

Systematic Design of Space-Time Trellis Codes for Diversity and Coding Advantages

Zoltan Safar

*Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA
Email: zsafar@eng.umd.edu*

K. J. Ray Liu

*Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA
Email: kjrlui@eng.umd.edu*

Received 15 June 2001 and in revised form 17 January 2002

The emerging need for high data rate wireless services has raised considerable interest in space-time coding. In this work, we propose a systematic code construction method that jointly considers diversity advantage and coding advantage for an arbitrary number of transmit antennas and any memoryless constellation. Our approach is to directly assign channel symbols to transmit antennas at different states by exploiting the properties of the state transitions in the trellis. The code construction problem is reduced to a combinatorial optimization problem and a computationally efficient suboptimal solution is proposed. The flexibility of the method is demonstrated by designing space-time trellis codes for QPSK, 8PSK, 16PSK, asymmetric QPSK and 4ASK constellations. Space-time code construction for a large number of transmit antennas (6, 8, and 10) is also considered. The simulations show that our design procedure results in codes that outperform the ones constructed by previously existing methods. The achievable performance gain is governed by the distance structure of the chosen constellation.

Keywords and phrases: space-time codes, trellis codes, transmit diversity, wireless communications.

1. INTRODUCTION

In wireless communications, the transmitted signal undergoes severe distortion due to multipath fading and interference from other users. The design of communication systems that offer reliable transmission at high data rates is a challenging task in such circumstances. To combat the adverse effects of the radio signal propagation environment, several diversity techniques have been developed [1]. These techniques offer increased protection against the channel-induced distortion by providing multiple versions of the transmitted signal to the receiver.

Time diversity, for example, represents redundancy in the time domain. A form of time diversity, channel coding combined with interleaving, adds redundancy with a certain algebraic structure that can be exploited to detect and correct transmission errors [2]. Another widely used diversity technique is spatial diversity, which adds redundancy in the spatial domain. By building a system with multiple transmit and/or receive antennas and making use of the larger number of signal propagation paths between the transmitter and the receiver, the quality of the wireless link can also be improved [1].

Information-theoretic works [3, 4, 5, 6, 7] have shown that the capacity of fading channels is substantially increased when using multiple transmit and/or receive antennas. The results of [6] for the quasi-static, flat Rayleigh fading channel model can be summarized as follows. If the channel gains between the transmit and receive antennas are assumed to be known at the receiver side, the channel capacity is approximately proportional to the minimum of the number of transmit and receive antennas.

Space-time (ST) trellis codes have been proposed as a means to exploit the potential for capacity increase in multi-antenna systems. The performance criteria were derived in [8, 9], characterizing the ST codes with two quantities: the diversity advantage, which describes the asymptotic error rate decrease as a function of the signal-to-noise ratio (SNR), and the coding advantage, which determines the vertical shift of the error performance curve. In [9], the authors proposed design rules for two transmit antennas to achieve the maximum diversity advantage. They also derived a lower bound on the complexity of the encoder and the decoder for the desired diversity advantage and data throughput. This lower bound states that in order to achieve a diversity advantage of K and transmit one B -ary source symbol per state transition,

the encoder and the decoder must have at least $N_{\min} = B^{K-1}$ states.

The repetition coded delay diversity scheme described in [10] was the first systematic design rule for an arbitrary number of transmit antennas. Using this method, ST codes achieving full diversity advantage can be designed for arbitrary constellations and encoders with N_{\min} states. This work also introduced the idea of zero symmetry to constrain computer search for ST codes with more than two antennas.

In [11], the design problem was transformed into the binary domain. The code design was based on the finite field counterpart of the ST code performance criteria for full spatial diversity. The authors proposed code design procedures for an arbitrary number of transmit antennas and an arbitrary number of states, but only for BPSK and QPSK constellations. Moreover, the design methods in [10, 11] for full diversity advantage uniquely determine the ST codes, not leaving room to improve the coding advantage.

In this paper, we propose a systematic design method based on an alternative approach: we exploit the structure of the trellis to design ST codes that provide full diversity advantage for an arbitrary number of transmit antennas, arbitrary number of encoder states (as long as it satisfies the lower bound) and arbitrary memoryless constellations. Our method can be treated as a generalization of the results of [9, 10]. The design rules for full diversity advantage do not specify the ST codes completely, offering the possibility to further optimize for coding advantage.

Moreover, we develop a code design procedure that benefits from this possibility for the important special case of encoders with N_{\min} states. Based on the design rules for diversity advantage, we reduce the code construction problem to a combinatorial optimization problem and propose a computationally efficient suboptimal solution. It seems that this is the first work that considers systematic code design for both diversity advantage and coding advantage.

The rest of the paper is organized as follows. Section 2 introduces the mathematical model of the communication system. The performance criteria for ST trellis codes are also described in this section. The code construction method will be developed in Sections 3 and 4. Section 5 describes specific ST code construction examples, and the simulation results demonstrating the performance of these codes are provided in Section 6. Finally, we draw some conclusions in Section 7.

2. SYSTEM MODEL AND PERFORMANCE CRITERIA

In this section, the mathematical model of the wireless communication system under study is described. The notation used throughout this paper is introduced. Then, we summarize the relevant results of the previous works and briefly restate the performance criteria derived in [8, 9]. These criteria serve as a basis for the development of our systematic design procedure.

Consider a wireless communication system with K transmit and L receive antennas (the transmit antennas are indexed by k , $k \in \{0, 1, \dots, K-1\}$, and the receive antennas are indexed by l , $l \in \{0, 1, \dots, L-1\}$). The input bit stream

is divided into b_s bit long blocks, forming B -ary ($B = 2^{b_s}$) source symbols. The ST encoder works as a finite state machine with N states: it takes the current b_s bit long source symbol, b_t ($b_t \in \{0, 1, \dots, B-1\}$) at discrete time t ($t = 0, 1, 2, 3, \dots$), and governed by this input and the current state, S_t ($S_t \in \{0, 1, \dots, N-1\}$), it moves to the next state, S_{t+1} . During this state transition, the encoder outputs K B -ary channel symbol indices, one for each transmit antenna. We denote by $i^k(S_t, b_t)$ the channel symbol index for transmit antenna k , generated during the state transition from S_t when the current input source symbol is b_t . We also use the channel symbol index vector, defined as

$$\mathbf{i}(S_t, b_t) = [i^0(S_t, b_t), i^1(S_t, b_t), \dots, i^{K-1}(S_t, b_t)]^T. \quad (1)$$

These channel symbol indices select one of the B different waveforms for each antenna, and the selected waveforms are transmitted simultaneously through the transmit antennas. In the sequel, $c(i)$ will be the complex baseband vector-space representation of the i th passband waveform ($i \in \{0, 1, \dots, B-1\}$). $c(i)$ will also be referred to as the i th constellation point or channel symbol. All the constellations are assumed to be normalized so that the average energy of the constellation is unity (if the channel symbols are equally likely). $c(i^k(S_t, b_t))$ will denote the constellation point output by antenna k when the current state is S_t and the current input is b_t . The vector of channel symbols is given by

$$\mathbf{c}(S_t, b_t) = [c(i^0(S_t, b_t)), c(i^1(S_t, b_t)), \dots, c(i^{K-1}(S_t, b_t))]^T. \quad (2)$$

The transmission medium is assumed to be flat (frequency nonselective), quasi-static, Rayleigh fading channel. The quasi-static property means that the channel remains constant over a certain time, called frame period, and changes independently from one frame to the other. α_{kl} will represent the path gain from transmit antenna k to receive antenna l . These path gains are modeled as independent, complex, zero mean, circularly symmetric Gaussian random variables with unit variance. Furthermore, some additional assumptions are made to facilitate the analysis. First, the receiver has knowledge of the α_{kl} propagation coefficients. Second, the receiver is perfectly synchronized with the transmitter.

Based on the above assumptions, after down-conversion, matched filtering and sampling, the received signal at receive antenna l , at discrete time t can be expressed as [9]

$$r_t^l = \sum_{k=0}^{K-1} \sqrt{\frac{E_0}{K}} \alpha_{kl} c(i^k(S_t, b_t)) + z_t^l = s_t^l + z_t^l, \quad (3)$$

where E_0 is the total average transmission energy per transmitted source symbol (each transmit antenna transmits the channel symbols with E_0/K average transmit energy), s_t^l and z_t^l stand for the received signal and noise components, respectively, and z_t^l 's are independent, complex, zero mean, circularly symmetric Gaussian random variables with variance

N_0 . Consequently, the average SNR per source symbol at receive antenna l becomes

$$\text{SNR}_l = \frac{E[|s_t^l|^2]}{E[|z_t^l|^2]} = \frac{E_0}{N_0}. \quad (4)$$

Assume that the previously described transmitter sends T ($T > K$) B -ary source symbols to the receiver. The ST encoder, while encoding the data, goes through the following sequence of states:

$$S_0 \xrightarrow{b_0} S_1 \xrightarrow{b_1} S_2 \xrightarrow{b_2} \dots \xrightarrow{b_{T-2}} S_{T-1} \xrightarrow{b_{T-1}} S_T. \quad (5)$$

In words, the encoder starts in S_0 , takes the first input b_s -tuple, b_0 , moves to S_1 , and so on. As a result of this state transition sequence, the encoder produces the channel symbol vector sequence

$$\mathbf{c}(S_0, b_0), \mathbf{c}(S_1, b_1), \dots, \mathbf{c}(S_{T-1}, b_{T-1}). \quad (6)$$

The above vector sequence can be arranged into a $K \times T$ matrix, \mathbf{C} :

$$\mathbf{C} = [\mathbf{c}(S_0, b_0), \mathbf{c}(S_1, b_1), \dots, \mathbf{c}(S_{T-1}, b_{T-1})]. \quad (7)$$

The decoder, due to decoding errors, goes through a different sequence of states,

$$S'_0 \xrightarrow{b'_0} S'_1 \xrightarrow{b'_1} S'_2 \xrightarrow{b'_2} \dots \xrightarrow{b'_{T-2}} S'_{T-1} \xrightarrow{b'_{T-1}} S'_T, \quad (8)$$

producing the erroneously decoded source symbol sequence $\{b'_t\}$ and the $K \times T$ channel symbol matrix \mathbf{C}' :

$$\mathbf{C}' = [\mathbf{c}(S'_0, b'_0), \mathbf{c}(S'_1, b'_1), \dots, \mathbf{c}(S'_{T-1}, b'_{T-1})]. \quad (9)$$

We can define \mathbf{D} , the channel symbol difference matrix as $\mathbf{D} = \mathbf{C} - \mathbf{C}'$, and a $K \times K$ matrix \mathbf{A} as $\mathbf{A} = \mathbf{D}\mathbf{D}^H$. Let \mathbf{A} (and \mathbf{D}) be of rank m . Since \mathbf{A} is Hermitian and nonnegative definite, its eigenvalues are real and nonnegative. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ be the nonzero eigenvalues of \mathbf{A} . Given the earlier described channel model, it can be shown [9] that the probability that the decoder erroneously decodes \mathbf{C}' if \mathbf{C} was sent can be upper bounded as

$$P(\mathbf{C}'|\mathbf{C}) \leq \left(\prod_{i=1}^m \lambda_i \right)^{-L} \left(\frac{E_0}{4KN_0} \right)^{-mL}. \quad (10)$$

The performance criteria [8, 9] were derived to minimize $P(\mathbf{C}'|\mathbf{C})$ for a given E_0 and N_0 :

(1) *Design for full spatial diversity (rank criterion)*: the matrix \mathbf{D} must be of full row rank for any distinct \mathbf{C} and \mathbf{C}' matrices. (Then we have $m = K$.) In this case, a diversity advantage of KL has been achieved.

