REFERENCES

[1] S. C. Swales, M. A. Beach, D. J. Edwards, and J. P. McGeehan, "The performance enhancement of multibeam adaptive base-stations antennas for cellular land mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 39, no. 1, Feb. 1990.

[2] S. Anderson, M. Millnert, M. Viberg, and B. Wahlberg, "An adaptive array for mobile communications system," *IEEE Trans. Veh. Technol.*, vol. 40, Feb. 1991.

[3] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 49–83, Nov. 1997.

[4] P. Balaban and J. Saltz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio. I. Theoretical considerations," *IEEE Trans. Commun.*, vol. 40, pp. 885–894, May 1992.

[5] G. E. Bottomley and K. Jamal, "Adaptive arrays and MLSE equalization," in *Proc. Vehicular Technology Conf. (VTC'95)*, vol. 1, 1995, pp. 50–54.

[6] E. Lindskog, "Space-time processing and equalization for wireless communications," Ph.D. dissertation, Uppsala Univ., Uppsala, Sweden, June 1999.

[7] L. L. Scharf, "The SVD and reduced-rank signal processing," in *Signal Processing (EURASIP)*, vol. 25, 1991, pp. 113–133.

[8] J. W. Liang, J. T. Chen, and A. J. Paulraj, "A two-stage hybrid approach for CCI/ISI reduction with space-time processing," *IEEE Commun. Lett.*, vol. 1, pp. 163–165, Nov. 1997.

[9] U. Girola and S. Parolari, "Riduzione delle interferenze nei sistemi GSM/DCS mediante elaborazione spazio-temporale," M.Sc. thesis (in Italian), Politecnico di Milano, Milan, Italy, 1996–1997.

[10] M. A. Lagunas, A. I. Perez-Neira, and J. Vidal, "Joint beamforming and Viterbi equalizer in wireless communications," in *Proc. 31st Asilomar Conf. Signals, Systems, and Computing*, vol. 1, Nov. 1997, pp. 915–919.

[11] F. Pipon, C. Pascal, P. Vila, and J. J. Monot, "Joint spatial and temporal equalization for channels with ISI and CCI—Theoretical and experimental results for a base station reception," *Proc. 1st IEEE Signal Processing Advances in Wireless Communications Workshop*, pp. 309–312, Apr. 1997.

[12] P. Stoica and M. Viberg, "Maximum likelihood parameter and rank estimation in reduced-rank multivariate linear regressions," *IEEE Trans. Signal Processing*, vol. 44, pp. 3069–3078, Dec. 1996.

[13] G. Castellini, F. Conti, E. Del Re, and L. Pierucci, "A continuously adaptive MLSE receiver for mobile communications: Algorithm and performance," *IEEE Trans. Commun.*, vol. 45, pp. 80–89, Jan. 1997.

[14] D. Giancola, F. Margherita, S. Parolari, A. Picciriello, and U. Spagnolini, "Analysis of the spectral efficiency of a fully-adaptive antenna array system in GSM/DCS networks," in *Proc. Vehicular Technology Conf (VTC'99)*, vol. 1, May 1999, pp. 812–815.

[15] G. L. Stuber, *Principles of Mobile Communication*. Norwell, MA: Kluwer, 1996.

[16] G. K. Kaleh, "Simple coherent receivers for partial response continuous phase modulation," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1427–1436, Dec. 1989.

[17] P. Jung, "Laurent's representation of binary digital continuous phase modulated signals with modulation index $1/2$ revisited," *IEEE Trans. Commun.*, vol. 42, pp. 221–224, Feb./Mar./Apr. 1994.

[18] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "Dual-polarized model of outdoor propagation environments for adaptive antennas," *Proc. IEEE Vehicular Technology Conf. (VTC'99)*, pp. 990–995, May 1999.

[19] COST259, Commission of the European Communities, "Modeling Unification Workshop," TD(99) 061, Vienna, Austria, Apr. 1999.

[20] *European digital cellular telecommunication system (phase $2+$): Radio transmission and reception*, July 1996. Std. GSM 05.05, draft.

[21] D. Giancola, A. Sanguanini, and U. Spagnolini, "Variable rank receiver structures for low rank ST channels," in *Proc. Vehicular Technology Conf. (VTC'99)*, vol. 1, May 1999, pp. 65–69.

