This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2009 proceedings

A Training-Based Iterative Detection/Channel Estimation Scheme for Large Non-Orthogonal STBC MIMO Systems

Ahmed Zaki, Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan
Department of ECE, Indian Institute of Science, Bangalore 560012, INDIA

***Abstract*—In this paper, we propose a training-based channel estimation scheme for large non-orthogonal space-time block coded (STBC) MIMO systems. The proposed scheme employs a block transmission strategy where an $N_t \times N_t$ pilot matrix is sent (for training purposes) followed by several $N_t \times N_t$ square data STBC matrices, where N_t is the number of transmit antennas. At the receiver, we iterate between channel estimation (using an MMSE estimator) and detection (using a low-complexity likelihood ascent search (LAS) detector) till convergence or for a fixed number of iterations. Our simulation results show that excellent bit error rate and nearness-to-capacity performance are achieved by the proposed scheme at low complexities. The fact that we could show such good results for large STBCs (e.g., 16×16 STBC from cyclic division algebras) operating at spectral efficiencies in excess of 20 bps/Hz (even after accounting for the overheads meant for pilot-based channel estimation and turbo coding) establishes the effectiveness of the proposed scheme.**

I. Introduction

High-rate non-orthogonal space-time block codes (STBC) from cyclic division algebras (CDA) are attractive to achieve high spectral efficiencies along with full diversity [1],[2]. The well known golden code [3] is an example of non-orthogonal STBC from CDA for 2 transmit antennas. Non-orthogonal STBCs from CDA for arbitrary number of transmit antennas is given in [2]. CDA STBCs with large number of antennas are of interest because of the high spectral efficiencies they can offer. An $N_t \times N_t$ CDA STBC matrix with N_t transmit antennas and N_t time slots contains N_t^2 complex data symbols [2]. So, a 16×16 STBC from CDA with 16-QAM and rate-3/4 turbo code can achieve a high spectral efficiency of 48 bps/Hz. A key challenge in exploiting the benefits of large CDA STBCs in practice, however, is *low-complexity detection and channel estimation*. Sphere decoding and its low-complexity variants are prohibitively complex to detect non-orthogonal STBCs with large dimensions (e.g., 32×32 CDA STBC with QAM symbols has 2048 real dimensions). Recently, we solved the problem of low-complexity detection in large non-orthogonal STBC MIMO systems by proposing a powerful Hopfield neural network based *likelihood ascent search (LAS) detector* [4]-[6], which achieved near-maximum likelihood (ML) performance at practically affordable low complexities in large STBC MIMO systems with tens of antennas. In [5],[6], we reported BER results for CDA STBCs of sizes up to 32×32 detected using the LAS detector.

Two assumptions are central to large STBC MIMO detection performance results reported in [5],[6]: i) i.i.d fading, and ii) perfect channel knowledge at the receiver. However, spatial correlation can affect the rank structure of the MIMO channel resulting in degraded capacity. In [7], we relaxed the i.i.d. fading assumption by considering a correlated MIMO channel model. Results in [7] showed that while spatial correlation degraded LAS detector's performance, this performance loss can be alleviated by using more receive dimensions. Large STBC MIMO detection results reported so far have assumed perfect channel state information at the receiver (CSIR). Our new contribution in this paper is that we relax the perfect CSIR assumption and propose a 'training-based iterative detection/channel estimation scheme' for large non-orthogonal STBC MIMO systems.

Training-based schemes, where a pilot signal known to the transmitter and the receiver is sent to get a rough estimate of the channel (training phase) has been studied for STBC MIMO systems in [8]-[11]. A lower bound on the capacity of a MIMO channel that is learned by training is derived in [11] as a function of received SNR, coherence time and number of transmit/receive antennas. It has been shown that the optimum number of channel uses for training is equal to the number of transmit antennas [11]. Blind, semi-blind, and training-based schemes are compared in [12]. Here, we propose a training-based scheme that employs a block transmission strategy where a pilot matrix is sent (for training purposes) followed by several data STBC matrices. At the receiver, we iterate between channel estimation (using an MMSE estimator) and detection (using LAS detector) till convergence or for a fixed number of iterations. We show that the proposed scheme achieves very good uncoded/coded BER performance at high spectral efficiencies in excess of 20 bps/Hz. For e.g., in slow fading with large coherence times (as will be experienced in fixed/low-mobility wireless applications) the proposed scheme using 16×16 STBC from CDA with 4-QAM and rate-3/4 turbo code achieves a spectral efficiency of 23.5 bps/Hz using 'estimated CSIR,' which amounts to only a small loss in spectral efficiency compared to that achieved with perfect CSIR (24 bps/Hz). With the feasibility of such low-complexity schemes, detection and channel estimation need not be a bottleneck in realizing large MIMO systems.