(2) *Design for coding advantage (determinant criterion)*: the minimum determinant of \mathbf{A} taken over all distinct \mathbf{C} and \mathbf{C}' matrices must be as large as possible. If the minimum determinant is γ , then a coding advantage of $\sqrt[\gamma]{\gamma}$ has been achieved.

Note that the above performance criteria are not constructive, that is, they do not provide a systematic method to construct good ST codes.

3. DESIGN FOR DIVERSITY ADVANTAGE

3.1. Trellis structure analysis

The goal of this subsection is to analyze the algebraic structure of the trellis of the ST encoder and find closed form expressions that relate the state sequence $\{S_t\}$ to the starting state S_0 and the input source symbol sequence $\{b_t\}$. The additions and the multiplications are assumed to be standard integer operations. The modulo operation will always be written explicitly to avoid ambiguity.

Assume that the encoder has $N = RB^{K+p-1}$ states, with $R = 2^r$, $B = 2^{b_s}$, $b_s > 0$, $p \geq 0$, and $0 \leq r < b_s$. Therefore, it satisfies the lower bound of [9] for desired diversity advantage of K with B -ary source symbols. Any large enough power of 2 number can be put into this form; the purpose of this representation is to make the analytical treatment easier. The number N is simply decomposed into the product of two numbers: the first number, R is less than B , and the other number is a power of B .

The state transition of the encoder at time t is determined by the previous state, S_{t-1} ($S_{t-1} \in \{0, 1, \dots, N-1\}$), and the previous B -ary input, b_{t-1} ($b_{t-1} \in \{0, 1, \dots, B-1\}$). Analytically it can be described as

$$\begin{aligned} S_t &= (BS_{t-1} + b_{t-1}) \bmod N \\ &= B(S_{t-1} \bmod (RB^{K+p-2})) + b_{t-1}. \end{aligned} \quad (11)$$

It is shown in the appendix that we can unfold this recursion and obtain a closed form expression for S_t , $1 \leq t \leq K+p-1$:

$$S_t = B^t (S_0 \bmod (RB^{K+p-t-1})) + \sum_{m=0}^{t-1} B^{t-1-m} b_m, \quad (12)$$

and for S_{K+p} :

$$S_{K+p} = B^{K+p-1} (b_0 \bmod R) + \sum_{m=1}^{K+p-1} B^{K+p-1-m} b_m. \quad (13)$$

Based on these analytical results, we can deduce some important information about the error path structure of the trellis. Assume that the first decoding error occurs at state S_0 , that is, the correct and the erroneous paths diverge at S_0 , and they merge at some later state. As a consequence of this assumption, we have $S_0 = S'_0$ and $b_0 \neq b'_0$. We have no information regarding the rest of the b_t 's and b'_t 's. In this case, (12) immediately tells us that for $1 \leq t \leq K+p-1$, $S_t \neq S'_t$. Moreover, from (13), it can be seen that S_{K+p} may or may not be equal to S'_{K+p} , since $b_0 \bmod R$ may be equal to $b'_0 \bmod R$, even though $b_0 \neq b'_0$. Thus, we have the following theorem.

Theorem 1. *If $R = 1$ (i.e., $r = 0$), the shortest error path is exactly $K+p$ long. If $R > 1$ (i.e., $r > 0$), the shortest error path is either $K+p$ long or longer. For arbitrary $p \geq 0$ and*

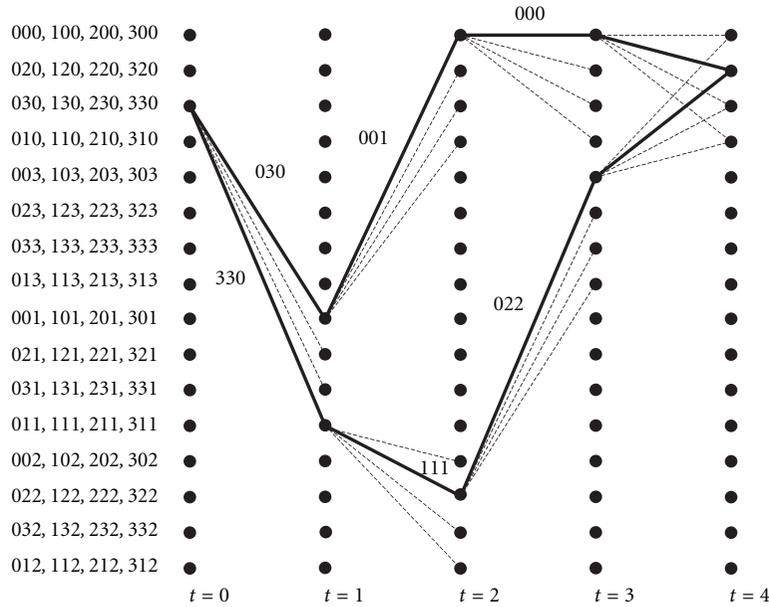


FIGURE 1: Example ST code for three antennas, QPSK.

$0 \leq r < b_s$, the shortest error path is at least K long, that is, the paths diverging at S_0 can merge only at S_K or later.

3.2. Design for full diversity

Using formulas (12) and (13), we derive design rules that guarantee that the ST trellis code achieves full diversity advantage. First, we obtain sufficient conditions to make the channel symbol difference matrix corresponding to the first K long error path segment of the first error event full rank. Afterwards, the results will be extended to arbitrary channel symbol difference matrices.

In the ST encoder, B channel symbol index vectors are assigned to each state, according to the branches emanating from that state. The current source symbol selects one of them, and the k th ($k = 0, 1, \dots, K-1$) index of the chosen vector determines the constellation point for antenna k . Figure 1 depicts an example ST code for three antennas and QPSK constellation ($K = 3, B = 4, N = 16$). In this case, if the current state is state 2 and the value of the current source symbol is 3, the ST encoder selects the third channel symbol index vector, $[3, 3, 0]^T$, and moves to state 11. The zeroth, first, and second antennas will transmit the channel symbols corresponding to the indices 3, 3, and 0, respectively.

Suppose that the transmitter sends T ($T > K$) source symbols. Without loss of generality, we can assume that the first decoding error event occurs at S_0 , making the correct and the decoded paths diverge. For now, we are concerned only about the first K long segment of all error paths of length K or larger immediately after the first error event has occurred. Our goal is to construct the $K \times K$ channel symbol difference matrix \mathbf{D}_1 , defined as

$$\mathbf{D}_1 = \begin{bmatrix} \mathbf{c}(S_0, b_0) - \mathbf{c}(S'_0, b'_0), \mathbf{c}(S_1, b_1) \\ -\mathbf{c}(S'_1, b'_1), \dots, \mathbf{c}(S_{K-1}, b_{K-1}) - \mathbf{c}(S'_{K-1}, b'_{K-1}) \end{bmatrix}, \quad (14)$$

in such a way that it is of full rank for any possible correct and erroneous paths through the trellis. Our method is to make \mathbf{D}_1 upper triangular with nonzero diagonal elements. We exclude all ST codes that do not produce upper triangular \mathbf{D}_1 matrices, so the resulting ST codes may not be optimal. However, what we gain is a problem formulation that leads to a simple solution.

The $S_0 \rightarrow S_1$ state transition is special since both the correct and the erroneous paths start at the same state. The goal is to set the zeroth entry of the zeroth column of \mathbf{D}_1 to a nonzero value and to zero out the rest of the entries in that column. This can be achieved by the following conditions that form the first half of the design rules:

- (1a) The zeroth indices of the channel symbol index vectors at the same state must be different.
- (1b) The remaining indices of the channel symbol index vectors at the same state must be the same.

In our example, assume that the $b_0 = 0$ (top) path is the correct path and the $b'_0 = 3$ (bottom) path is the erroneously decoded path. The channel symbol index vectors $[0, 3, 0]^T$ and $[3, 3, 0]^T$ have different zeroth indices, but the first and second indices are the same; therefore, the zeroth column of the \mathbf{D}_1 matrix will be $[1 + j, 0, 0]^T$.

For the rest of the state transitions $S_t \rightarrow S_{t+1}$, $t = 1, 2, \dots, K-1$, the objective is to set the t th entry of the t th column of \mathbf{D}_1 to a nonzero value and to zero out all the entries below the t th entry in that column. To facilitate the explanation, we introduce the following definitions.

Definition 1. A level t group is a collection of all destination states that can be reached at state transition t from a given S_0 starting state through all possible b_0, b_1, \dots, b_{t-1} input sequences.

Definition 2. A subgroup of a level t group is a collection of all destination states that can be reached at state transition t from a given S_0 starting state and a given b_0 starting branch through all possible b_1, b_2, \dots, b_{t-1} input sequences.