[22] J. Chun and T. Kailath, *Generalized Displacement Structure for Block-Toeplitz, Toeplitz-Block and Toeplitz-Derived Matrices*, ser. NATO ASI, vol. F70. Berlin, Germany: Springer-Verlag, 1991.

[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

[24] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud, "Least sum of squared errors channel estimation," *Proc. Inst. Elec. Eng.–F*, vol. 138, pp. 371–378, Aug. 1991.

[25] S. Y. Kung, *VLSI Array Processors*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[26] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Trans. Commun.*, vol. 42, pp. 1000–1011, Feb./Mar./Apr. 1994.

[27] J. H. Winters, "Optimum combining in digital mobile radio with co-channel interference," *IEEE J. Select. Areas Commun.*, vol. 2, pp. 528–539, July 1984.

[28] M. Nicoli and U. Spagnolini, "Multiuser space-time channel estimation for CDMA under the reduced-rank constraint," in *Proc. IEEE GLOBECOM 2000*, vol. 1, Nov. 2000, pp. 147–151.

Optimal Binary Training Sequence Design for Multiple-Antenna Systems Over Dispersive Fading Channels

Shan-An Yang and Jingshown Wu

***Abstract*—Accurate and efficient channel estimation is important in multiple-antenna communication systems in order to effectively reduce the mutual interference among different transmitting antennas. For a nondispersive channel that is modeled by a single tap for each transmitting and receiving antenna pair, the well-known Hadamard sequences can be applied to estimate the channel coefficients. However, for a dispersive channel that has multipath problem and is modeled by multiple taps, the optimal sequences must have both good autocorrelations and cross correlations. The existence of binary sequences with such good property is an open problem. In this paper, we devise an algorithm to find these sequence sets. These codes can be applied in multiple-antenna systems.**

***Index Terms*—Channel estimation, fading, MIMO, training sequence.**

I. INTRODUCTION

Efficient channel estimation is important for multiple-antenna systems especially when the number of antennas increases. To avoid the degradation of estimation accuracy due to interference, an intuitive way is to transmit training sequences for each transmitting antenna in turn [1]. For a system with M antennas, this scheme requires M times bandwidth compared with a single antenna transmitter system. However, orthogonal training sequences can be simultaneously applied for each transmitter antenna to estimate the channel efficiently [2], [3]. For a single tap coefficient discrete channel model, it is well known that orthogonal sequences are the optimal training sequences that minimize the estimation errors if the additive noises are identical independent Gaussian random processes. In this case, a Hadamard matrix can be applied. However, in the case of multipath channel, the channel for each pair of transmitting and receiving antennas should be modeled by several taps. It can be proven that the training sequences should have both good autocorrelation and cross correlation. Existence of such training sequence sets is still an open problem. In this paper, we discuss the existence of such optimal binary training sequence sets and propose a search algorithm.

Manuscript received October 9, 2000; revised July 10, 2001. This work was supported in part by the National Science Council and the Ministry of Education, Taiwan, R.O.C., under Grants NSC89-2213-E002-061 and 89-E-FA06-2-4.

The authors are with the Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: wujsh@cc.ee.ntu.edu.tw).

Digital Object Identifier 10.1109/TVT.2002.800638

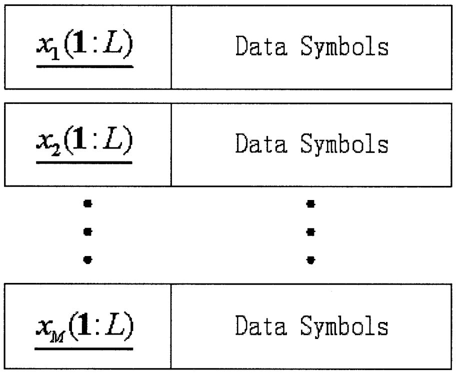

Fig. 1. The burst data structure for M-antenna transmission.