II. Non-Orthogonal STBC MIMO System Model

Consider a STBC MIMO system with multiple transmit and multiple receive antennas. An (n, p, k) STBC is represented by a matrix $\mathbf{X}_c \in \mathbb{C}^{n \times p}$, where n and p denote the number of transmit antennas and number of time slots, respectively, and k denotes the number of complex data symbols sent in one STBC matrix. The (i, j)th entry in $\mathbf{X}_c$ represents the complex number transmitted from the ith transmit antenna in the jth time slot. The rate of an STBC is given by $\frac{k}{p}$. Let N_r and $N_t = n$ denote the number of receive and transmit antennas, respectively. Let $\mathbf{H}_c \in \mathbb{C}^{N_r \times N_t}$ denote the channel gain matrix, where the (i, j)th entry in $\mathbf{H}_c$ is the complex channel gain from the jth transmit antenna to the ith receive antenna. We assume that the channel gains remain constant over one STBC matrix duration. Assuming rich scattering, we model the entries of $\mathbf{H}_c$ as i.i.d $\mathcal{CN}(0,1)$. The received space-time signal matrix, $\mathbf{Y}_c \in \mathbb{C}^{N_r \times p}$, can be written as

$$\mathbf{Y}_c \quad = \quad \mathbf{H}_c \mathbf{X}_c + \mathbf{N}_c, \tag{1}$$

where $\mathbf{N}_c \in \mathbb{C}^{N_r \times p}$ is the noise matrix at the receiver and its entries are modeled as i.i.d $\mathcal{CN}\big(0, \sigma^2 = \frac{N_t E_s}{\gamma}\big)$, where E_s is the average energy of the transmitted symbols, γ is the average received SNR per receive antenna, and the (i,j)th entry in $\mathbf{Y}_c$ is the received signal at the ith receive antenna in the jth time slot. Consider linear dispersion STBCs, where $\mathbf{X}_c$ can be written in the form

$$\mathbf{X}_c \quad = \quad \sum_{i=1}^{k} x_c^{(i)} \mathbf{A}_c^{(i)}, \tag{2}$$

where $x_c^{(i)}$ is the ith complex data symbol, and $\mathbf{A}_c^{(i)} \in \mathbb{C}^{N_t \times p}$ is its weight matrix. The received signal model in (1) can be written in an equivalent V-BLAST form as

$$\mathbf{y}_c \quad = \quad \sum_{i=1}^{k} x_c^{(i)} \, (\widehat{\mathbf{H}}_c \, \mathbf{a}_c^{(i)}) + \mathbf{n}_c \quad = \quad \widetilde{\mathbf{H}}_c \mathbf{x}_c + \mathbf{n}_c, \tag{3}$$

where $\mathbf{y}_c \in \mathbb{C}^{N_r p \times 1} = vec\,(\mathbf{Y}_c)$, $\widehat{\mathbf{H}}_c \in \mathbb{C}^{N_r p \times N_t p} = (\mathbf{I} \otimes \mathbf{H}_c)$, $\mathbf{a}_c^{(i)} \in \mathbb{C}^{N_t p \times 1} = vec\,(\mathbf{A}_c^{(i)})$, $\mathbf{n}_c \in \mathbb{C}^{N_r p \times 1} = vec\,(\mathbf{N}_c)$, $\mathbf{x}_c \in \mathbb{C}^{k \times 1}$ whose ith entry is the data symbol $x_c^{(i)}$, and $\widetilde{\mathbf{H}}_c \in \mathbb{C}^{N_r p \times k}$ whose ith column is $\widehat{\mathbf{H}}_c \, \mathbf{a}_c^{(i)}$, $i = 1, 2, \cdots, k$. Each element of $\mathbf{x}_c$ is an M-PAM or M-QAM symbol. Let $\mathbf{y}_c$, $\widetilde{\mathbf{H}}_c$, $\mathbf{x}_c$, and $\mathbf{n}_c$ be decomposed into real and imaginary parts as

$$\begin{aligned} \mathbf{y}_c &= \mathbf{y}_I + j\mathbf{y}_Q, & \mathbf{x}_c &= \mathbf{x}_I + j\mathbf{x}_Q, \\ \mathbf{n}_c &= \mathbf{n}_I + j\mathbf{n}_Q, & \widetilde{\mathbf{H}}_c &= \widetilde{\mathbf{H}}_I + j\widetilde{\mathbf{H}}_Q. \end{aligned} \tag{4}$$

