Design of Delay-Tolerant Space-time Codes with Linear MMSE Receivers

Wenjin Wang and Fu-Chun Zheng
School of Systems Engineering, The University of Reading, Reading, RG6 6AY, UK
{w.wang,f.zheng}@reading.ac.uk

***Abstract*—This paper presents a new design method for delay-tolerant linear dispersion codes (DT-LDCs) for asynchronous cooperative communication networks. We consider a system that is equipped with linear minimum mean square error (MMSE) receiver. Based on the DT-LDC framework, we propose a new design method to yield DT-LDCs that approach near-optimal capacity as well as minimum average MSE. The proposed design employs stochastic gradient algorithm to guarantee the local optimum. Moreover, it is improved by using simulated annealing type optimization to approach the global optimum. Simulation results confirm the performance of the newly-proposed delay-tolerant LDCs.**

I. Introduction

Multiple-input multiple-output (MIMO) technology has been widely accepted as one of the effective ways to increase the capacity and reliability of wireless communications in the presence of fading [1]. However, in wireless communications, the individual nodes may be equipped with only a single antenna due to cost and size constraints. The cooperative communication technique allows single-antenna terminals to obtain the benefits of MIMO systems. The basic idea is that single-antenna terminals can share their antennas and create a virtual MIMO system, where MIMO technologies, such as space-time codes (STCs), can be implemented in a distributed manner [2], [3].

Unlike MIMO systems, the nodes in cooperative communications are usually arbitrary geographically located so that perfect synchronization between them is difficult to achieve. This brings a new challenge of designing distributed STCs: the codes that are designed for conventional synchronous multi-antenna systems are no longer effective in such asynchronous scenarios. Instead, an efficient code design should satisfy the full-diversity order for any delay profile, which is called the delay-tolerant property [4].

In [4] and [5], delay-tolerant space-time trellis codes and linear convolutional STCs were introduced respectively. In comparison, space-time block coding (STBC) based schemes involve much shorter codewords and tend to have relatively lower computational complexity. Damen and Hammons constructed a new set of delay-tolerant threaded algebraic space-time (DT-TAST) codes in [6] and an extension of this framework with minimum length has appeared in [7]. To achieve diversity-multiplexing trade-off optimality, Sarkiss *et al.* recently used the cyclic division algebras of perfect codes to construct new STBC for asynchronous systems, particularly for 2×2, 3×3 and 4×4 cases [8]. These efficient delay-tolerant STBCs are proposed for scenarios that have specified constraints on relations among the length of codewords, the number of antennas (relay nodes) and the number of modulated symbols. To overcome this drawback, we proposed a new method in [9] to design delay-tolerant linear dispersion codes (DT-LDCs), which provided a more generalized framework allowing for flexible parameters of STCs.

The good error performance of both DT-TAST codes and DT-LDCs depends on the maximum-likelihood (ML) detector. However, the computational complexity of ML detector, even when it is implemented via sphere decoding, will be prohibitive if the number of transmitter antennas or the constellation size is large. On the other hand, the minimum mean square error (MMSE) receiver is relatively simple to implement. In this paper, we investigate LDC design methods for asynchronous relay networks with linear MMSE in the receiver. Based on the DT-LDC framework, we propose a new design method to yield DT-LDCs that approach near-optimal capacity as well as minimum average MSE. As it is difficult to find the close-form solution because of the non-concave objective function, we use a stochastic gradient algorithm (SGA) to solve the programming, where the real gradient is estimated by Monte-Carlo simulations [10]. Since the SGA can reach a local minimum only, we employ simulated annealing optimization in SGA to approach the global minimum.

II. Delay-Tolerant Linear Dispersion Codes

A. Asynchronous Cooperative Systems

We consider a wireless network constitutes of $M+2$ nodes, where there is one source node, one destination node and M relay nodes. Each relay node is equipped only with a single antenna and the destination is equipped with N_R receive antennas. The transmission has two phases. In the first phase, the source node broadcasts the complex modulated symbols $(u_1, \cdots, u_Q)$ to the potential relay nodes. In the second phase, relay nodes simultaneously send their transmitted to the destination. In this paper, we use the DF protocol: the relay nodes decode the source message then send them to the destination node if they are correctly decoded. We assume that among M relay nodes there are N_T nodes that are able to achieve error free decoding.