II. Problem Description

The multiple-antenna system under consideration has M transmitting antennas and N receiving antennas. The burst data structure for each transmitting antenna is shown in Fig. 1, where $x_\alpha(1:L)$, $\alpha = \mathbf{1}, \mathbf{2} \ldots M$, denotes the training sequence to be transmitted from the αth antenna. L is the length of each training sequence. The training sequences are embedded in each burst. Data bursts from different antennas will have different training sequences that are designed together so that the coexistence of the training sequences does not affect the channel estimation accuracy. Provided that the burst is short and the channel is quasi-static within a burst, the output of discrete equivalent channel can be expressed as

$$y_\beta(k) = \sum_{\alpha=\mathbf{1}}^{M} \sum_{i=\mathbf{0}}^{V} h_{\alpha\beta}(i) x_\alpha(k-i) + n_\beta(k) \tag{1}$$

where V is the order of channel memory and $h_{\alpha\beta}(i)$ denotes the response of the βth receiving antenna of the receiver to a discrete unit sample applied in the αth transmitting antenna. $n_\beta(k)$ is assumed to be identical independently distributed Gaussian random noise. With good synchronization, small value of V is enough to well approximate the channel.

With simple manipulations, we can prove that the training sequence set is optimal if the training sequence in each antenna is not only orthogonal to its shifts within V taps but also orthogonal to the training sequences in other antennas and their shifts within V taps (see Appendix). In other words, the optimal training sequences should satisfy the following equation:

$$\begin{aligned} &\sum_{k=1}^{L-V} x_\alpha(k) x_\beta(k+s) = 0 \\ &\text{where } \alpha, \beta = 1 \sim M \\ &s = \begin{cases} 0 & \sim V, \quad \text{when } \alpha \neq \beta \\ 1 & \sim V, \quad \text{when } \alpha = \beta. \end{cases} \end{aligned} \tag{2}$$

III. Properties and a Search Algorithm of the Optimal Training Sequences for BPSK

In this section, we will discuss existence of the optimal training sequence selected from the binary phase-shift keying (BPSK) constellation. In other words, we will discuss the binary $\{1, -1\}$-sequence sets satisfying (2). First, we prove that $x_\alpha(k) = x_\alpha(k+P)$ for $k \leq V$. For convenience, we denote $P = (L - V)$. Consider $s = 1$ and $\beta = \alpha$ in (2); we have

$$\sum_{k=1}^{P} x_\alpha(k) x_\alpha(k+1) = 0. \tag{3}$$

Let $y(k) = x_\alpha(k)x_\alpha(k+1)$. Thus, $y(k) = -1$ implies $x_\alpha(k) = -x_\alpha(k+1)$. Since $\sum_{k=1}^{P} y(k) = 0$, half of $y(k), k = 1 \sim P$, are -1. Since $P/2$ is an even number (see Property 1 in Table I), we have $x_\alpha(P+1) = (-1)^{P/2} x_\alpha(1) = x_\alpha(1)$. As a consequence, finding optimal sequences $x_\alpha(1:L)$ is equivalent to obtain $x_\alpha(1:P)$, which satisfies the following equation:

$$\begin{aligned} &\sum_{k=1}^{p} x_\alpha(k) x_\beta(k + s \bmod P) = 0 \\ &\text{where } \alpha, \beta = 1 \sim M \\ &s = \begin{cases} 0 & \sim V, \quad \text{when } \alpha \neq \beta \\ 1 & \sim V, \quad \text{when } \alpha = \beta. \end{cases} \end{aligned} \tag{4}$$

Hereafter, we will call a sequence set that satisfies (4) as a (P, V, M) code. Some properties of this code are listed in Table I. For $M \geq 2$ and $V \geq 1$, we require that $x_1(1:P), x_1(2:P+1), x_2(1:P)$, and $x_2(2:P+1)$ are mutually orthogonal. This leads to result that P must be a multiple of four. (This is the same reason that the order of a Hadamard matrix is multiples of four.) Thus, we have Property 1. Property 2 states that multiplying -1 to any sequence in a (P, V, M) code can result another (P, V, M) code. Properties 3, 4, and 5 are quite trivial and can be easily verified. Property 3 implies that we can find all the $(P, V, M-1)$ codes and then find all the (P, V, M) codes based on all the $(P, V, M-1)$ codes. Property 4 says that if we reverse every sequence simultaneously, the result is still a (P, V, M) code. Property 5 implies that if we shift every sequence by the same amount, the result is still a (P, V, M) code. Property 6 gives an upper bound of M when the values P and V are given.