In order to effectively use these definitions in the design procedure, we need to describe the relationship between the encoder states and the groups and subgroups at different levels. Equation (12) expresses the state transition at time t , $t = 1, 2, \dots, K - 1$, as a function of the starting state, S_0 , and the source symbol sequence b_0, b_1, \dots, b_{t-1} . Because S_0 is kept constant in Definition 1 for all possible b_0, b_1, \dots, b_{t-1} sequences, we can eliminate the effect of the starting state by taking modulo B^t of both sides of (12). Therefore, the expression $S_t \bmod B^t$ will describe how S_t depends on the b_0, b_1, \dots, b_{t-1} sequence for an arbitrary, but fixed, S_0 starting state. From (12), we obtain

$$S_t \bmod B^t = B^{t-1}b_0 + B^{t-2}b_1 + \dots + Bb_{t-2} + b_{t-1}. \quad (15)$$

The above quantity can be thought of as a t digit B -ary number. As the input B -tuples $(b_0, b_1, \dots, b_{t-1})$ vary from $(0, 0, \dots, 0)$ to $(B - 1, B - 1, \dots, B - 1)$, the value of $S_t \bmod B^t$ varies from 0 to $B^t - 1$. Consequently, for $t = 1, 2, \dots, K - 1$, any level t group starts at state m such that $m \bmod B^t = 0$ and consists of B^t consecutive states.

Similarly, S_0 and b_0 are kept constant in Definition 2, so the expression $S_t \bmod B^{t-1}$ will describe how S_t changes as a function of b_1, b_2, \dots, b_{t-1} . From (12), we have

$$S_t \bmod B^{t-1} = B^{t-2}b_1 + B^{t-3}b_2 + \dots + Bb_{t-2} + b_{t-1}. \quad (16)$$

The above quantity can be thought of as a $t - 1$ digit B -ary number. As the input B -tuples $(b_1, b_2, \dots, b_{t-1})$ vary from $(0, 0, \dots, 0)$ to $(B - 1, B - 1, \dots, B - 1)$, the value of $S_t \bmod B^{t-1}$ varies from 0 to $B^{t-1} - 1$. Therefore, we conclude that for $t = 1, 2, \dots, K - 1$, any subgroup of a level t group starts at state m such that $m \bmod B^{t-1} = 0$ and consists of B^{t-1} consecutive states.

Since $b_0 \in \{0, 1, \dots, B - 1\}$, by definition every group consists of B subgroups according to different b_0 values. From (15), it can be seen that different b_0 values result in disjoint sets of $S_t \bmod B^t$ values as the B -tuples $(b_1, b_2, \dots, b_{t-1})$ vary from $(0, 0, \dots, 0)$ to $(B - 1, B - 1, \dots, B - 1)$. Thus, for $t = 1, 2, \dots, K - 1$, every level t group consists of B disjoint subgroups. We index the subgroups within a group by the zeroth source symbol, so S_t belongs to the b_0 th subgroup and S'_t belongs to the b'_0 th subgroup of the same level t group.

In the case of the ST code of Figure 1, the level 1 groups consist of 4 consecutive states, starting at states 0, 4, 8, and 12. The subgroups consist of only one state. The only level 2 group is comprised of all the 16 states, and its subgroups are made up of 4 consecutive states, starting at states 0, 4, 8, and 12. The level 1 and level 2 groups and subgroups of the example ST code are depicted in Figures 2b and 2c, respectively.

Because both the correct and the erroneous paths start from the same state ($S_0 = S'_0$), at state transition t , $t = 1, 2, \dots, K - 1$, both the correct path (S_t) and the erroneous

path (S'_t) go through states that belong to the same level t group. This means that if the m th indices of the channel symbol index vectors at states belonging to any level t group are the same, then the m th entry of the t th column of \mathbf{D}_1 will be zero. For example, in Figure 1, states 8 and 11 belong to the same level 1 group, and the second indices of the channel symbol index vectors $[0, 0, 1]^T$ and $[1, 1, 1]^T$ are the same. As a consequence, the first column of the \mathbf{D}_1 matrix becomes $[1 - j, 1 - j, 0]^T$.

Since the first decoding error occurs at S_0 ($b_0 \neq b'_0$), at state transition t , $t = 1, 2, \dots, K - 1$, the correct path (S_t) and the erroneous path (S'_t) go through states that belong to different subgroups of the same level t group. We can take advantage of this fact as follows: if the m th indices of the channel symbol index vectors at states belonging to different subgroups of the same level t group are different, then the m th entry of the t th column of \mathbf{D}_1 will be nonzero. To continue the example, states 0 and 13 belong to different subgroups of the same (only) level 2 group. The second indices of the channel symbol index vectors $[0, 0, 0]^T$ and $[0, 2, 2]^T$ are different, so the second column of the matrix \mathbf{D}_1 will be $[0, 2, 2]^T$.

Having produced the methods to place zero and nonzero entries into the matrix \mathbf{D}_1 , we can state the second half of the design rules:

- (2a) For $t = 1, 2, \dots, K - 1$, the t th indices of the channel symbol index vectors at states belonging to the same subgroup of any level t group must be the same, and they must be different from the t th indices of the channel symbol index vectors at states belonging to any other subgroup of that group.
- (2b) For $t = 1, 2, \dots, K - 2$, the $(t+1)$ st, $(t+2)$ nd, \dots , $(K-1)$ st indices of the channel symbol index vectors at states belonging to the same level t group must be the same. (Note that criterion (2b) is omitted for $t = K - 1$.)

After making the matrix \mathbf{D}_1 full rank, the final task is to show that the channel symbol difference matrix \mathbf{D} corresponding to the transmission of all T source symbols is also of full rank. The matrix \mathbf{D} can be decomposed as

$$\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2], \quad (17)$$

where \mathbf{D}_1 is defined in (14), and \mathbf{D}_2 is a $K \times (T - K)$ matrix. Since \mathbf{D}_2 is arbitrary, this description includes the cases when the correct and the decoded paths diverge and merge several times. From linear algebra, it is well known that if \mathbf{D}_1 is of full rank, then \mathbf{D} is also of full (row) rank. Consequently, the design rules will produce codes that provide full diversity advantage.

Figure 2a illustrates design rules (1a) and (1b) for the ST code shown in Figure 1. At each state, the zeroth indices of the channel symbol index vectors are different, and the 1st and second indices are the same. Figure 2b shows how the rules (2a) and (2b) are applied for $t = 1$. In each subgroup of the level 1 groups, the first indices of the channel symbol index vectors are the same, and they are different from the first indices of the index vectors of any other subgroup of that

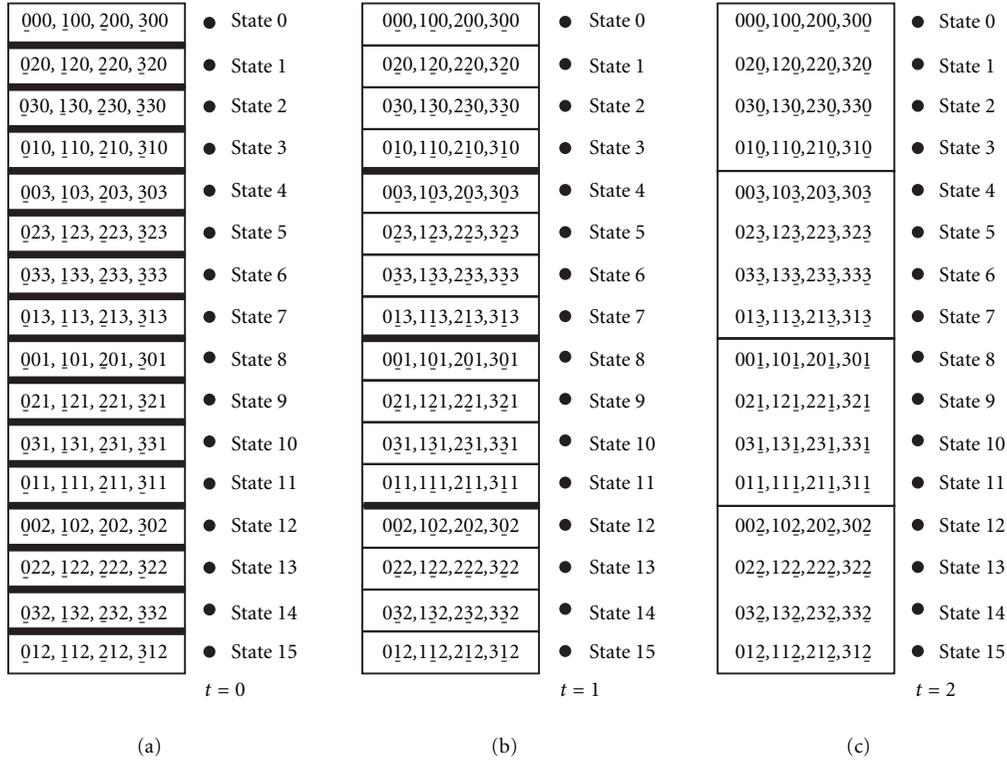


FIGURE 2: The group/subgroup structure of the example ST code.

group. Moreover, the second indices of the channel symbol index vectors in each level 1 group are the same. Finally, rule (2a) for $t = 2$ is illustrated in Figure 2c. In each subgroup of the only level 2 group, the second indices of the channel symbol index vectors are the same, and they are different from the second indices of the index vectors of the other subgroups.

By observing the group/subgroup structure of the state transitions in the trellis, we can make the design rules independent of the state evolution in time. The above design method describes relationships between channel symbol indices of different antennas at different states. Furthermore, these design rules do not fully determine the state-channel symbol assignment, providing the possibility to further optimize for coding gain.

Design rules (1a) and (1b) are similar to the design rules described in [9] for two transmit antennas. Therefore, our approach can be treated as a generalization of the method of [9] to an arbitrary number of transmit antennas.

4. DESIGN FOR CODING ADVANTAGE

In general, finding the best way to assign channel symbol indices to antennas and states is not a simple task. If $N > N_{\min}$, the shortest error path is longer than K , so the corresponding code difference matrix does not have any special structure. As a consequence, expressing the minimum determinant of the code becomes very difficult. However, in the $N = N_{\min} = B^{K-1}$ case (i.e., $r = 0$ and $p = 0$), it is possible to find an efficient method to maximize the coding gain, so

from now on, it is assumed that the encoder has N_{\min} states.