Similarly to [6], [9], we also assume that the timing errors are integer multiples of the symbol duration and the fractional

This work was supported by the UK EPSRC under Grant EP/H016945/1.

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

timing errors are absorbed in the channel dispersion. Note that regarding quasi-synchronous relay networks with fractional delays, low complexity detectors such as [11] can be employed to deal with the inter-symbol-interference. Let $s_n(k)$ denote the transmit signal of nth relay node at kth instant. Then the received signal at mth antenna and kth instant at the destination can be written as

$$y_m(k) = \sum_{n=1}^{N_T} h_{m,n} s_n(k-\delta_n) + z_m(k) \tag{1}$$

where $z_m(k)$ is the corresponding additive white Gaussian noise with variance σ_z^2, $h_{m,n}$ is the channel coefficient between the nth relay node to mth receive antenna and δ_n is the relative delay of the received signal from nth relay node as referenced to the earliest received relay signal. In the rest of the paper, we use $\Delta \triangleq (\delta_1, \cdots, \delta_{N_T})$ to denote the delay profile of an asynchronous relay network and $\delta_{\max} \triangleq \max\{\delta_1, \cdots, \delta_{N_T}\}$ to denote the maximum of relative delays.

B. Linear Dispersion Codes

Let Q and T denote the number of modulated symbols and the length of a STC codeword. We can construct a linear dispersion code (LDC) as in [12]

$$\mathbf{S} = \sum_{q=1}^{Q} \left(u_q \mathbf{C}_q + u_q^* \mathbf{D}_q \right) \tag{2}$$

where the $\mathbf{C}_q$ and $\mathbf{D}_q$ are fixed $T \times N_T$ complex matrices. The (k,n)th element of $\mathbf{S}$ is the transmit signal $s_n(k)$. The code is completely determined by the set of dispersion matrices $\{\mathbf{C}_q, \mathbf{D}_q\}$, whereas each individual codeword is determined by the symbols $u_1, \cdots, u_Q$.

Define $\mathbf{u} \triangleq [u_1, \cdots, u_Q]^T$, $\mathbf{C} \triangleq [\mathrm{vec}(\mathbf{C}_1), \cdots, \mathrm{vec}(\mathbf{C}_Q)]$, $\mathbf{D} \triangleq [\mathrm{vec}(\mathbf{D}_1), \cdots, \mathrm{vec}(\mathbf{D}_Q)]$ and $T' = T + \delta_{\max}$. Submitting (2) into (1), the signal model of a linear dispersion coded asynchronous relay network can be written into a vector-matrix form:

$$\mathbf{y} = (\mathbf{HGCu} + \mathbf{HGDu}^*) + \mathbf{z} \tag{3}$$

where $\mathbf{y} = [\mathbf{y}_1^T, \cdots, \mathbf{y}_{N_R}^T]^T$, $\mathbf{z} = [\mathbf{z}_1^T, \cdots, \mathbf{z}_{N_R}^T]^T$, $\mathbf{H} = \mathbf{H}_w \otimes \mathbf{I}_{T'}$ and $\mathbf{G} = \mathrm{diag}(\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_{N_T})$, where

$$\begin{aligned}
\mathbf{y}_m &= [y_m(1), y_m(2), \cdots, y_m(T')]^T \\
\mathbf{z}_m &= [z_m(1), z_m(2), \cdots, z_m(T')]^T \\
\mathbf{H}_w &= \begin{bmatrix} h_{1,1} & \cdots & h_{1,N_T} \\ \vdots & \ddots & \vdots \\ h_{N_R,1} & \cdots & h_{N_R,N_T} \end{bmatrix} \\
\mathbf{G}_i &= \begin{bmatrix} \mathbf{O}_{\delta_i \times T} \\ \mathbf{I}_T \\ \mathbf{O}_{(\delta_{\max}-\delta_i) \times T} \end{bmatrix}
\end{aligned}$$