Based on the above properties, we can restrict the first entry of each seqeunce to -1 (or 1) and restrict the sequences permutation in a (P, V, M) code to the order in the $(P, V, 1)$ code table without loss of generality. We describe the search algorithm as follows.

Step 4) Find all of the $(P, V, 1)$ codes for the given values of P and V. Construct the $(P, V, 1)$ code table; each code in the code table is presented by an unique number. As in Fig. 2, for example, $(P, V, 1)$ #2 represents the second $(P, V, 1)$ code in the $(P, V, 1)$ code table.

Step 5) Find all the $(P, V, 2)$ codes from all the pairs of $(P, V, 1)$ by checking the orthogonality conditions. At the same time, we construct an index table in order to reduce the complexity for $M = 3$.

Step 6) Let $m = 3$.

Step 7) Construct the (P, V, m) code table and index table by applying the $(P, V, m-1)$ code table and index table. Note that a (P, V, m) code is composed of two $(P, V, m-1)$ codes, where the last $m - 2(P, V, 1)$ codes of a $(P, V, m-1)$ code is identical to the first $m - 2(P, V, 1)$ codes of the other $(P, V, m-1)$ code. As a result, we only have to check the orthogonality of the first $(P, V, 1)$ code in the former $(P, V, m-1)$ code and the last $(P, V, 1)$ code in the latter $(P, V, m-1)$ code to determine whether these two codes constitute a (P, V, m) code when $m \geq 3$.

Step 8) If $m > M$ or the newly constructed code table is empty, then the process is finished. Otherwise, increase m by one and go back to Step 4).

In Fig. 2, we illustrate how the code table and index table are constructed. As illustated in the first three columns, code $(P, V, 2)$ #1 is composed of code $(P, V, 1)$ #1 and #3. We build the code table for $M = 2$ by testing all possible pairs of $(P, V, 1)$ and list them in the third column. At the same time, an index table is constructed with its contents pointing to the starting positions of the corresponding codes in the code table. In order to find the $(P, V, 3)$ codes, we start from the

TABLE I
SOME PROPERTIES OF A (P, V, M) CODE

Property 1	P is a multiple of 4 if a (P, V, M) code exists for $M \geq 2$ and $V \geq 1$.
Property 2	If $\{x_i \mid i = 1 \sim M\}$ is a (P, V, M) code, then multiplying x_i by -1 yields another (P, V, M) code.
Property 3	Any subset composed of $M-1$ sequences of a (P, V, M) code is a $(P, V, M-1)$ code.
Property 4	If $\{x_i \mid i = 1 \sim M\}$ is a (P, V, M) code and $y_i(k) = x_i(P+1-k)$ then $\{y_i \mid i = 1 \sim M\}$ is a (P, V, M) code.
Property 5	If $\{x_i \mid i = 1 \sim M\}$ is a (P, V, M) code and $y_i(k) = x_i((k+1) \bmod P)$ then $\{y_i \mid i = 1 \sim M\}$ is a (P, V, M) code.
Property 6	If there is a (P, V, M) code, then $M(V+1) \leq P$.

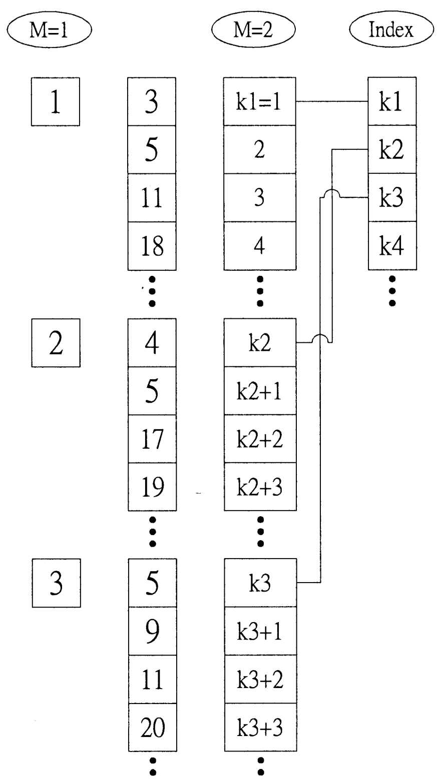

Fig. 2. The construction of the code table and the index table.