The channel symbol difference matrix corresponding to the first K long segment of the error paths after the first decoding error has occurred is the matrix \mathbf{D}_1 , defined in (14). It is square and upper triangular, so its determinant is the product of its diagonal elements:

$$\det(\mathbf{D}_1) = \prod_{k=0}^{K-1} \left(c(i^k(S_k, b_k)) - c(i^k(S'_k, b'_k)) \right). \quad (18)$$

We define the $K \times K$ matrix \mathbf{A}_1 as $\mathbf{A}_1 = \mathbf{D}_1 \mathbf{D}_1^H$. Then γ_1 , the determinant of \mathbf{A}_1 , is

$$\begin{aligned} \gamma_1 &= \det(\mathbf{A}_1) = \det(\mathbf{D}_1) \det(\mathbf{D}_1)^* \\ &= \prod_{k=0}^{K-1} \left| c(i^k(S_k, b_k)) - c(i^k(S'_k, b'_k)) \right|^2. \end{aligned} \quad (19)$$

Considering the transmission of all T source symbols, and using the decomposition of (17), the matrix $\mathbf{A} = \mathbf{D} \mathbf{D}^H$, whose minimum determinant is to be maximized, can be expressed as

$$\mathbf{A} = \mathbf{D}_1 \mathbf{D}_1^H + \mathbf{D}_2 \mathbf{D}_2^H = \mathbf{A}_1 + \mathbf{A}_2, \quad (20)$$

where $\mathbf{A}_2 = \mathbf{D}_2 \mathbf{D}_2^H$. By construction, both \mathbf{A}_1 and \mathbf{A}_2 are Hermitian and nonnegative definite. To continue the argument, we will use the following theorem from linear algebra [12]: Let \mathbf{X} and \mathbf{Y} be $K \times K$, Hermitian, and nonnegative

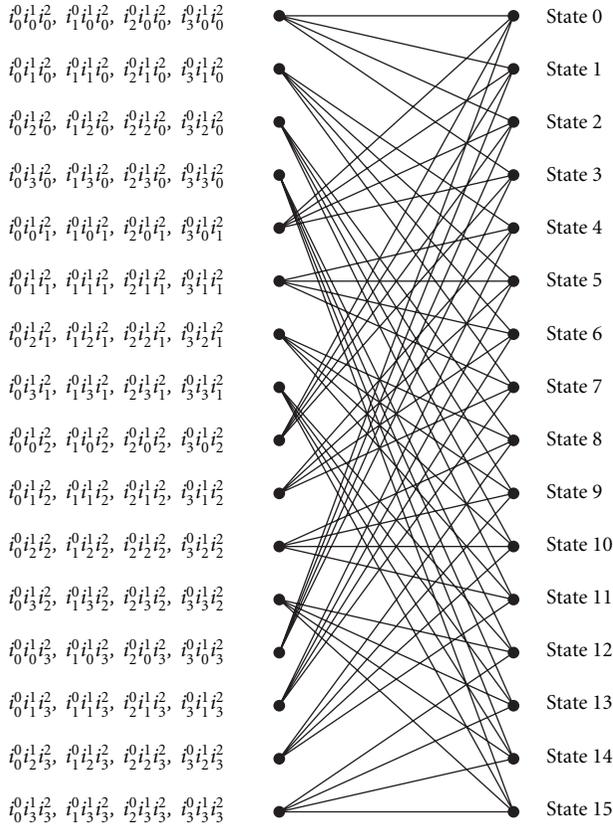


FIGURE 3: Example ST code template for three antennas, 4-ary constellations.

definite matrices. Moreover, let $\lambda_0(\mathbf{X}) \geq \lambda_1(\mathbf{X}) \geq \dots \geq \lambda_{K-1}(\mathbf{X})$ denote the real and nonnegative eigenvalues of \mathbf{X} . Then we have the following inequality for $i = 0, 1, \dots, K-1$:

$$\lambda_i(\mathbf{X} + \mathbf{Y}) \geq \lambda_i(\mathbf{X}) + \lambda_{K-1}(\mathbf{Y}). \quad (21)$$

In our case, (21) becomes

$$\lambda_i(\mathbf{A}_1 + \mathbf{A}_2) \geq \lambda_i(\mathbf{A}_1) + \lambda_{K-1}(\mathbf{A}_2) \quad \text{for } i = 0, 1, \dots, K-1. \quad (22)$$

Since \mathbf{A}_2 is nonnegative definite, $\lambda_{K-1}(\mathbf{A}_2) \geq 0$. This means that

$$\lambda_i(\mathbf{A}_1 + \mathbf{A}_2) \geq \lambda_i(\mathbf{A}_1) \quad \text{for } i = 0, 1, \dots, K-1. \quad (23)$$

From this, we can conclude that γ , the determinant of \mathbf{A} , satisfies the inequality

$$\gamma = \det(\mathbf{A}) = \prod_{i=0}^{K-1} \lambda_i(\mathbf{A}_1 + \mathbf{A}_2) \geq \prod_{i=0}^{K-1} \lambda_i(\mathbf{A}_1) = \det(\mathbf{A}_1) = \gamma_1. \quad (24)$$

We can fix an arbitrary correct path and pick an arbitrary error path that is longer than K state transitions. Both this error path and the error path corresponding to the K long error event that starts from the same S_0 starting state

and the same b'_0 starting branch go through states that belong to the same subgroups of the same groups, resulting in \mathbf{D}_1 matrices with the same diagonal elements (design rules (1a) and (2a)). Therefore, for any error event that is longer than K state transitions, it is possible to find a K long error event with the same $\det(\mathbf{A}_1)$ value. As a consequence of this observation and (24), γ_{\min} , the minimum determinant of the code, can be determined by taking into account only the shortest error events:

$$\gamma_{\min} = \min_{\{S_l, b_l\}, \{S'_l, b'_l\}} \prod_{k=0}^{K-1} |c(i^k(S_k, b_k)) - c(i^k(S'_k, b'_k))|^2. \quad (25)$$

The minimum is taken over all possible K long correct and incorrect paths.

The S_k and S'_k state transition sequences can also be described by making use of the group/subgroup structure of the trellis. The results of Section 3 allow us to map the first K long segment of the correct and erroneous paths of the first decoding error event onto different groups and subgroups of states. Toward this end, we introduce a channel symbol index based notation that does not explicitly depend on the state transition sequence.

Let $i_l^0, i_l^0 \in \{0, 1, \dots, B-1\}$, be the zeroth indices of the channel symbol index vectors at the same state corresponding to source symbol l ($l \in \{0, 1, \dots, B-1\}$). For simplicity, it is assumed that the 0th indices of the channel symbol index vectors at different states corresponding to the same source symbol values are the same. Moreover, let $i_l^k, k = 1, 2, \dots, K-1, i_l^k \in \{0, 1, \dots, B-1\}$, denote the k th indices of the channel symbol index vectors at the states belonging to the l th subgroup of the same level k group ($l \in \{0, 1, \dots, B-1\}$). According to design rules (1a) and (2a), the relation $i_l^k \neq i_m^k$ must hold for any $l \neq m$. Therefore, the B -tuple $(i_0^k, i_1^k, \dots, i_{B-1}^k), k = 0, 1, \dots, K-1$, is an (arbitrary) permutation of the numbers $0, 1, \dots, B-1$.

Applying the design method of Section 3 and using the above index notation, we can create a "template" ST code. It is called template because the design rules for full spatial diversity do not specify the codes completely. It contains channel symbol index templates at each state for each antenna. For the ST code example of Figure 1, this template is shown in Figure 3. Here, the 4-tuples $(i_0^0, i_1^0, i_2^0, i_3^0)$, $(i_0^1, i_1^1, i_2^1, i_3^1)$, and $(i_0^2, i_1^2, i_2^2, i_3^2)$ can be any permutations of the numbers $(0, 1, 2, 3)$. The ST code will achieve full diversity advantage for arbitrary permutations. The objective is to find those permutations that result in maximizing the coding advantage of the ST code.

Using the simplified notation $l = b_0$ and $m = b'_0$ ($l \neq m$), it was shown earlier that for $k = 1, 2, \dots, K-1$, S_k belongs to the l th subgroup and S'_k belongs to the m th subgroup of the same level k group. Therefore, using the above defined index notation, we can make the following substitutions:

$$i^k(S_k, b_k) = i_l^k, \quad i^k(S'_k, b'_k) = i_m^k, \quad \text{for } k = 0, 1, \dots, K-1. \quad (26)$$

Consequently, the expression for the minimum determinant

can be rewritten as

$$\gamma_{\min} = \min_{\substack{l, m \in \{0, 1, \dots, B-1\} \\ l < m}} \prod_{k=0}^{K-1} \left| c(i_l^k) - c(i_m^k) \right|^2. \quad (27)$$

The $l < m$ condition can be used since the squared distance function is symmetric in its arguments. The final goal is to maximize the minimum determinant. Therefore, if Ω_B denotes the set of all permutations of the numbers $0, 1, \dots, B-1$, and $\sigma^k \in \Omega_B$ ($k = 0, 1, \dots, K-1$) stands for a particular permutation $(i_0^k, i_1^k, \dots, i_{B-1}^k)$, then γ_{\min}^* , the optimal minimum determinant, can be expressed as

$$\gamma_{\min}^* = \max_{\sigma^0, \sigma^1, \dots, \sigma^{K-1}} \left[\min_{\substack{l, m \in \{0, 1, \dots, B-1\} \\ l < m}} \prod_{k=0}^{K-1} \left| c(i_l^k) - c(i_m^k) \right|^2 \right]. \quad (28)$$

This combinatorial optimization problem can be interpreted as follows. The design rules for diversity advantage and K permutations of the numbers $0, 1, \dots, B-1$ together uniquely determine the code. The task is to find those permutations that offer the largest minimum determinant. Because the numbers $0, 1, \dots, B-1$ can be arranged in $B!$ different ways, the size of the search space is $(B!)^K$. This means that exhaustive search may be impractical in certain cases. For example, if $B = 16$ and $K = 5$, the search has to be done over approximately $4 \cdot 10^{66}$ possibilities.

To get around this complexity growth, we propose a sub-optimal approach that offers a practical solution. The basic idea is to restrict the search space such that the resulting complexity is not prohibitive. Toward this end, we define the notion of the parametric permutation function.

The parametric permutation function is a function that generates a subset of all possible permutations of the numbers $0, 1, \dots, B-1$. Different parameters produce different permutations, so the problem will be reduced to a search for the best parameter. Therefore, the parametric permutation function is required to have the following properties:

- It must be a bijective map: for any parameter, it must map the set $\{0, 1, \dots, B-1\}$ onto itself in a one-to-one manner.
- Two different parameters must generate two different permutations.