By separating the real and imaginary parts of the signals, the signal model in (3) can also be written as real-valued one as

$$\tilde{\mathbf{y}} = \tilde{\mathbf{H}} \tilde{\mathbf{G}} \boldsymbol{\mathcal{X}} \tilde{\mathbf{u}} + \tilde{\mathbf{z}} \tag{4}$$

where $\tilde{\mathbf{y}} = [\Re(\mathbf{y}^T)\ \Im(\mathbf{y}^T)]^T$, $\tilde{\mathbf{z}} = [\Re(\mathbf{z}^T)\ \Im(\mathbf{z}^T)]^T$, $\tilde{\mathbf{u}} = [\Re(\mathbf{u}^T)\ \Im(\mathbf{u}^T)]^T$, $\tilde{\mathbf{G}} = \mathbf{I}_2 \otimes \mathbf{G}$,

$$\tilde{\mathbf{H}} = \begin{bmatrix} \Re(\mathbf{H}) & -\Im(\mathbf{H}) \\ \Im(\mathbf{H}) & \Re(\mathbf{H}) \end{bmatrix},$$

and

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix} \Re(\mathbf{C}+\mathbf{D}) & -\Im(\mathbf{C}-\mathbf{D}) \\ \Im(\mathbf{C}+\mathbf{D}) & \Re(\mathbf{C}-\mathbf{D}) \end{bmatrix}. \tag{5}$$

Notice that (4) represents a generalized $N_R T' \times Q$ linear signal model for asynchronous relay networks. Assuming that $u_1, \cdots u_Q$ have unit-variance and the transmit signal power is normalized, then we have $\mathrm{Tr}(\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{X}}^T) \leq 2N_T T$. If $\tilde{\mathbf{G}} = \mathbf{I}_{2N_T T}$, then (4) reduces to the signal model of LDCs in synchronous networks. Hence, the LDCs for synchronous systems can be considered as a special case and can also be designed within this framework.

C. Linear MMSE Equalizer

Based on the real-valued equivalent MIMO channels in (4), we employ the linear MMSE equalizer at the receiver to obtain the estimation of transmitted symbol $\tilde{u}_k$ $(k = 1, \cdots, 2Q)$, which is the kth element of symbol vector $\tilde{\mathbf{u}}$. The MMSE equalizer is to find $\hat{\tilde{u}}_k = \mathbf{w}_k \tilde{\mathbf{y}}$ so that $\mathbb{E}\left[|\hat{\tilde{u}}_k - \tilde{u}_k|^2\right]$ is minimized, where $\mathbf{w}_k$ is the MMSE coefficients to be determined for $\tilde{u}_k$. The solution is given by [13]

$$\mathbf{w}_k = \mathbf{e}_k^T \left(\boldsymbol{\mathcal{X}}^T \tilde{\mathbf{G}}^T \tilde{\mathbf{H}}^T \tilde{\mathbf{H}} \tilde{\mathbf{G}} \boldsymbol{\mathcal{X}} + \sigma_z^2 \mathbf{I}_{2Q} \right)^{-1} \boldsymbol{\mathcal{X}}^T \tilde{\mathbf{G}}^T \tilde{\mathbf{H}}^T. \tag{6}$$

With the MMSE equalizer, the interference caused by the other transmitted symbols is suppressed but cannot be totally eliminated, which is usually considered as residual Gaussian distributed interference and noise. Thus, the symbol-by-symbol equivalent channel between $\hat{\tilde{u}}_k$ and $\tilde{u}_k$ is given by

$$\hat{\tilde{u}}_k = \rho_k \tilde{u}_k + v_k, \tag{7}$$

where ρ_k is given by

$$\rho_k = \mathbf{w}_k \tilde{\mathbf{H}} \tilde{\mathbf{G}} \boldsymbol{\mathcal{X}} \mathbf{e}_k, \tag{8}$$

v_k is the corresponding zero-mean Gaussian residual interference and noise. It is easily to be verified that the variance of v_k can be calculated by

$$\sigma_{v,k}^2 = \frac{1}{2} \rho_k (1 - \rho_k) \tag{9}$$

respectively. Then the MSE of estimation is given by

$$\mathrm{MSE}_k \triangleq \mathbb{E}_{\tilde{u}_k} \left(|\hat{\tilde{u}}_k - \tilde{u}_k|^2 \right) = \frac{1}{2}(1 - \rho_k) \tag{10}$$

III. Design of MMSE Optimal Codes

Now we consider the design criteria of LDCs in asynchronous systems with linear MMSE receiver. In this section, we will propose a method to design DT-LDCs which can approach near-optimal capacity as well as minimum MSE.