first element in the code table for $M = 2$. Since the second element of $(P, V, 2)$ #1 is $(P, V, 1)$ #3, we refer to the third and fourth elements in the index table which are denoted by $k3$ and $k4$. So, we search from $k3$th through $(k4-1)$th codes in the code table to see if any code can pair with $(P, V, 2)$ #1 to constitute a $(P, V, 3)$ code. We only have to check the orthogonality between the first $(P, V, 1)$ code in $(P, V, 2)$ #1 and the last $(P, V, 1)$ code in the other $(P, V, 2)$ codes between $k3$th through $(k4-1)$th. After finding all the $(P, V, 2)$ codes which can be paired with $(P, V, 2)$ #1, we continue with $(P, V, 2)$ #2 and then #3 until reaching the last element of the code table.

A. Computational Complexity Analysis

For (P, V, M) codes, restricton on the first entry of the sequence has a reduction factor of 2^M times. In addition, restriction on the permutation reduces the computation complexity by M factorial times. It can be proven that the total number of $(P, V = 1, M = 1)$ sequences is ${}_nC_{n/2}$. For greater values of M, the computational complexity depends on the total number of codes that exist and is hard to analyze. Thus, we take $(P, V) = (16, 1)$ as an example and give the numerical results in Table II. Columns (A) and (B) demonstrate the total number of codes found by the algorithm and the total number of codes without any restriction, respectively. Column (C) shows the reduction factor defined as a ratio between Column (B) and Column (A). Column (D) gives the probability of finding such (P, V, M) code by random guess. It is observed that the values in Column (A) increase at first, reaches its maximum value at $M = 3$, and decreases thereafter. On the contrary, the values in Column (B) increase all the way until $M = 8$. To determine that two sequences satisfy (4), we have to test orthogonality $(2V + 1)$ times. Each time it requires an operation of bit-wise XOR between two P-bits words and an operation to count the total number of ones in a P-bit word. In Columns (E) and (F), we show the total numbers of orthogonality tests required in each round for the proposed algorithm and for a method by directly searching all combinations of $\{-1, 1\}$ in each bit. This implies that our algorithm provides a feasible approach for a personal computer to determine existence or nonexistence of (P, V, M) codes for a small value of P. Although the proposed algorithm will become computationally impractical for a large value of P, it suffices for the purpose of designing training sequences in multiple-antenna systems since they are often short.

B. An Alternative Way of Construction

If we do not intend to find the maximum value of M, we can construct such codes by applying shifts of a binary almost perfect sequence (BAPS) [4], [5]. Take $M = 2$ as an example; we can choose $x_1(k) = a_k$ and $x_2(k) = a_{k+p/4}$, where $k = 1 \sim p, V \leq p/4 - 1$, and a_k is a BAPS. For example, the

TABLE II
COMPARISON BETWEEN THE ALGORITHM AND THE DIRECT EXHAUSIVE SEARCH WITHOUT ANY RESTRCTION

M	(A)	(B)	(C)	(D)	(E)	(F)
1	12870	25740	2	0.392	3.28E4	1.31E5
2	2217250	17738000	8	4.1E-3	8.28E7	4.29E9
3	11262888	540618624	48	1.9E-6	9.48E8	1.41E14
4	3206212	1231185408	384	6.7E-11	1.61E9	4.61E18
5	292832	1124474880	3840	9.3E-16	3.75E6	1.51E23
6	37408	1723760640	46080	2.2E-20	2.16E5	4.95E27
7	8000	5160960000	645120	9.9E-25	3.32E4	1.62E32
8	960	9909043200	10321920	2.9E-29	3.84E3	5.31E36

(A) Number of Codes with the proposed restrictions
(B) Actual Number of Codes without any restriction
(C) Reduction Factor
(D) Probability of finding by random guess
(E) Tests required in each round with the proposed algorithm
(F) Tests required in each round with by the most naïve method

TABLE III
THE MAXIMUM VALUES OF M GIVEN (P, V)