Further, we define $\mathbf{x}_r \in \mathbb{R}^{2k \times 1}$, $\mathbf{y}_r \in \mathbb{R}^{2N_r p \times 1}$, $\mathbf{H}_r \in \mathbb{R}^{2N_r p \times 2k}$, and $\mathbf{n}_r \in \mathbb{R}^{2N_r p \times 1}$ as

$$\begin{aligned} \mathbf{x}_r &= [\mathbf{x}_I^T \ \ \mathbf{x}_Q^T]^T, & \mathbf{y}_r &= [\mathbf{y}_I^T \ \ \mathbf{y}_Q^T]^T, \\ \mathbf{H}_r &= \begin{pmatrix} \widetilde{\mathbf{H}}_I & -\widetilde{\mathbf{H}}_Q \\ \widetilde{\mathbf{H}}_Q & \widetilde{\mathbf{H}}_I \end{pmatrix}, & \mathbf{n}_r &= [\mathbf{n}_I^T \ \ \mathbf{n}_Q^T]^T. \end{aligned} \tag{5}$$

Now, (3) can be written as

$$\mathbf{y}_r \quad = \quad \mathbf{H}_r \mathbf{x}_r + \mathbf{n}_r. \tag{6}$$

Henceforth, we work with the real-valued system in (6). For notational simplicity, we drop subscripts r in (6) and write

$$\mathbf{y} \quad = \quad \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{7}$$

where $\mathbf{H} = \mathbf{H}_r \in \mathbb{R}^{2N_r p \times 2k}$, $\mathbf{y} = \mathbf{y}_r \in \mathbb{R}^{2N_r p \times 1}$, $\mathbf{x} = \mathbf{x}_r \in \mathbb{R}^{2k \times 1}$, and $\mathbf{n} = \mathbf{n}_r \in \mathbb{R}^{2N_r p \times 1}$. The channel coefficients are assumed to be known at the receiver but not at the transmitter.

A. High-Rate Non-Orthogonal STBCs from CDA

We consider square (i.e., $n{=}p{=}N_t$), full-rate (i.e., $k{=}pn{=}N_t^2$), circulant (where the weight matrices $\mathbf{A}_c^{(i)}$'s are permutation type), non-orthogonal STBCs from CDA [2], whose construction for arbitrary number of transmit antennas n is given by the matrix in (7.a) given at the bottom of the next page [2]. In (7.a), $\omega_n = e^{\frac{\mathbf{j}2\pi}{n}}$, $\mathbf{j} = \sqrt{-1}$, and $x_{u,v}$, $0 \leq u, v \leq n-1$ are the data symbols from a QAM alphabet. When $\delta = t = 1$, the code in (7.a) is information lossless, and when $\delta = e^{\sqrt{5}\mathbf{j}}$ and $t = e^{\mathbf{j}}$, it is full-diversity and information lossless [2]. High spectral efficiencies with large N_t can be achieved using this code construction; e.g., 16×16 STBC from (7.a) using 16-QAM and rate-3/4 turbo code offers a spectral efficiency of 48 bps/Hz along with the full-diversity of order $N_t N_r$ under ML detection. However, since these STBCs are non-orthogonal, ML detection gets increasingly impractical for large N_t. Hence, a key challenge in realizing the benefits of these large STBCs in practice is that of achieving near-ML performance for large N_t at low detection complexities; the LAS detector in [6] has been shown to essentially achieve this. The detection part in this paper is carried out using the LAS algorithm. For details of the LAS algorithm for large MIMO detection, please refer to [6],[13].

