. 9, pp. 1618–1630, 2001.

[17] M. Harrison and K. Kuchi, "Open and closed loop transmit diversity at high data rates on 2 and 4 elements," Third-Generation Partnership Project 2 (3GPP2) Contribution 3GPP2-C30-19990817-017, August 1999, Portland, Ore, USA.

[18] G. D. Mandyam, "Application of unitary transforms to quasi-closed-loop transmit diversity systems," in *Asilomar Conference on Signals, Systems, and Computers*, pp. 97–101, Pacific Grove, Calif, USA, October 2000.

Giridhar D. Mandyam is the Research Manager of the Wireless Data Access Group at Nokia Research Center, Irving, Texas. He received his B.S. degree *Magna Cum Laude* in electrical engineering from Southern Methodist University (Dallas, Texas) in 1989, the M.S. degree in the same field from the University of Southern California (Los Angeles, California) in 1993, and the Ph.D. degree in electrical engineering from the University of New Mexico (Albuquerque, New Mexico) in 1996. He has worked for several companies on wireless communications equipment, including Qualcomm and Texas Instruments. In 1998, he joined Nokia, where he has worked on standardization and implementation concepts for cdma2000, 1X-EV, and WCDMA. He has authored or coauthored over 40 journal and conference publications and four book chapters. He also holds four US patents in the area of wireless communications technology.