M \ V, P	1	2	3	4	5	6	7	8	9
4	2	1	1	N	N	N	N	N	N
8	4	2	1	N	N	N	N	N	N
12	5	2	1	1	1	N	N	N	N
16	8	4	2	2	1	1	1	N	N

sequence $\{-1,-1,1,-1,-1,-1,-1,-1,1,1,-1,1,-1,1,1,1\}$ is a BAPS with $p = 16$. Thus,

$$x_1 = \{-1,-1,1,-1,-1,-1,-1,-1,1,1,-1,1,-1,1,1,1\}$$

and

$$x_2 = \{-1,-1,-1,-1,1,1,-1,1,-1,1,1,1-1,-1,1,-1\}$$

satisfy the conditions of the optimal training sequences for $(P = 16, V = 3, M = 2)$. However, we see in Table III that the maximal number of achievable V is 4, not 3. Therefore, although applying BAPS helps to construct such code, it does not achieve the maximum value of M or V.

IV. SEARCH RESULT AND NUMERICAL SIMULATION

In Table III, we list the maximal numbers of achievable M given P and V. In the table, N means that no such (P, V, M) code exists. One can observe that the M value for some (P, V) combinations achieve the upper bound given by Property 6, while others do not. For example, when $P = 16$ and $V = 1$, the maximum achievable M is 8, which is just the upper bound given by Property 6. However, when $(P, V) = (12, 1)$, the maximal achievable M is only 5, not 6.

In Table IV, we list at least one example for all existing codes with $P \leq 16$. Although we have proved the optimal property of the proposed training seqeunces under the assumption of quasi-coherent channel, we are also interested in their performance in an environment with Doppler frequency shift. We perform numerical simulation to compare three different training sequence sets. The sequences under test are listed in Table V. The first one is the optimal sequence set with $(P, V, M) = (8, 1, 4)$ as proposed. The second sequence set

TABLE IV
EXAMPLES OF THE (P, V, M) CODES

P	V	M	(P,V,M) code
4	1	2	X1={1 1 1 0}, X2={1 0 1 1}
8	1	4	X1={10000010},X2={10110001} X3={11010111},X4={11100100}
12	1	5	X1={1 0 0 0 0 0 0 1 0 0 1 0} X2={1 0 0 1 0 1 0 0 1 1 1 1} X3={1 0 0 1 1 0 1 1 1 0 0 1} X4={1 0 1 0 1 1 0 0 0 0 0 1} X5={1 1 1 1 0 0 0 1 0 1 0 1}
12	2	2	X1={1 1 1 0 1 0 0 0 1 1 1 0} X2={1 1 1 1 0 1 1 0 1 0 0 0}
16	1	8	X1={0000000001010101} X2={0000111101011010} X3={0011001101100110} X4={0011110001101001} X5={0101010100000000} X6={0101101000001111} X7={0110011000110011} X8={0110100100111100}
16	2	4	X1={0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1} X2={0 0 1 1 0 0 0 1 0 1 1 1 1 0 1 0} X3={0 1 0 1 0 1 1 1 0 0 0 1 1 1 0 0} X4={0 1 1 0 0 1 0 0 0 0 1 0 1 1 1 1}
16	4	2	X1={1 0 0 0 0 0 1 0 0 1 1 1 0 0 1 0} X2={1 0 1 1 0 0 0 1 0 1 0 0 0 0 0 1}

Ps: '-1' is represented by '0' in the table.

TABLE V
THE CODES UNDER TEST IN THE NUMERICAL SIMULATION

PROPOSED CODE	x1={1 -1 -1 -1 -1 -1 1 -1} x2={1 -1 1 1 -1 -1 -1 1} x3={1 1 -1 1 -1 1 1 1} x4={1 1 1 -1 -1 1 -1 -1}
PRBS	x1={-1 -1 1 -1 1 1 1} x2={1 -1 1 1 1 -1 -1} x3={1 1 1 -1 -1 1 -1}
ARBITRARY	x1={1 -1 1 -1 -1 -1 1 -1} x2={1 -1 1 1 -1 1 -1 1} x3={1 -1 -1 1 -1 1 1 1} x4={1 1 -1 1 -1 1 1 -1}

is constructed with a well-known pseudorandom binary sequence (PRBS) with different shifts in different transmitter antennas. Since we want to keep the cross correlation low between shifts, the maximal number of transmitting antennas is three with a PRBS of length seven. The transmitted power is increased to compensate the shorter length for a fair comparison. The third one is an arbitrarily chosen sequence set. The channel tap coefficients are assumed to be independent complex Gaussian random variable with uniformly distributed phaseand Rayleigh distributed amplitude. The transmitted power from each transmitting antenna is assumed to be the same. Here, we use the well-known Jakes' model to perform the simulation. The results are shown in Fig. 3. We see that when the Doppler frequency shift