III. Iterative Detection/Channel Estimation

In this section, we propose to estimate the channel matrix based on a training-based iterative detection/channel estimation scheme. In the proposed scheme, transmission is carried out in frames, where one $N_t \times N_t$ pilot matrix (for training purposes) followed by N_d data STBC matrices are sent in each frame as shown in Fig. 1. One frame length, T, (taken to be the channel coherence time) is $T = (N_d+1)N_t$ channel uses. A frame of transmitted pilot and data matrices is of dimension $N_t \times N_t(1+N_d)$, which can be written as

$$\mathcal{X}_c \quad = \quad \Big[\mathbf{X}_c^{(\mathrm{P})} \, \mathbf{X}_c^{(1)} \, \mathbf{X}_c^{(2)} \cdots \mathbf{X}_c^{(N_d)} \Big]. \tag{8}$$

As in [11], let γ_p and γ_d denote the average SNR during pilot and data phases, respectively, which are related to the average received SNR γ as $\gamma(N_d+1) = \gamma_p + N_d\gamma_d$. Define $\beta_p \triangleq \frac{\gamma_p}{\gamma}$, and $\beta_d \triangleq \frac{\gamma_d}{\gamma}$. Let E_s denote the average energy of the transmitted symbol during the data phase. The average received signal power during the data phase is given by $\mathbb{E}\big[\mathrm{tr}\big(\mathbf{X}_c^{(\mathrm{i})} {\mathbf{X}_c^{(\mathrm{i})}}^H\big)\big] = N_t^2 E_s$, and the average received signal power during the pilot phase is $\mathbb{E}\big[\mathrm{tr}\big(\mathbf{X}_c^{(\mathrm{P})} {\mathbf{X}_c^{(\mathrm{P})}}^H\big)\big] = \frac{N_t^2 E_s \beta_p}{\beta_d} = \mu N_t$, where $\mu \triangleq \frac{N_t E_s \beta_p}{\beta_d}$. For optimal training, the pilot matrix should be such that $\mathbf{X}_c^{(\mathrm{P})} {\mathbf{X}_c^{(\mathrm{P})}}^H = \mu \mathbf{I}_{N_t}$ [11].

$$\begin{bmatrix}
\sum_{i=0}^{n-1} x_{0,i}\, t^i & \delta \sum_{i=0}^{n-1} x_{n-1,i}\, \omega_n^i\, t^i & \delta \sum_{i=0}^{n-1} x_{n-2,i}\, \omega_n^{2i}\, t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{1,i}\, \omega_n^{(n-1)i}\, t^i \\
\sum_{i=0}^{n-1} x_{1,i}\, t^i & \sum_{i=0}^{n-1} x_{0,i}\, \omega_n^i\, t^i & \delta \sum_{i=0}^{n-1} x_{n-1,i}\, \omega_n^{2i}\, t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{2,i}\, \omega_n^{(n-1)i}\, t^i \\
\sum_{i=0}^{n-1} x_{2,i}\, t^i & \sum_{i=0}^{n-1} x_{1,i}\, \omega_n^i\, t^i & \sum_{i=0}^{n-1} x_{0,i}\, \omega_n^{2i}\, t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{3,i}\, \omega_n^{(n-1)i}\, t^i \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\sum_{i=0}^{n-1} x_{n-2,i}\, t^i & \sum_{i=0}^{n-1} x_{n-3,i}\, \omega_n^i\, t^i & \sum_{i=0}^{n-1} x_{n-4,i}\, \omega_n^{2i}\, t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{n-1,i}\, \omega_n^{(n-1)i} t^i \\
\sum_{i=0}^{n-1} x_{n-1,i}\, t^i & \sum_{i=0}^{n-1} x_{n-2,i}\, \omega_n^i\, t^i & \sum_{i=0}^{n-1} x_{n-3,i}\, \omega_n^{2i}\, t^i & \cdots & \sum_{i=0}^{n-1} x_{0,i}\, \omega_n^{(n-1)i}\, t^i
\end{bmatrix}. \tag{7.a}$$

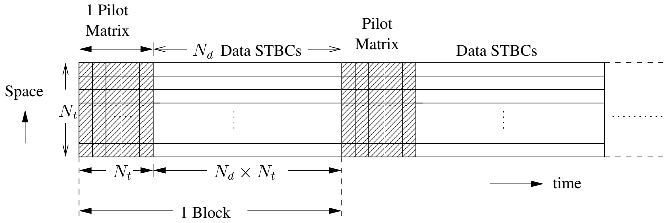

Fig. 1. Transmission scheme with one pilot matrix followed by N_d data STBC matrices in each frame.