A. Ergodic Capacity

From signal model (4), the LDC matrix $\boldsymbol{\mathcal{X}}$ determines the ergodic capacity for the given delay profiles and channel types. Hence, one possible criterion of designing $\boldsymbol{\mathcal{X}}$ is to maximize the ergodic capacity. In [14], Heath *et al.* presented a near-optimal capacity condition for the LDC design in synchronous systems,

$$\boldsymbol{\mathcal{X}}^T\boldsymbol{\mathcal{X}} = \frac{2TN_T}{Q}\mathbf{I}_{2Q} \tag{11}$$

Particularly, for the special case where $Q = TN_T$ and $\boldsymbol{\mathcal{X}}$ is a square matrix, it was proved to be the sufficient condition to be a capacity-optimal LDC for synchronous MIMO systems.

However, the conclusion may not hold in asynchronous systems. In fact, it is even difficult to obtain a similar result because the statistics of delay profiles are required, which is usually unavailable. In [9], we have employed a widely-adopted Jensen's upper bound of ergodic capacity criterion instead and proved that if condition in (11) is satisfied, the resulting LDCs achieve optimal Jensen's upper bound of ergodic capacity in asynchronous relay networks. Although such DT-LDCs can provide near optimal ergodic capacity, they usually can not guarantee minimal error performance for specific systems. In the following, we will present a new criterion of DT-LDC design based on the minimum MSE.

B. MMSE Optimal DT-LDCs

From (7) and (10), we know that the error performance, e.g. bit error rate (BER) of $\tilde{u}_k$ will be determined by MSE_k. Hence, the LDC matrix will take effects on the final error performance. Given the delay profile $\boldsymbol{\Delta}$ and the channel $\tilde{\mathbf{H}}$, the total MSE of all $2Q$ real-valued transmitted symbols can be written as

$$\begin{aligned}\mathrm{MSE} &= \sum_{k=1}^{2Q}\mathrm{MSE}_k = Q - \frac{1}{2}\sum_{k=1}^{2Q}\rho_k \\ &= \mathrm{Tr}\left(\frac{\sigma_z^2}{2}\left(\boldsymbol{\mathcal{X}}^T\tilde{\mathbf{G}}^T\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\tilde{\mathbf{G}}\boldsymbol{\mathcal{X}} + \sigma_z^2\mathbf{I}_{2Q}\right)^{-1}\right)\end{aligned} \tag{12}$$

Then the unconditional average MSE is given by

$$\overline{\mathrm{MSE}} \triangleq \mathbb{E}_{\tilde{\mathbf{H}},\boldsymbol{\Delta}}\mathrm{MSE}. \tag{13}$$

The DT-LDC design should minimize average MSE:

$$\mathrm{Minimize}_{\boldsymbol{\mathcal{X}}} \qquad \overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}}), \quad s.t.\ \mathrm{Tr}(\boldsymbol{\mathcal{X}}^T\boldsymbol{\mathcal{X}}) = 2TN_T, \tag{14}$$

where the power constraint is to ensure normalized average transmit power. As we will show later, $\overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}})$ is continuously differentiable in $\boldsymbol{\mathcal{X}}$. Hence, the objective function attains a minimum on the compact set $\boldsymbol{\mathcal{X}}$. The objective functions in (14) are non-concave. Hence, the closed-form solution for such problem is usually not available. We employ the stochastic gradient algorithm (SGA) to approach the optimum [10] since the calculated gradient is an unbiased estimate of the real gradient.

C. Spherical Coordinate Mapping

The programming of (14) is a constrained optimization problem. In order to turn it into an equivalent unconstrained optimization problem, we use spherical coordinate mapping as in [9], [15]. Let $\boldsymbol{\chi} = \mathrm{vec}(\boldsymbol{\mathcal{X}}) = [\chi_1, \cdots, \chi_{4TN_TQ}]$ and $\boldsymbol{\phi} \triangleq [\phi_1, \cdots, \phi_{4TN_TQ-1}]^T \in [0, 2\pi]^{4TN_TQ-1}$. We use the following coordinate mapping,

$$\chi_l = \sqrt{2TN_T}\cos\phi_l\prod_{k=1}^{l-1}\sin\phi_l, \quad l = 1, \cdots, 4TN_TQ-1. \tag{15}$$

To satisfy the power constraint $\mathrm{Tr}(\boldsymbol{\mathcal{X}}^T\boldsymbol{\mathcal{X}}) = 2TN_T$, we have

$$\chi_{4TN_TQ} = \sqrt{2TN_T}\prod_{k=1}^{4TN_TQ-1}\sin\phi_k. \tag{16}$$

Therefore, the optimization problem in (14) becomes the following unconstrained optimization problem

$$\mathrm{Minimize}_{\boldsymbol{\phi}} \qquad \overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}}(\boldsymbol{\phi})). \tag{17}$$

We will now introduce SGA to solve such an optimization problem.