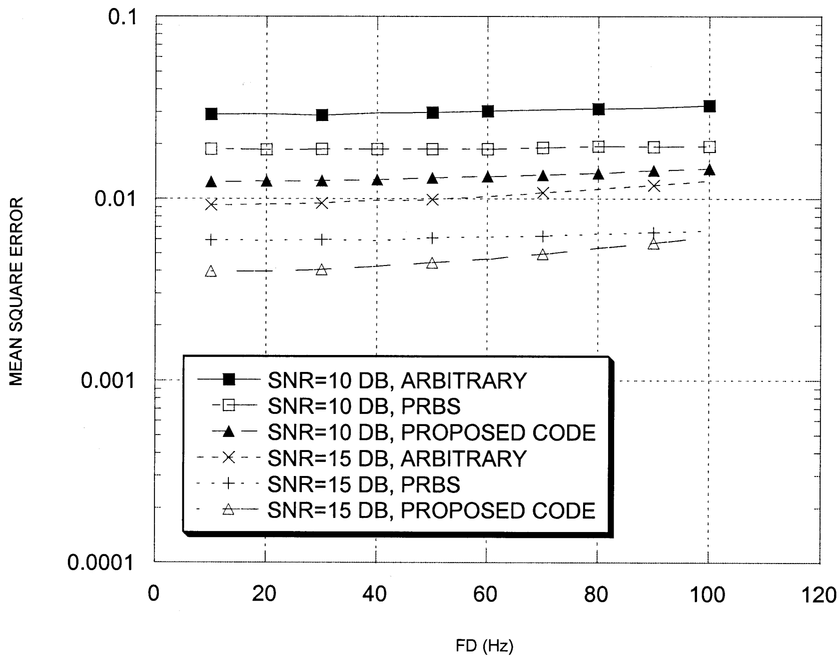

Fig. 3. Performance comparison of three sequence sets with Doppler frequency shift.

is not severe, the advantage of the proposed sequence set remains unchanged.

V. Conclusion

In this paper, we study the design of optimal trianing sequence sets for multiple-antenna communication systems in a dispersive fading environment. The conditions of the optimal training sequences for the multiple-antenna systems are proposed and proven. We also propose an algorithm to search the optimal training sequences and analyze the complexity of the algorithm. Existence of such codes is shown by an exhaustive search for code length less than or equal to 16. Examples of the search result are listed in a table. Numerical tests are performed to test their performance in a nonideal environment. We believe that these sequences can be used for channel estimation in multiple-antenna communication systems.

Appendix
Proof of the Condition for Optimal Training Sequences

The first V sampled values are discarded to avoid the interference from symbols prior to the first training symbol. Thus, we can rewrite (1) in matrix form as

$$\underline{y_\beta} = \underline{\underline{X}}\,\underline{H_\beta} + \underline{N_\beta} \tag{A.1}$$

where $\underline{y_\beta} = [y_\beta(V+1)\ \ y_\beta(V+2)\ \ \cdots\ \ y_\beta(L)]^T$ (see the equation at the bottom of the page). If $\underline{\underline{X}}^H\underline{\underline{X}}$ is invertible, the least squares estimation for the channel matrix $\underline{H_\beta}$ is given by

$$\underline{\hat{H}_\beta} = (\underline{\underline{X}}^H\underline{\underline{X}})^{-1}(\underline{\underline{X}}^H\underline{y_\beta}). \tag{A.2}$$