We want to estimate the channel matrix, $\mathbf{H}_c \in \mathbb{C}^{N_r \times N_t}$. We assume block fading, where the channel gains remain constant over one frame consisting of $(1 + N_d)N_t$ channel uses, which can be viewed as the channel coherence time. This assumption can be valid in slow fading fixed wireless applications (e.g., as in possible applications like BS-to-BS backbone connectivity and BS-to-CPE wireless IPTV/HDTV distribution). For this training-based system and channel model, Hassibi and Hochwald presented a lower bound on the capacity in [11]; we will illustrate the nearness of the performance achieved by the proposed scheme to this bound. The received frame is of dimension $N_r \times N_t(1 + N_d)$, and can be written as

$$\mathcal{Y}_c = \left[\mathbf{Y}_c^{(\mathrm{P})}\, \mathbf{Y}_c^{(1)}\, \mathbf{Y}_c^{(2)} \cdots \mathbf{Y}_c^{(N_d)}\right] = \mathbf{H}_c \mathcal{X}_c + \mathcal{N}_c\,, \quad (9)$$

where $\mathcal{N}_c = \left[\mathbf{N}_c^{(\mathrm{P})}\, \mathbf{N}_c^{(1)}\, \mathbf{N}_c^{(2)} \cdots \mathbf{N}_c^{(N_d)}\right]$ is the $N_r \times N_t(1+N_d)$ noise matrix and its entries are modeled as i.i.d. $\mathcal{CN}(0, \sigma^2 = \frac{N_t E_s}{\gamma \beta_d})$, where γ is the average SNR per receive antenna. Equation (9) can be decomposed into two parts, namely, the pilot matrix part and the data matrices part, as

$$\mathbf{Y}_c^{(\mathrm{P})} = \mathbf{H}_c \mathbf{X}_c^{(\mathrm{P})} + \mathbf{N}_c^{(\mathrm{P})}, \quad (10)$$

and

$$\begin{aligned}\mathbf{Y}_c^{(\mathrm{D})} &= \left[\mathbf{Y}_c^{(1)}\, \mathbf{Y}_c^{(2)} \cdots \mathbf{Y}_c^{(N_d)}\right] \\ &= \mathbf{H}_c \left[\mathbf{X}_c^{(1)}\, \mathbf{X}_c^{(2)} \cdots \mathbf{X}_c^{(N_d)}\right] + \left[\mathbf{N}_c^{(1)}\, \mathbf{N}_c^{(2)} \cdots \mathbf{N}_c^{(N_d)}\right]. \quad (11)\end{aligned}$$

A. MMSE Estimation Scheme

A straight-forward way to achieve detection of data symbols with estimated channel coefficients is as follows:

1) Estimate the channel gains via an *MMSE estimator* from the signal received during the first N_t channel uses (i.e., during pilot transmission); i.e., given $\mathbf{Y}_c^{(\mathrm{P})}$ and $\mathbf{X}_c^{(\mathrm{P})}$, an estimate of the channel matrix $\mathbf{H}_c$ is found as

$$\mathbf{H}_c^{est} = \mathbf{Y}_c^{(\mathrm{P})} (\mathbf{X}_c^{(\mathrm{P})})^H \left[\sigma^2 \mathbf{I}_{N_t} + \mathbf{X}_c^{(\mathrm{P})} (\mathbf{X}_c^{(\mathrm{P})})^H\right]^{-1}. \quad (12)$$

2) Use the above $\mathbf{H}_c^{est}$ in place of $\mathbf{H}_c$ in the LAS algorithm (reported in [6],[13]) and detect the transmitted data symbols.

B. Proposed Iterative Detection/Estimation Scheme

Techniques that employ iterations between channel estimation and detection can offer improved performance. Here, we propose an *'iterative detection/estimation scheme,'* which works as follows:

1) Obtain an initial estimate of the channel matrix using the MMSE estimator in (12) from the pilot part.