In the sequel, we will use the notation $i_2 = \psi_B(n, i_1)$ for one possible realization of the parametric permutation function. The value n is the parameter to be optimized ($n \in \{1, 2, \dots, B-1\}$), and i_1 and i_2 are the input and output indices, respectively ($i_1, i_2 \in \{0, 1, \dots, B-1\}$). Note that the obvious $\psi_B(n, i_1) = (ni_1) \bmod B$, for odd n , is not a good choice because it does not “shuffle” the indices well enough.

The output index is generated according to the following description. The input index i_1 and the parameter n are treated as binary vector representations of two field elements in $\text{GF}(B) = \text{GF}(2^b)$. These field elements are multiplied together according to field arithmetic, and the output index i_2 will be the binary vector representation of the product. For example, for $B = 4$, the field $\text{GF}(4)$ can be built up using α , a

TABLE 1: Element representation in $\text{GF}(4)$.

Value	Binary vector representation	Polynomial representation	Power representation	Power
0	00	0	0	–
1	01	1	α^0	0
2	10	α	α^1	1
3	11	$\alpha + 1$	α^2	2

TABLE 2: The generated permutations.

i_1	$n = 1$	$n = 2$	$n = 3$
0	0	0	0
1	1	2	3
2	2	3	1
3	3	1	2

root of the primitive polynomial $p(x) = x^2 + x + 1$, as shown in Table 1. In this case, the function $\psi_4(\cdot, \cdot)$ will generate the permutations given in Table 2. The table entries are the function values for different input index and parameter values.

The above example shows two general properties of the function $\psi_B(\cdot, \cdot)$. First, the zero index always maps to itself: $\psi_B(n, 0) = 0$. Second, if the parameter is one, the function value is equal to the input index value: $\psi_B(1, i_1) = i_1$. Since the above definition of the parametric permutation function exploits the algebraic properties of the underlying Galois field, it can be easily seen that it has the required properties. Moreover, the elements of a Galois field can also be represented as powers of a primitive element, providing the possibility of efficient implementation by turning the multiplications into modulo $B-1$ additions.

Replacing the permutation operation by the parametric permutation function, the optimization problem reduces to

$$\gamma_{\min}^* = \max_{\substack{n_0, n_1, \dots, n_{K-1} \\ \in \{1, 2, \dots, B-1\}}} \left[\min_{\substack{l, m \in \{0, 1, \dots, B-1\} \\ l < m}} \prod_{k=0}^{K-1} \left| c(\psi_B(n_k, l)) - c(\psi_B(n_k, m)) \right|^2 \right]. \quad (29)$$

To find γ_{\min}^* , we only have to search over $(B-1)^K$ possibilities. In the case of the $B = 16, K = 5$ example, the size of the search space will be less than $8 \cdot 10^5$. Once the $n_0^*, n_1^*, \dots, n_{K-1}^*$ parameter values that maximize the minimum determinant have been calculated, the channel symbol indices can be determined as

$$i_l^k = \psi_B(n_k^*, l), \quad k = 0, 1, \dots, K-1, l = 0, 1, \dots, B-1. \quad (30)$$

5. CODE DESIGN EXAMPLES

This section provides some examples that demonstrate the code design method described in the previous sections. First,

we describe the relationship between the repetition coded delay diversity scheme of [10] and our approach. The delay diversity scheme was chosen as a basis for comparison because, to our knowledge, this is the only existing procedure that can be used to construct ST trellis codes for any number of transmit antennas and any constellation. Then, we give specific code design examples.

The delay diversity scheme is a special case of our design method. For $B = 2$ (e.g., the BPSK constellation) and $N = N_{\min}$, the two methods are equivalent: the design rules for full spatial diversity uniquely determine the code. For $B > 2$, the delay diversity scheme corresponds to $n_k^* = 1$ ($k = 0, 1, \dots, K-1$), which leads to $i_l^k = l$ ($k = 0, 1, \dots, K-1$, $l = 0, 1, \dots, B-1$). In [10], it was shown that, if Δ denotes the minimum Euclidean distance of the chosen constellation, then the minimum determinant of the resulting delay diversity ST code will be $\gamma_{\min}^D = \Delta^{2K}$.

To characterize the theoretical performance improvement of our method over the delay diversity ST codes, we will use the *relative coding advantage*, β , defined as $\beta = \sqrt[K]{\gamma_{\min}^*/\gamma_{\min}^D}$. This quantity describes the normalized vertical shift between the two error performance curves as the SNR becomes large.

5.1. Code design for B-ary PSK

In case of B-ary PSK modulation, the squared distance between constellation points l and m ($l, m \in \{0, 1, \dots, B-1\}$) is given by

$$d^2(l, m) = |c(l) - c(m)|^2 = 4 \sin^2 \left(\frac{(l-m)\pi}{B} \right). \quad (31)$$

A pictorial representation of the QPSK and 8PSK constellations are shown in Figures 4a and 4b. Using (31) to express the minimum determinant of the code, the optimization problem becomes

$$\gamma_{\min}^* = \max_{\substack{n_0, n_1, \dots, n_{K-1} \\ \in \{1, 2, \dots, B-1\}}} \left[\min_{\substack{l, m \in \{0, 1, \dots, B-1\} \\ l < m}} 4^K \times \prod_{k=0}^{K-1} \sin^2 \left(\frac{(\psi_B(n_k, l) - \psi_B(n_k, m))\pi}{B} \right) \right]. \quad (32)$$

If this optimization procedure is used for three antennas ($K = 3$) with QPSK ($B = 4$, $N = 16$), and Table 2 is used to generate the parametric permutation function values, then the result of the optimization will be $n_0^* = 1$, $n_1^* = 2$, and $n_2^* = 3$ with $\gamma_{\min}^* = 16$. Note that this maximum is not unique: several other sets of $\{n_k^*\}$ values exist. This is not surprising because of the symmetry of the QPSK constellation and the commutativity of multiplication. The obtained permutations are: $(i_0^0, i_1^0, i_2^0, i_3^0) = (0, 1, 2, 3)$, $(i_0^1, i_1^1, i_2^1, i_3^1) = (0, 2, 3, 1)$, and $(i_0^2, i_1^2, i_2^2, i_3^2) = (0, 3, 1, 2)$. These permutations generate the ST code example depicted in Figure 1. The minimum determinant of the corresponding delay diversity ST code is $\gamma_{\min}^D = \Delta^6 = 8$, resulting in a relative coding advantage of

$\beta = 1.26$. For this particular case, the authors of [13] found a better code ($\gamma_{\min} = 32$) through computer search. However, exhaustive search cannot be used to find good ST codes for a larger number of transmit antennas and larger constellation sizes because of its computational complexity.

For easy description, we need to find an efficient and concise representation for our ST codes. In [14], a generator matrix based approach was used. However, some of our codes belong to a more general class of ST codes because they cannot be described by generator matrices. To see this, consider the ST code example of Figure 1. The constellation point for antenna 2 is determined by the two most significant bits of the 4 bit state information. We denote these two bits by s_2 and s_3 ($s_2, s_3 \in \{0, 1\}$). In order to be able to describe this ST code by a generator matrix, the function $F(s_2, s_3)$, defined as

$$F(s_2, s_3) = (a_2 s_2 + a_3 s_3) \bmod 4, \quad (33)$$

should generate the indices 0, 3, 1, 2 in this order for some a_2 and a_3 ($a_2, a_3 \in \{0, 1, 2, 3\}$). We trivially have $F(0, 0) = 0$. The $F(1, 0) = 3$ relation forces a_2 to be 3. Similarly, a_3 has to be 1, as a consequence of $F(0, 1) = 1$. Finally, the function will result in $F(1, 1) = 0$, which is not the desired value 2. Therefore, this ST code cannot be put in a generator matrix form. The possibly large number of encoder states prevents us from using the trellis diagram, so the channel symbol index permutations will be used to describe the ST codes. Note that this representation is unique, and due to the regular structure of the proposed ST codes, it is easy to design simple encoders with $O(KB)$ hardware complexity.

Using the same procedure, ST codes for three transmit antennas ($K = 3$), 8PSK ($B = 8$, $N = 64$), and 16PSK ($B = 16$, $N = 256$) constellations have also been constructed. In the 8PSK case, the channel symbol index permutations

$$\begin{aligned} (i_0^0, i_1^0, i_2^0, i_3^0, i_4^0, i_5^0, i_6^0, i_7^0) &= (0, 1, 2, 3, 4, 5, 6, 7), \\ (i_0^1, i_1^1, i_2^1, i_3^1, i_4^1, i_5^1, i_6^1, i_7^1) &= (0, 1, 2, 3, 4, 5, 6, 7), \\ (i_0^2, i_1^2, i_2^2, i_3^2, i_4^2, i_5^2, i_6^2, i_7^2) &= (0, 2, 4, 6, 3, 1, 7, 5) \end{aligned} \quad (34)$$

are one of the possible sets of permutations that maximize the objective function, yielding the minimum determinant $\gamma_{\min}^* = 0.6863$. The minimum determinant of the delay diversity scheme with the same design parameters is $\gamma_{\min}^D = 0.2010$. These coding advantage values yield $\beta = 1.51$. For the 16PSK constellation, our design method resulted in the ST code given by the channel symbol indices

$$\begin{aligned} (i_0^0, i_1^0, i_2^0, i_3^0, i_4^0, i_5^0, i_6^0, i_7^0, i_8^0, i_9^0, i_{10}^0, i_{11}^0, i_{12}^0, i_{13}^0, i_{14}^0, i_{15}^0) \\ &= (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15), \\ (i_0^1, i_1^1, i_2^1, i_3^1, i_4^1, i_5^1, i_6^1, i_7^1, i_8^1, i_9^1, i_{10}^1, i_{11}^1, i_{12}^1, i_{13}^1, i_{14}^1, i_{15}^1) \\ &= (0, 2, 4, 6, 8, 10, 12, 14, 3, 1, 7, 5, 11, 9, 15, 13), \\ (i_0^2, i_1^2, i_2^2, i_3^2, i_4^2, i_5^2, i_6^2, i_7^2, i_8^2, i_9^2, i_{10}^2, i_{11}^2, i_{12}^2, i_{13}^2, i_{14}^2, i_{15}^2) \\ &= (0, 4, 8, 12, 3, 7, 11, 15, 6, 2, 14, 10, 5, 1, 13, 9). \end{aligned} \quad (35)$$

TABLE 3: ST codes for QPSK modulation.