D. Stochastic Gradient Algorithm

The SGA for (17) can be written as

$$\boldsymbol{\phi}^{(n+1)} = \boldsymbol{\phi}^{(n)} - a_n\hat{\mathbf{g}}(\boldsymbol{\phi}^{(n)}) \tag{18}$$

where $\boldsymbol{\phi}^{(n)}$ is the value of $\boldsymbol{\phi}$ at nth iteration, $\hat{\mathbf{g}}(\boldsymbol{\phi}^{(n)})$ is the estimate of gradient, $\{a_n\}$ is a decreasing step-size sequence of positive real numbers satisfying $\sum_{n=1}^{+\infty} a_n = +\infty$ and $\sum_{n=1}^{+\infty} a_n^2 < +\infty$. The aim of gradient estimation is to compute an unbiased estimate of the true gradient, i.e., $\mathbb{E}\left\{\hat{\mathbf{g}}(\boldsymbol{\phi}^{(n)})\right\} = \nabla_{\boldsymbol{\phi}}\overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}}(\boldsymbol{\phi}))$. To this end, we firstly consider the gradient $\nabla_{\boldsymbol{\mathcal{X}}}\mathrm{MSE}(\boldsymbol{\mathcal{X}})$,

Proposition 1: Let $\mathbf{W} = \tilde{\mathbf{G}}^T\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\tilde{\mathbf{G}}$. $\mathrm{MSE}(\boldsymbol{\mathcal{X}})$ in (12) is differentiable and the gradient $\nabla_{\boldsymbol{\mathcal{X}}}\mathrm{MSE}(\boldsymbol{\mathcal{X}})$ is given by

$$\nabla_{\boldsymbol{\mathcal{X}}}\mathrm{MSE}(\boldsymbol{\mathcal{X}}) = -\sigma_z^2\mathbf{W}\boldsymbol{\mathcal{X}}(\boldsymbol{\mathcal{X}}^T\mathbf{W}\boldsymbol{\mathcal{X}} + \sigma_z^2\mathbf{I}_{2Q})^{-2} \tag{19}$$

Proof: See Appendix. ■

According to (13), the true gradient $\nabla_{\boldsymbol{\mathcal{X}}}\overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}})$ can be written as

$$\nabla_{\boldsymbol{\mathcal{X}}}\overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}}) = \mathbb{E}_{\Delta}\left\{\mathbb{E}_{\tilde{\mathbf{H}}}\left\{\nabla_{\boldsymbol{\mathcal{X}}}\mathrm{MSE}(\boldsymbol{\mathcal{X}})\right\}\right\} \tag{20}$$

Theoretically, the gradient in (20) requires an expectation over unlimited samples of $\tilde{\mathbf{H}}$. The proposed stochastic gradient algorithm employs a subset of samples of $\tilde{\mathbf{H}}$ to calculate the estimation of gradient by Monte Carlo simulations. In this way, the estimate of gradient is given by

$$\hat{\mathbf{g}}(\boldsymbol{\mathcal{X}}) \triangleq \frac{1}{N_\Delta N_c}\sum_{\Delta\in\mathcal{D}}\sum_{\mathbf{H}_w\in\mathcal{H}}\nabla_{\boldsymbol{\mathcal{X}}}\mathrm{MSE}(\boldsymbol{\mathcal{X}}) \tag{21}$$

TABLE I
DT-LDC DESIGN WITH SGA

Step 0: Given the set of parameters N_T, Q, N_R and T, randomly generate the $2TN_T \times 2Q$ matrix $\mathbf{A}$, calculate $\mathcal{X} = \sqrt{2TN_T}\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}$ such that $\mathcal{X}$ satisfy the condition of maximal mutual information.

Step 1: For each nth iteration, calculate $\hat{\mathbf{g}}(\boldsymbol{\phi}^{(n+1)})$. Update $\boldsymbol{\phi}^{(n+1)} = \boldsymbol{\phi}^{(n)} - a_n\hat{\mathbf{g}}(\boldsymbol{\phi}^{(n)})$. Then update $\mathcal{X}$ according to (15).