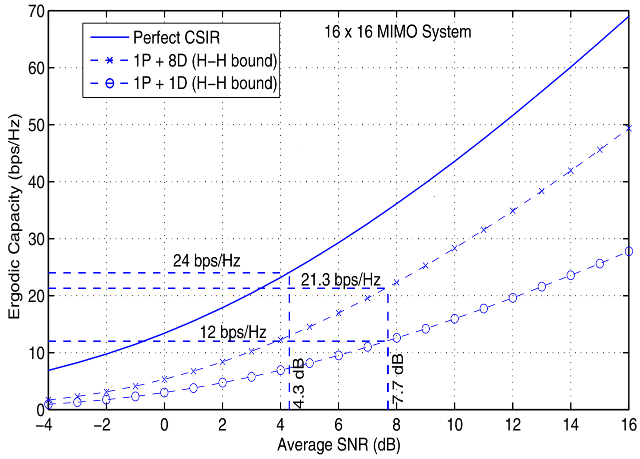

Fig. 2. Hassibi-Hochwald (H-H) capacity bound for 1P+8D ($T = 144, \tau = 16, \beta_p = \beta_d = 1$) and 1P+1D ($T = 32, \tau = 16, \beta_p = \beta_d = 1$) training for a 16×16 MIMO channel. Perfect CSIR capacity is also shown.

2) Using the estimated channel matrix, detect the data STBC matrices $\mathbf{X}_c^{(i)}$, $i = 1, 2, \cdots, N_d$ using the LAS detector. Substituting these detected STBC matrices into (8), form $\mathcal{X}_c^{est}$.
3) Re-estimate the channel matrix using $\mathcal{X}_c^{est}$ from the previous step, via

$$\mathbf{H}_c^{est} = \mathcal{Y}_c (\mathcal{X}_c^{est})^H \left[\sigma^2 \mathbf{I}_{N_t} + \mathcal{X}_c^{est} (\mathcal{X}_c^{est})^H\right]^{-1}. \quad (13)$$

4) Iterate steps 2 and 3 until convergence or for a specified number of iterations.

The complexity of obtaining the MMSE estimates in (12) and (13) is less than the LAS detection complexity. Since the number of detection/estimation iterations is typically small (our simulations showed that the performance gain saturates beyond 4 iterations), the overall order of complexity remains same as that of the LAS algorithm with MMSE initial vector.

C. BER Performance with Estimated CSIR

We evaluated the BER performance of the LAS detector using estimated CSIR, where we estimate the channel gain matrix through the training-based estimation schemes described in the previous two subsections. We consider the BER performance under three scenarios, namely, i) under perfect CSIR, ii) under CSIR estimated using the MMSE estimation scheme in Sec. III-A, and iii) under CSIR estimated using the iterative detection/estimation scheme in Sec. III-B. In the case of estimated CSIR, we show plots for 1P+N_dD training, where by 1P+N_dD training we mean a training scheme with a frame of size $1 + N_d$ matrices with 1 pilot matrix followed N_d data STBC matrices from CDA. For this 1P+N_dD training scheme, a lower bound on the capacity is given by [11]

$$C \geq \frac{T-\tau}{T}\, \mathbb{E}\left[\operatorname{logdet}\left(\mathbf{I}_{N_t} + \frac{\gamma^2 \beta_d \beta_p \tau}{N_t(1+\gamma\beta_d) + \gamma\beta_p\tau} \frac{\mathbf{H}_c^{'} \mathbf{H}_c^{'H}}{N_t \sigma^2_{\mathbf{H}_c^{'}}}\right)\right], \quad (14)$$

where T and τ, respectively, are the frame size (i.e., channel coherence time) and pilot duration in number of channel uses, and $\sigma^2_{\mathbf{H}_c^{'}} = \frac{1}{N_t N_r}\, \mathbb{E}\left[tr\{\mathbf{H}_c^{'} \mathbf{H}_c^{'H}\}\right]$ where $\mathbf{H}_c^{'} = \mathbb{E}\left[\mathbf{H}_c \,\middle|\, \mathbf{X}_c^{(\mathrm{P})}, \mathbf{Y}_c^{(\mathrm{P})}\right]$ is the MMSE estimate of the channel gain matrix, given by (12). We computed the capacity bound in (14) through simulations for 1P+8D and 1P+1D training for a 16×16 MIMO channel. For 1P+8D training $T = (1+8)16 =$