K	N	Permutations	γ_{\min}^*	γ_{\min}^D	Decoding depth
2	4	P_1, P_1	4	4	25
4	64	P_1, P_1, P_2, P_3	32	16	25
6	1024	$P_1, P_1, P_2, P_2, P_3, P_3$	256	64	25
8	2^{14}	$P_1, P_1, P_1, P_1, P_2, P_2, P_3, P_3$	1024	256	35
10	2^{18}	$P_1, P_1, P_1, P_1, P_2, P_2, P_2, P_3, P_3, P_3$	8192	1024	45

The minimum determinant of this code is $\gamma_{\min}^* = 0.110105$, while the delay diversity construction gives $\gamma_{\min}^D = 0.003529$. The relative coding advantage is $\beta = 3.15$.

The reduced computational complexity of the proposed design procedure allowed us to construct ST codes for a large number of transmit antennas. Table 3 contains the brief description of the codes designed for QPSK constellation. The symbols P_1 , P_2 , and P_3 denote the permutations $(0, 1, 2, 3)$, $(0, 2, 3, 1)$, and $(0, 3, 1, 2)$, respectively. The permutations are assigned to transmit antennas from left to right. According to this notation, the three antenna ST code example of Figure 1 can be described as: P_1, P_2, P_3 . The γ_{\min}^* and γ_{\min}^D values are also shown.

5.2. Code design for asymmetric QPSK

The next two examples employ constellations that are not used in current wireless communication systems, but they can illustrate the flexibility of the proposed design method. The first ST code was constructed for four transmit antennas ($K = 4$) and asymmetric QPSK modulation ($B = 4, N = 64$). The pictorial representation of the asymmetric QPSK constellation can be observed in Figure 4c. The parameter α , which is the angle between the signal points and the real axis, was set to $\pi/8$ (rad). The Euclidean distances between two arbitrary constellation points cannot be expressed in a closed form, but they can be easily calculated. The minimum distance of the constellation is $\Delta = 2 \sin \alpha = 0.7654$. The optimization procedure results in

$$\begin{aligned}
 (i_0^0, i_1^0, i_2^0, i_3^0) &= (0, 1, 2, 3), \\
 (i_0^1, i_1^1, i_2^1, i_3^1) &= (0, 1, 2, 3), \\
 (i_0^2, i_1^2, i_2^2, i_3^2) &= (0, 2, 3, 1), \\
 (i_0^3, i_1^3, i_2^3, i_3^3) &= (0, 3, 1, 2)
 \end{aligned} \tag{36}$$

permutations with $\gamma_{\min}^* = 4.6863$. The minimum determinant of the delay diversity scheme is $\gamma_{\min}^D = 0.1177$, so a relative coding advantage of $\beta = 2.51$ is achieved.

5.3. Code design for 4ASK

We also designed a ST code for four transmit antennas ($K = 4$), and 4ASK constellation ($B = 4, N = 64$), shown in Figure 4d. The minimum distance of the normalized constellation was $\Delta = \sqrt{4/5}$. The squared distance between constel-

lation points l and m ($l, m \in \{0, 1, 2, 3\}$) can be expressed as

$$d^2(l, m) = |c(l) - c(m)|^2 = \Delta^2(l - m)^2. \tag{37}$$

The ST code design method found the same permutations as for the asymmetric QPSK case. The minimum determinant of the code is $\gamma_{\min}^* = 1.6384$, and the delay diversity method yields $\gamma_{\min}^D = 0.4096$. The resulting relative coding advantage is $\beta = 1.41$.

5.4. Discussion

The definition of the relative coding advantage allows us to predict the performance improvement before performing any simulation. If we compare the β values in the three antenna case for B -ary PSK, we can see that as the number of constellation points (B) increases, the relative coding advantage also increases, and, therefore, more significant improvement is expected.

Based on the relative coding advantage values of the four antenna ST codes for QPSK, 4ASK and asymmetric QPSK modulations, improvement comparison can be made for a fixed constellation size ($B = 4$). The β values suggest that the proposed QPSK ST code will perform a little better than the delay diversity scheme, and the improvement will be more pronounced in the case of the 4ASK codes. Finally, the asymmetric QPSK ST code (whose actual performance depends on the value of α) seems to offer the largest improvement.

Due to the structure of the proposed ST codes, the minimum determinants are functions of product distances. The code design method tries to assign channel symbol indices to antennas at different states in such a way that the minimum value of the product distances is as large as possible. The minimum determinant of the delay diversity construction is only a function of the minimum distance of the constellation. Therefore, if the maximum distance of the chosen constellation is much larger than the minimum distance, our design method can exploit the additional degrees of freedom effectively, producing ST codes that perform much better than the delay diversity scheme. On the other hand, if the distances in the constellation have similar magnitudes, the proposed design method may not result in significant improvement.

6. SIMULATION RESULTS

To illustrate the performance of the codes designed using the above described method, we show some simulation results. We compare our approach with the delay diversity scheme of [10], since, to our knowledge, this is the only method that can be used to construct ST codes for any number of transmit antennas and any memoryless constellation. The simulated communication system had one receive antenna. The source symbols were transmitted in frames of length 130, and the Viterbi algorithm (ML sequence detection) with decoding depth of 20 state transitions was used to decode the received signals. For each frame, the path gains between the transmit antennas and the receive antenna were modeled as independent, complex, zero-mean, circularly symmetric Gaussian random variables with unit variance.

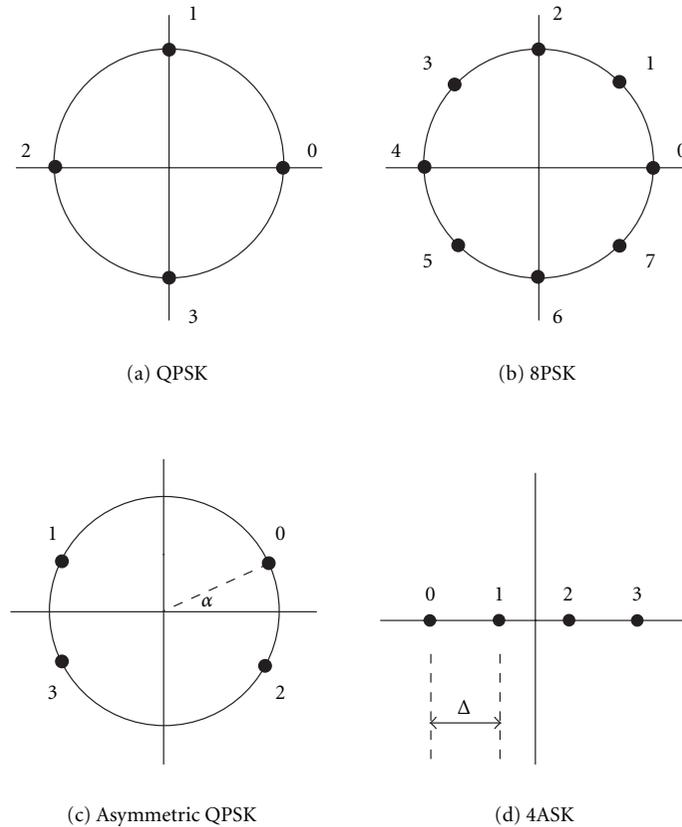


FIGURE 4: Example constellations.

Since the frame error probability depends on the length of the frame and it does not seem very informative, we present probability of bit error curves as functions of the average signal-to-noise ratio (SNR) per source symbol at the receive antenna. In the sequel, the expression *coding gain* will refer to the difference (in dB) of transmit energies to achieve the same probability of bit error value. Both the coding advantage and the coding gain give information about the performance improvement, but the coding advantage is a theoretical quantity characterizing the *vertical* shift of the error performance curve, while the coding gain is experimental and it describes the *horizontal* shift.

The repetition coded delay diversity of [10] is a special case of our design rules. Figure 5a shows the performance of this scheme for different number of transmit antennas ($K = 2, 3, 4$ and $N = 4, 16, 64$, respectively) and QPSK modulation. We observe that the codes indeed provide different spatial diversity advantages since the steepness of the bit error rate curves is different.

The rest of the figures compare the performance of the delay diversity construction and our approach using the example codes described in the previous section. Figure 5b depicts the results for three transmit antennas and QPSK modulation. The two probability of bit error curves are shifted versions of each other, as expected. Approximately 0.4–0.5 dB coding gain is observed over the delay diversity scheme.

Figure 6a shows the bit error rate curves of a three transmit antenna system with 8PSK modulation. At low SNR, the two error performance curves are close to each other, and they behave according to the expectations at medium and high SNRs. This phenomenon can be explained as follows. The definition of coding advantage [9] is based on an upper bound on the $Q(x)$ Gaussian tail probability function, and this bound is loose at low SNR. Moreover, the large number of transmission errors and the small minimum distance of the constellation may prevent the Viterbi algorithm with finite decoding depth from working properly at low SNR. The simulation shows that the performance improvement is more pronounced; at higher SNR, more than 1 dB coding gain can be achieved.

The performance of the ST code for three antennas and 16PSK constellation can be observed in Figure 6b. Our ST code yields 2–2.5 dB coding gain compared to the delay diversity scheme. Figures 7a and 7b depict the bit error rate curves for the four antenna ST codes using asymmetric QPSK and 4ASK modulation, respectively. The first figure shows approximately 3 dB coding gain from medium SNR, and the second figure demonstrates 2 dB improvement over the delay diversity construction.