Step 2: Choose $\mathcal{X}$ after all iterations complete.

---

where $\mathcal{D}$ and $\mathcal{H}$ are the sets of all possible Δ and $\mathbf{H}_w$ respectively, and N_D and N_c are their cardinalities. It is easy to verify that the resulting estimate of the gradient satisfying $\mathbb{E}\{\hat{\mathbf{g}}(\boldsymbol{\mathcal{X}})\} = \nabla_{\boldsymbol{\mathcal{X}}}\overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}})$ and for a relatively large N_c, $\hat{\mathbf{g}}(\boldsymbol{\mathcal{X}})$ can approach the exact value of gradient $\nabla_{\boldsymbol{\mathcal{X}}}\overline{\mathrm{MSE}}(\boldsymbol{\mathcal{X}})$. From our experience, choosing N_c to be around 10000 works quite well and the computational complexity is acceptable. Thus, the gradient $\hat{\mathbf{g}}(\boldsymbol{\phi}^{(n)})$ can be calculated by

$$\hat{\mathbf{g}}(\boldsymbol{\phi}) = \frac{\partial \boldsymbol{\mathcal{X}}}{\partial \boldsymbol{\phi}}\hat{\mathbf{g}}(\boldsymbol{\mathcal{X}}). \tag{22}$$

The DT-LDC design can be summarized as in Table I.

E. Simulated Annealing Optimization

The above introduced SGA can reach the locally optimal DT-LDCs, which depends on the initial values of $\boldsymbol{\mathcal{X}}$. In our work, we employ the simulated annealing (SAN) algorithm to enhance the likelihood of finding the global optimum. The SAN is an iterative algorithm able to find the global optimum solution, even when the problem is non-concave. Such a general SAN algorithm, however, can be implemented in various ways. One of them is to combine with the SGA framework as in [16]. In particular, the estimate of the gradient is modified to include a random input $\mathbf{w}^{(n)}$ scaled by an SGA coefficient b_n that decays to zeros. Thus, (18) becomes

$$\boldsymbol{\phi}^{(n+1)} = \boldsymbol{\phi}^{(n)} - a_n\hat{\mathbf{g}}(\boldsymbol{\phi}^{(n)}) + b_n\mathbf{w}^{(n)} \tag{23}$$

where $\mathbf{w}^{(n)}$ to be independent, identically distributed (i.i.d.) Gaussian variables. To make the SGA be convergent, a_n and b_n should be properly selected [10]. In our paper, we set $a_n = 1/\log(n+1)$ and $b_n = 0.05/\log(n+1)$.

When we introduce the random perturbation $\mathbf{w}^{(n)}$ in the SGA, the resulting LDC matrix may lose the condition of (11). To avoid such performance degradation, in each iteration, we firstly calculate $\tilde{\boldsymbol{\mathcal{X}}} = \sqrt{2TN_T}\boldsymbol{\mathcal{X}}(\boldsymbol{\mathcal{X}}^T\boldsymbol{\mathcal{X}})^{-1}$, then update $\boldsymbol{\phi}^{(n+1)}$ with $\tilde{\boldsymbol{\mathcal{X}}}$.

IV. SIMULATIONS

In this section, we will present the bit error rate performance of the newly designed DT-LDC matrices with linear MMSE detector in the receivers. Notice that the operating SNR is also a parameter in our designs. We choose the operating SNR as 20dB, and the resulting optimized codes will be proved to be fine for a wide range of SNR of interest.

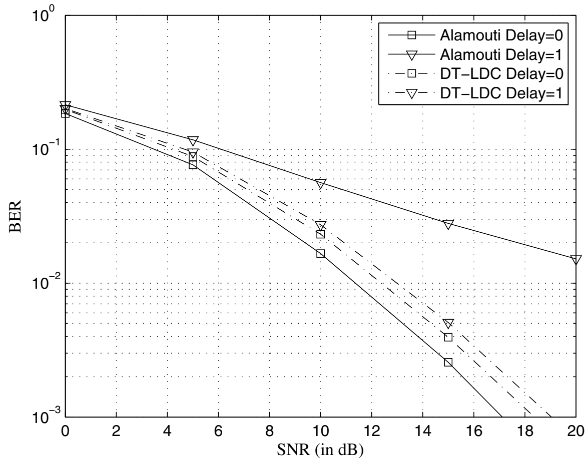

Fig. 1. Performance comparison of DT-LDC with Alamouti Code, $N_T = 2$, $N_R = 1$ $T = 2$ and $Q = 2$.