The above estimation is unbiased. Thus, the estimation error and variance are given by

$$\underline{e_\beta} = \underline{\hat{H}_\beta} - \underline{H_\beta} \tag{A.3}$$

and

$$\begin{aligned} E[\|\underline{e_\beta}\|^2] &= E\left[\mathrm{tr}\left[\underline{e_\beta e_\beta^H}\right]\right] \\ &= \mathrm{tr}\left[(\underline{\underline{X}}^H\underline{\underline{X}})^{-1}\underline{\underline{X}}^H E\left[\underline{N_\beta N_\beta^H}\right]\underline{\underline{X}}(\underline{\underline{X}}^H\underline{\underline{X}})^{-1}\right]. \end{aligned} \tag{A.4}$$

In the case of independent white discrete Gaussian noises, we have

$$E\left[\underline{N_\beta N_\beta^H}\right] = \sigma_n^2 \underline{\underline{I_{L-V}}}. \tag{A.5}$$

Therefore, (A.4) is simplified as

$$E\left[\|\underline{e_\beta}\|^2\right] = \sigma_n^2\,\mathrm{tr}[(\underline{\underline{X}}^H\underline{\underline{X}})^{-1}] = \sigma_n^2 \sum_{k=1}^{(MV+M)} \frac{1}{\lambda_k} \tag{A.6}$$

$$\underline{\underline{X}} = \begin{pmatrix} \underline{x_1(1:V+1)} & \underline{x_2(1:V+1)} & \cdots & \underline{x_M(1:V+1)} \\ \underline{x_1(2:V+2)} & \underline{x_2(2:V+2)} & \cdots & \underline{x_M(2:V+2)} \\ \vdots & \vdots & \vdots & \vdots \\ \underline{x_1(L-V:L)} & \underline{x_2(L-V:L)} & \cdots & \underline{x_M(L-V:L)} \end{pmatrix}$$

$$\underline{H_\beta} = [\underline{h_{1\beta}}\ \ \underline{h_{2\beta}}\ \ \cdots\ \ \underline{h_{M\beta}}]^T,\quad \underline{N_\beta} = [n_\beta(V+1)\ \ n_\beta(V+2)\ \ \cdots\ \ n_\beta(L)]^T$$

$$\underline{x_\alpha(i:j)} \hat{=} [x_\alpha(i)\ \ x_\alpha(i+1)\ \ \cdots\ \ x_\alpha(j)],\quad \underline{h_{\alpha\beta}} = [h_{\alpha\beta}(V)\ \ h_{\alpha\beta}(V-1)\ \ \cdots h_{\alpha\beta}(0)]$$

$$i = 1 \sim L-V,\quad j = i+V,\quad \alpha = \mathbf{1}, \mathbf{2}, \ldots, M \quad \text{and} \quad \beta = \mathbf{1}, \mathbf{2}, \ldots, N.$$

where λ_k denotes the eigenvalues of $\underline{\underline{X}}^H \underline{\underline{X}}$. By applying Cauchy inequality, we have

$$\left(\sum_{k=1}^{(MV+M)} \frac{1}{\lambda_k}\right)\left(\sum_{k=1}^{(MV+M)} \lambda_k\right) \geq M^2(V+1)^2. \tag{A.7}$$

Thus, minimization of (A.7) requires that λ_k is constant for all $k = \mathbf{1} \sim (MV+M)$. This implies that $\underline{\underline{X}}^H \underline{\underline{X}}$ is a diagonal matrix with all the diagonal entries equal to a constant.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their valuable comments and suggestion.

REFERENCES

[1] R. Cusani, E. Baccarelli, G. D. Blasio, and S. Galli, "A simple polarization-recovery algorithm for dual-polarized cellular mobile-radio systems in time-variant faded environments," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 220–228, Jan. 2000.

[2] M. Kavehrad, "Performance of cross-polarized $\boldsymbol{M}$-ary QAM signals over nondispersive fading channels," *AT&T Bell Lab. Tech. J.*, vol. 63, no. 3, pp. 499–521, Mar. 1984.

[3] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space-time coding modem for high-data-rate wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1459–1478, Oct. 1998.

[4] J. Wolfmann, "Almost perfect autocorrelation sequences," *IEEE Trans. Inform. Theory*, vol. 28, pp. 1412–1418, July 1992.

[5] A. Pott and S. P. Bradley, "Existence and nonexistence of almost-perfect autocorrelation sequences," *IEEE Trans. Inform. Theory*, vol. 41, pp. 301–304, Jan. 1995.