Since the number of states is exponential in the number of transmit antennas, it is not possible to decode the ST trellis codes designed for a large number of transmit antennas

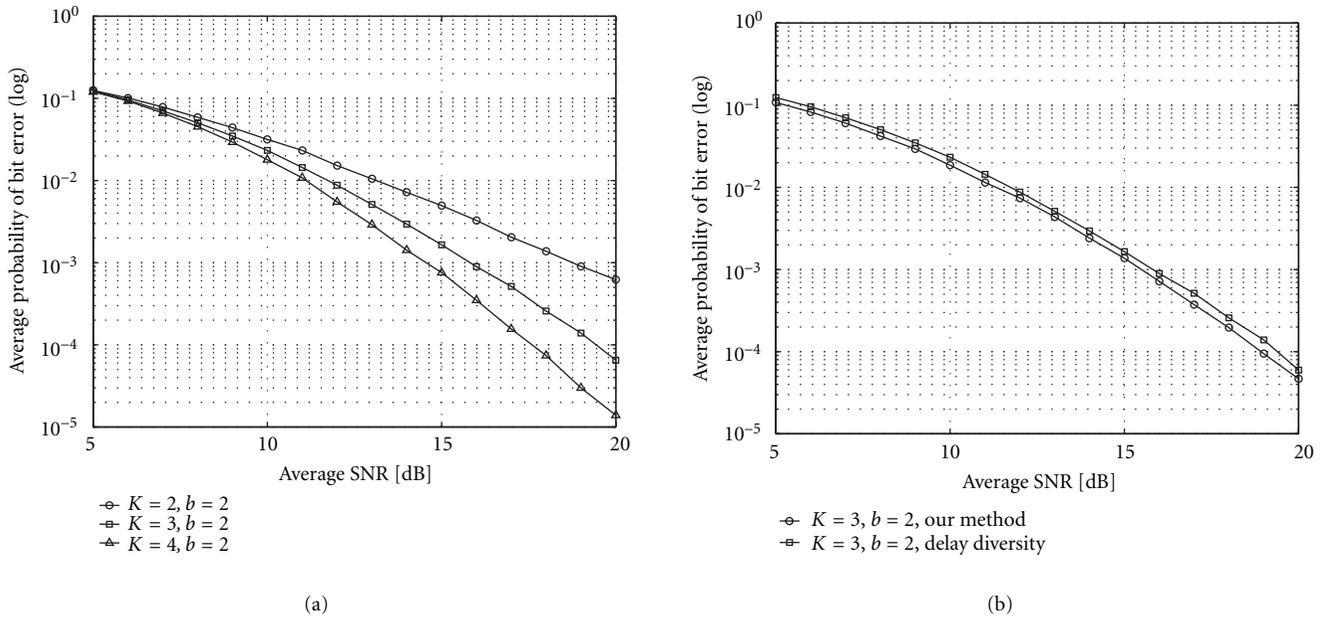


FIGURE 5: Simulation results I. (a) Delay diversity with QPSK, (b) three transmit antennas with QPSK.

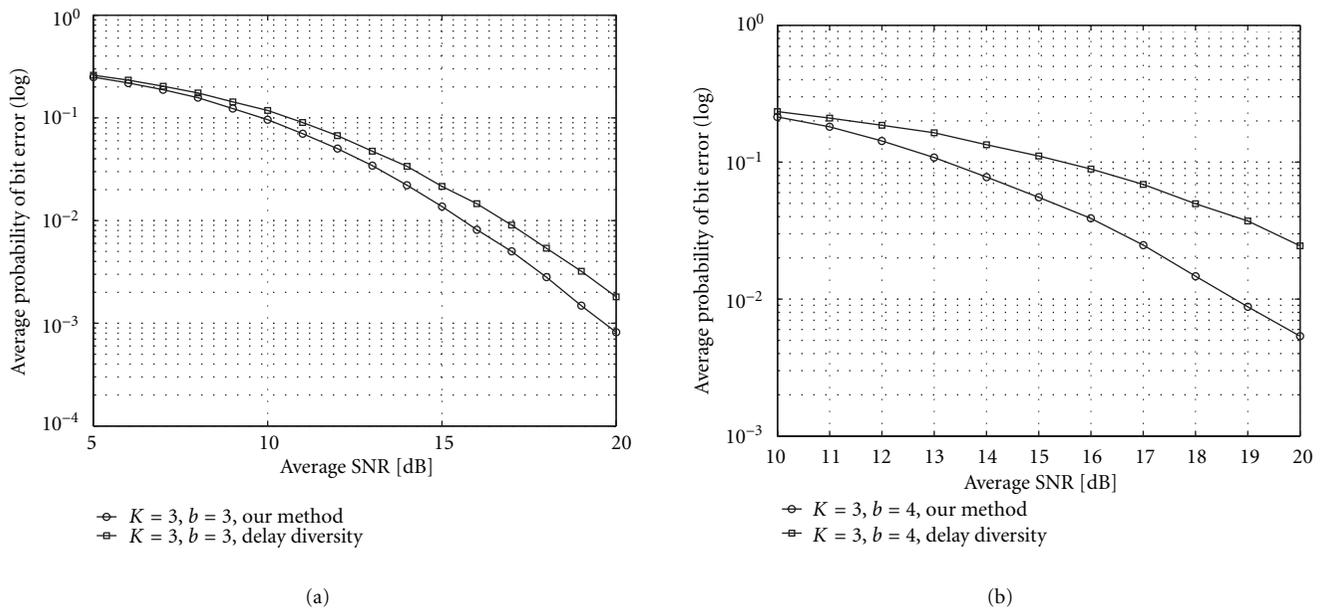


FIGURE 6: Simulation results II. (a) Three transmit antennas with 8PSK, (b) three transmit antennas with 16PSK.

using ML sequence detection (Viterbi algorithm). Therefore, we chose a suboptimal tree decoding algorithm developed for convolutional coding and trellis coding: the M -algorithm [15]. This algorithm uses a tree structure to evaluate the metrics (in our case: the Euclidean distances) for the allowable channel symbol sequences. At each stage, it keeps at most M partial paths with the best metrics. Thus, the decoding complexity is $O(M)$, which is independent of the number of encoder states.

Figure 8a shows the performance of the four antennas, QPSK ST code for ML decoding and for suboptimal decoding. The bit error curve of the three antennas, QPSK ST code with ML decoding is also included for comparison. It can be observed that reducing the computational complexity (reducing the value of M) results in performance degradation.

Finally, the simulation results for the ST codes described in Table 3 are depicted in Figure 8b. For 2 and 4 antennas,

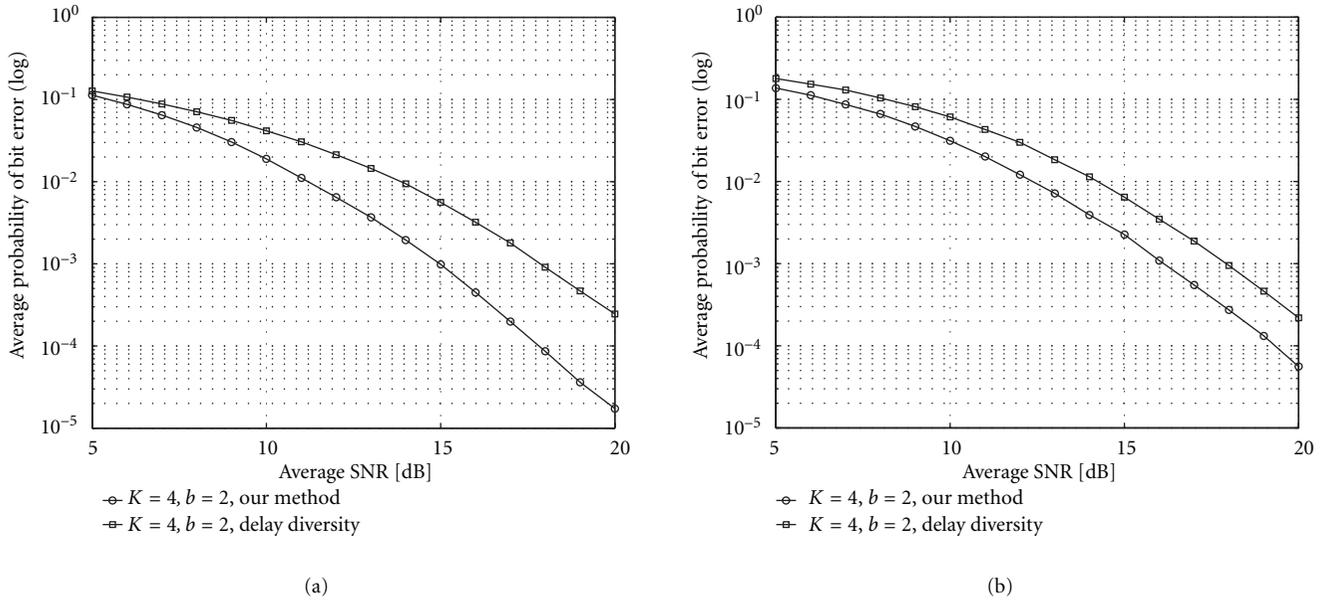


FIGURE 7: Simulation results III. (a) Four transmit antennas with asymmetric QPSK, (b) four transmit antennas with 4ASK.

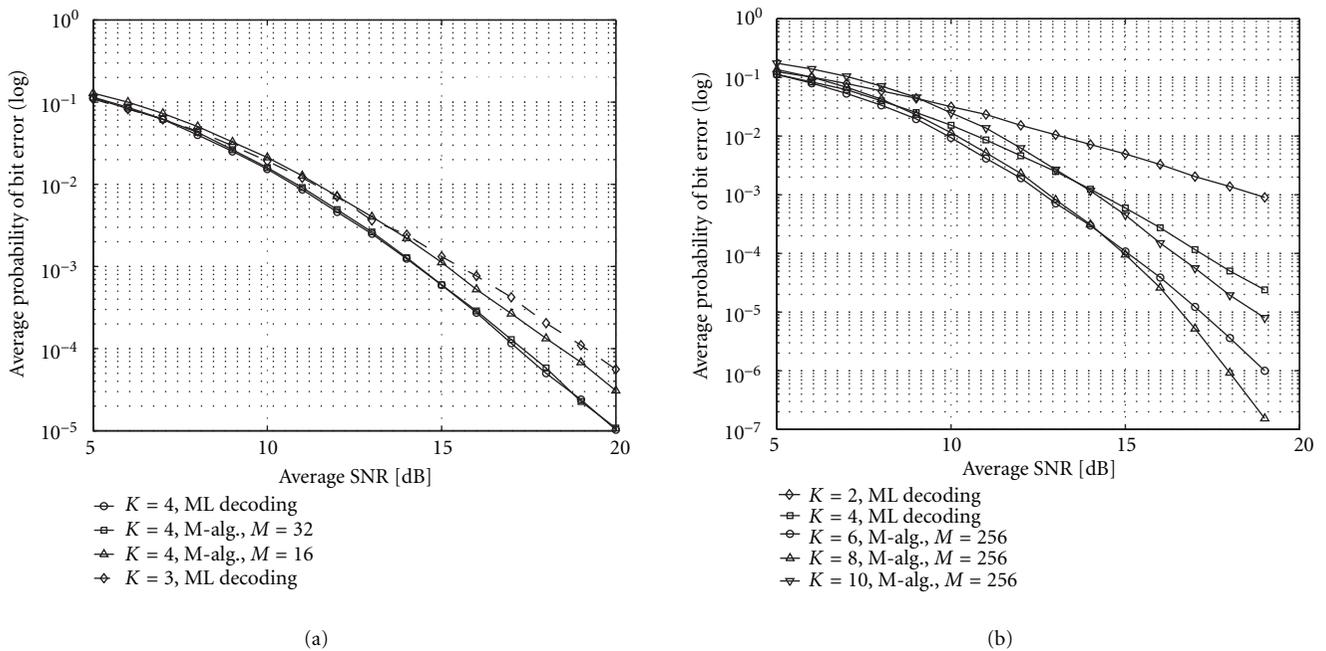


FIGURE 8: Simulation results IV. (a) Four antennas, suboptimal decoding, (b) ST codes for QPSK modulation.

the ML decoding algorithm was used, and for the 6, 8, and 10 antenna cases, the ST codes were decoded by the suboptimal M-algorithm. The decoding complexity was kept approximately constant by setting $M = 256$. The used decoding depth values can be found in the last column of Table 3. The results show that as the N/M ratio increases, the performance loss increases.