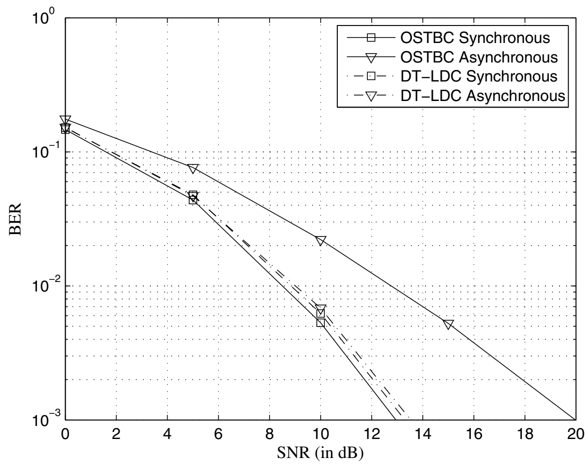

Fig. 2. Performance comparison of DT-LDC with OSTBC. $N_T = 3$, $N_R = 1$, $T = 4$ and $Q = 3$.

Fig. 1 compares the error performance of Alamouti Code with that of our newly designed DT-LDC when QPSK modulation is used. The designed LDC matrix is

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix} 0.7390 & -0.2784 & -0.2996 & -0.6782 \\ -0.2177 & -0.8877 & 0.3270 & -0.0298 \\ 0.1257 & 0.8627 & 0.1227 & -0.1742 \\ 1.0198 & -0.1399 & 0.1224 & 0.0278 \\ -0.0247 & 0.4913 & 0.7766 & -0.5695 \\ 0.5691 & 0.0045 & 0.4599 & 0.6217 \\ 0.0779 & 0.0907 & 0.4358 & 0.7907 \\ 0.1416 & 0.3478 & -0.8768 & 0.4148 \end{bmatrix}.$$

Both schemes use two transmit antennas and one receive antennas and $T = Q = 2$. Notice that Alamouti Code is the only orthogonal space-time code for 2×1 MIMO channels, and the MMSE detector can achieve the optimal performance as the ML detector. When the system is synchronous, the performance of designed DT-LDC is slightly worse than that of Alamouti Code. However, when there is a delay between two nodes, a significant gain of the proposed DT-LDC is observed.

In Fig. 2, we present an example comparing the BER performance of designed DT-LDC with another minimum delay orthogonal space-time block codes (OSTBC) in [1]. In this case, $N_T = 3$, $N_R = 1$, $T = 4$, $Q = 3$ and QPSK modulation is employed. We can see in the synchronous case, the performance of newly designed DT-LDC can approach that of OSTBC. However, when the system is asynchronous, the performance of OSTBC degrades severely while no significant performance degradation is observed with designed DT-LDC.

V. Conclusion

In this paper, we have proposed a novel method to design DT-LDCs for asynchronous cooperative communication networks with linear MMSE receiver. Based on the DT-LDC framework, we derived a design method that can generate LDCs approaching the Jensen's upper bound of ergodic capacity as well as minimizing MSE. The proposed design employed stochastic gradient algorithm to guarantee a local optimum. Moreover, it has been improved by using simulated annealing type of optimization to approach the global optimum. Simulation results confirmed the performance of the new LDCs.

Appendix A
Proof of Proposition 1

For any real-valued matrix $\mathbf{X}$, we have the following rules on derivative of matrices [17].

$$\partial(\mathrm{Tr}(\mathbf{X})) = \mathrm{Tr}(\partial(\mathbf{X})) \tag{24}$$

$$\partial(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}(\partial(\mathbf{X}))\mathbf{X}^{-1} \tag{25}$$

$$\frac{\partial \mathbf{X}^T\mathbf{B}\mathbf{X}}{\partial[\mathbf{X}]_{i,j}} = \mathbf{X}^T\mathbf{B}\mathbf{e}_i\mathbf{e}_j^T + \mathbf{e}_j\mathbf{e}_i^T\mathbf{B}\mathbf{X} \tag{26}$$