7. CONCLUSION

Having observed the group/subgroup structure of the state transitions, we proposed systematic design rules for ST trellis codes that achieve full spatial diversity. For encoders having N_{\min} states, we developed a code construction method that allows for ST code design for both diversity advantage and

coding advantage. Due to the low complexity of the proposed design method, ST codes for a large number of antennas were also constructed. Based on the theoretical coding advantage values and the simulation results, we can draw the following conclusions.

First, if the ratio of the maximum and the minimum distances of the chosen constellation is large, the additional optimization for coding advantage results in codes that substantially outperform the codes that were designed only for diversity advantage.

Second, if the M-algorithm is used to decode the ST codes with constant decoding complexity in a quasi-static Rayleigh fading environment, increasing the number of transmit antennas will provide diminishing returns. The choice of the number of transmit antennas will depend on the actual allowed maximum computational complexity.

APPENDIX

We will prove (12) by induction. Using (11), the first state change can be expressed easily

$$S_1 = B(S_0 \bmod (RB^{K+p-2})) + b_0. \quad (\text{A.1})$$

Assume that the formula holds for S_{t-1} , formally

$$S_{t-1} = B^{t-1}(S_0 \bmod (RB^{K+p-t})) + \sum_{m=0}^{t-2} B^{t-2-m} b_m. \quad (\text{A.2})$$

It only remains to show that the above described relationship also holds for S_t . Using the symbol Q to denote RB^{K+p-2} , S_t can be expressed as

$$\begin{aligned} S_t &= B(S_{t-1} \bmod Q) + b_{t-1} \\ &= B \left(\left[B^{t-1}(S_0 \bmod (RB^{K+p-t})) \right. \right. \\ &\quad \left. \left. + \sum_{m=0}^{t-2} B^{t-2-m} b_m \right] \bmod Q \right) + b_{t-1}. \end{aligned} \quad (\text{A.3})$$

By applying the identity

$$\left(\sum_{i=1}^n a_i \right) \bmod b = \left(\sum_{i=1}^n a_i \bmod b \right) \bmod b, \quad (\text{A.4})$$

the above expression becomes

$$\begin{aligned} S_t &= B \left(\left[\left(B^{t-1}(S_0 \bmod (RB^{K+p-t})) \right) \bmod Q \right. \right. \\ &\quad \left. \left. + \sum_{m=0}^{t-2} \left(B^{t-2-m} b_m \right) \bmod Q \right] \bmod Q \right) + b_{t-1}. \end{aligned} \quad (\text{A.5})$$

Recognizing that for any $0 \leq n \leq K+p-2$ and $S \in \{0, 1, \dots, N-1\}$,

$$(B^n S) \bmod (RB^{K+p-2}) = B^n (S \bmod (RB^{K+p-2-n})), \quad (\text{A.6})$$

the state transition equation can be rewritten as

$$\begin{aligned} S_t &= B \left(\left[B^{t-1} \left((S_0 \bmod (RB^{K+p-t})) \bmod (RB^{K+p-t-1}) \right) \right. \right. \\ &\quad \left. \left. + \sum_{m=0}^{t-2} \left(B^{t-2-m} b_m \right) \bmod Q \right] \bmod Q \right) + b_{t-1}. \end{aligned} \quad (\text{A.7})$$

The next step is to make use of the following simple result: if $b \in \{0, 1, \dots, B-1\}$ then for any $0 \leq n < K+p-2$

$$(B^n b) \bmod (RB^{K+p-2}) = B^n b. \quad (\text{A.8})$$

This allows for further simplification:

$$\begin{aligned} S_t &= B \left(\left[B^{t-1} \left((S_0 \bmod (RB^{K+p-t})) \bmod (RB^{K+p-t-1}) \right) \right. \right. \\ &\quad \left. \left. + \sum_{m=0}^{t-2} B^{t-2-m} b_m \right] \bmod Q \right) + b_{t-1}. \end{aligned} \quad (\text{A.9})$$

Now we use the fact that if $S \in \{0, 1, \dots, N-1\}$ and $0 \leq n \leq K+p-1$,

$$\begin{aligned} (S \bmod (RB^{K+p-n})) \bmod (RB^{K+p-n-1}) \\ = S \bmod (RB^{K+p-n-1}), \end{aligned} \quad (\text{A.10})$$

and we can obtain the following form:

$$\begin{aligned} S_t &= B \left(\left[B^{t-1} (S_0 \bmod (RB^{K+p-t-1})) \right. \right. \\ &\quad \left. \left. + \sum_{m=0}^{t-2} B^{t-2-m} b_m \right] \bmod Q \right) + b_{t-1}. \end{aligned} \quad (\text{A.11})$$

Finally, the identity

$$\begin{aligned} (B^n x + B^{n-1} b_0 + B^{n-2} b_1 + \dots + b_{n-1}) \bmod (RB^{K+p-2}) \\ = B^n x + B^{n-1} b_0 + B^{n-2} b_1 + \dots + b_{n-1} \end{aligned} \quad (\text{A.12})$$

that holds for $b_i \in \{0, 1, \dots, B-1\}$, $0 < n \leq K+p-2$, and $x \in \{0, 1, \dots, RB^{K+p-2-n}-1\}$ will give us the final step that completes the proof:

$$\begin{aligned} S_t &= B \left(B^{t-1} (S_0 \bmod (RB^{K+p-t-1})) + \sum_{m=0}^{t-2} B^{t-2-m} b_m \right) + b_{t-1} \\ &= B^t (S_0 \bmod (RB^{K+p-t-1})) + \sum_{m=0}^{t-1} B^{t-1-m} b_m. \end{aligned} \quad (\text{A.13})$$

The closed form expression for S_{K+p} can be derived using similar ideas.

REFERENCES

- [1] S. Saunders, *Antennas and Propagation for Wireless Communication Systems*, John Wiley & Sons, New York, NY, USA, 1999.
- [2] S. Wicker, *Error Control Systems for Digital Communication and Storage*, Prentice Hall, Englewood Cliffs, NJ, USA, 1995.
- [3] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 5, pp. 871–878, 1987.
- [4] J. Winters, J. Salz, and R. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Communications*, vol. 42, no. 2-4, pp. 1740–1750, 1994.
- [5] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 139–157, 1999.
- [6] I. Telatar, "Capacity of multi-antenna Gaussian channels," Technical report, AT&T Bell Labs., 1995.
- [7] G. Foschini and M. Gans, "On the limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Communication*, vol. 6, no. 3, pp. 311–335, 1998.
- [8] J.-C. Guey, M. Fitz, M. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE Vehicular Technology Conference*, vol. 1, pp. 136–140, Atlanta, Ga, USA, 1996.
- [9] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [10] J. Grimm, M. Fitz, and J. Krogmeier, "Further results on space-time coding for Rayleigh fading," in *Proc. 36th Allerton Conference on Communications, Control, and Computing*, pp. 391–400, Monticello, Ill, USA, September 1998.
- [11] A. R. Hammons, Jr. and H. El Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 524–542, 2000.
- [12] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, UK, 1985.
- [13] Q. Yan and R. Blum, "Optimum space-time convolutional codes," *Proc. IEEE Wireless Communications and Networking Conference*, vol. 3, pp. 1351–1355, 2000.
- [14] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis coded modulation," *IEEE Communications Letters*, vol. 4, no. 1, pp. 20–22, 2000.
- [15] C. Schlegel, *Trellis Coding*, IEEE Press, 1997.

K. J. Ray Liu received his B.S. degree from the National Taiwan University, and the Ph.D. degree from UCLA, both in electrical engineering. He is Professor of Electrical and Computer Engineering Department of the University of Maryland, College Park. His research interests span broad aspects of signal processing architectures; multimedia signal processing; wireless communications and networking; information security; and bioinformatics in which he has published over 230 refereed papers, of which over 70 are in archival journals. Dr. Liu is the recipient of numerous awards including the 1994 National Science Foundation Young Investigator, the IEEE Signal Processing Society's 1993 Senior Award, IEEE 50th Vehicular Technology Conference Best Paper Award, Amsterdam, 1999. He also received the George Corcoran Award in 1994 for outstanding contributions to electrical engineering education and the Outstanding Systems Engineering Faculty Award in 1996 in recognition of outstanding contributions in interdisciplinary research, both from the University of Maryland. Dr. Liu is Editor-in-Chief of EURASIP Journal on Applied Signal Processing, and has been an Associate Editor of IEEE Transactions on Signal Processing, a Guest Editor of special issues on Multimedia Signal Processing of Proceedings of the IEEE, a Guest Editor of special issue on Signal Processing for Wireless Communications of IEEE Journal of Selected Areas in Communications, a Guest Editor of special issue on Multimedia Communications over Networks of IEEE Signal Processing Magazine, a Guest Editor of special issue on Multimedia over IP of IEEE Trans. on Multimedia, and an Editor of Journal of VLSI Signal Processing Systems. Dr. Liu is the Past Chair of Multimedia Signal Processing Technical Committee.



Zoltan Safar received the University Diploma in electrical engineering from the Technical University of Budapest, Hungary in 1996. After one year of working experience, he started his graduate studies at the University of Maryland, where he is currently pursuing his Ph.D. in electrical engineering. His research interests include wireless communications and multimedia signal processing, with a particular focus on MIMO systems and space-time coding.