Then we can get

$$\frac{\partial \mathrm{Tr}\left(\boldsymbol{\mathcal{X}}^{\mathrm{T}}\mathbf{W}\boldsymbol{\mathcal{X}}+\mathbf{I}\right)^{-1}}{\partial[\boldsymbol{\mathcal{X}}]_{i,j}} = \mathrm{Tr}\left(\frac{\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}+\mathbf{I}\right)^{-1}}{\partial[\boldsymbol{\mathcal{X}}]_{i,j}}\right)$$
$$= -\mathrm{Tr}\left(\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}+\mathbf{I}\right)^{-1}\frac{\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}}{\partial[\boldsymbol{\mathcal{X}}]_{i,j}}\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}+\mathbf{I}\right)^{-1}\right) \tag{27}$$

It is easy to be verified that

$$\mathrm{Tr}\left(\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}\right)^{-1}\mathbf{X}^{T}\mathbf{W}\mathbf{e}_i\mathbf{e}_j^T\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{A}\boldsymbol{\mathcal{X}}\right)^{-1}\right)$$
$$= \mathbf{e}_i^T\mathbf{W}^T\boldsymbol{\mathcal{X}}\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}\right)^{-2}\mathbf{e}_j \tag{28}$$

In a similar way, we have

$$\mathrm{Tr}\left(\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}\right)^{-1}\mathbf{e}_j\mathbf{e}_i^T\mathbf{W}^T\mathbf{X}\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}\right)^{-1}\right)$$
$$= \mathbf{e}_i^T\mathbf{W}^T\boldsymbol{\mathcal{X}}\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{W}\boldsymbol{\mathcal{X}}\right)^{-2}\mathbf{e}_j \tag{29}$$

Substitute (26), (28) and (29) into (27), yields

$$\frac{\partial \mathrm{Tr}\left(\boldsymbol{\mathcal{X}}^{\mathrm{T}}\mathbf{A}\boldsymbol{\mathcal{X}}+\mathbf{I}\right)^{-1}}{\partial[\boldsymbol{\mathcal{X}}]_{i,j}} = -2\mathbf{e}_i^T\mathbf{A}^T\boldsymbol{\mathcal{X}}\left(\boldsymbol{\mathcal{X}}^{T}\mathbf{A}\boldsymbol{\mathcal{X}}+\mathbf{I}\right)^{-2}\mathbf{e}_j \tag{30}$$

Thus we can obtain Proposition 1.

References

[1] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[2] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.

[3] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, 2006.

[4] Y. Li and X.-G. Xia, "A family of distributed space-Time trellis Codes with asynchronous cooperative diversity," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 790–800, 2007.

[5] X. Guo and X.-G. Xia, "Distributed linear convolutive space-time codes for asynchronous cooperative communication networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1857–1861, 2008.

[6] M. O. Damen and A. R. Hammons, "Delay-tolerant distributed-TAST codes for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3755–3773, 2007.

[7] M. Torbatian and M. O. Damen, "On the design of delay-tolerant distributed space-time codes with minimum length," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 931–939, 2009.

[8] M. Sarkiss, R. B. Othman, M. O. Damen, and J. C. Belfiore, "Construction of new delay-tolerant space-time codes," in *Proceedings of International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2010, pp. 379–384.

[9] W. Wang and F.-C. Zheng, "Linear Dispersion Codes Design for Asynchronous Cooperative Communications," in *Proceedings of IEEE PIMRC'11*, Sept. 2011.

[10] J. C. Spall, *Introduction to stochastic search and optimization: estimation, simulation, and control*. John Wiley and Sons, 2003.

[11] F.-C. Zheng, A.-G. Burr, and S. Olafsson, "Near-optimum detection for distributed space-time block coding under imperfect synchronization," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1795–1799, 2008.

[12] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, 2002.

[13] S. Kay, "Fundamentals of statistical signal processing: estimation theory," 1993.

[14] R. W. Heath Jr and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2429–2441, 2002.

[15] X. Wang, V. Krishnamurthy, and J. Wang, "Stochastic gradient algorithms for design of minimum error-rate linear dispersion codes in MIMO wireless systems," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1242–1255, 2006.

[16] H. J. Kushner, "Asymptotic global behavior for stochastic approximation and diffusions with slowly decreasing noise effects: global minimization via Monte Carlo," *SIAM Journal on Applied Mathematics*, vol. 47, no. 1, pp. 169–185, 1987.

[17] K. B. Petersen and M. S. Pedersen, "The matrix cookbook, 2008," *URL http://matrixcookbook. com*.