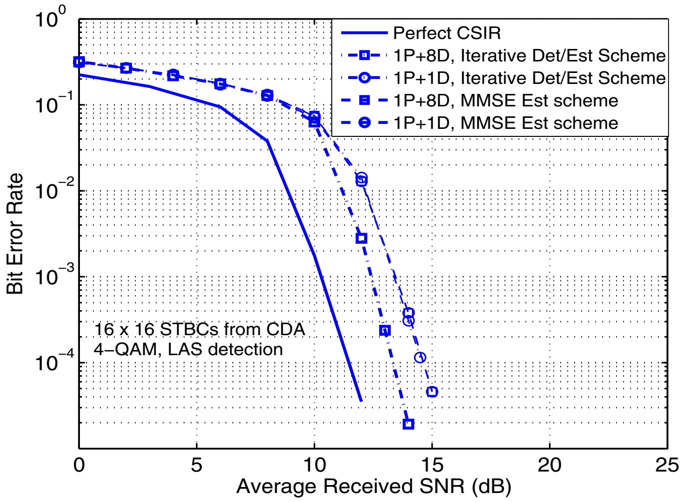

Fig. 3. Uncoded BER of LAS detector for 16×16 STBC with $i)$ perfect CSIR, $ii)$ CSIR using MMSE estimation scheme, and $iii)$ CSIR using iterative detection/channel estimation scheme (4 iterations). $N_t = N_r = 16$, 4-QAM, 1P+1D $(T = 32, \tau = 16, \beta_p = \beta_d = 1)$ and 1P+8D $(T =$

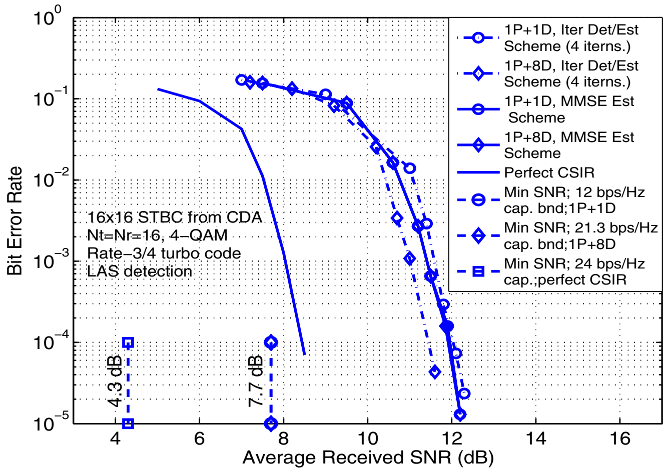

Fig. 4. Turbo coded BER performance of LAS detector for 16×16 STBC with $i)$ perfect CSIR, $ii)$ CSIR using MMSE estimation, and $iii)$ CSIR using iterative detection/channel estimation (4 iterations). $N_t = N_r = 16$, 4-QAM, rate-3/4 turbo code, 1P+1D $(T = 32, \tau = 16, \beta_p = \beta_d = 1)$ and 1P+8D $(T = 144, \tau = 16, \beta_p = \beta_d = 1)$ training.

144, $\tau = 16$, and for 1P+1D training $T = (1+1)16 = 32$, $\tau = 16$. In computing the bounds (shown in Fig. 2) and in BER simulations (in Figs. 3, 4), we have used $\beta_p = \beta_d = 1$.

In Fig. 2, we plot the computed capacity bounds, along with the capacity under perfect CSIR. We obtain the minimum SNR for a given capacity bound in (14) from the plots in Fig. 2, and show (later in Fig. 4) the nearness of the coded BER of the proposed scheme to this SNR limit. We note that improved capacity and BER performance can be achieved if optimum pilot/data power allocation derived in [11] is used instead of the allocation used in Figs. 2 to 4 (i.e., $\beta_p = \beta_d = 1$). We have used the optimum power allocation in [11] for generating the BER plots in Figs. 5 and 6. In all the BER simulations with training, $\sqrt{\mu}\,\mathbf{I}_{N_t}$ is used as the pilot matrix.

First, in Fig. 3, we plot the uncoded BER performance of LAS detector when 1P+1D and 1P+8D training are used for channel estimation in a 16×16 STBC (with $\delta = t = 1$) MIMO system with $N_t = N_r = 16$ and 4-QAM. BER performance with perfect CSIR is also plotted for comparison.

From Fig. 3, it can be observed that, as expected, the BER degrades with estimated CSIR compared to that with perfect CSIR. With MMSE estimation scheme, the performance with 1P+1D and 1P+8D are same because of the one-shot nature of the MMSE estimation. Also, with 1P+1D training, both the MMSE estimation scheme as well as the iterative detection/estimation scheme (with 4 iterations between detection and estimation) perform almost the same, which is about 3 dB worse compared to that of perfect CSIR at an uncoded BER of 10^{-3}. This indicates that with 1P+N_dD training, iteration between detection and estimation does not improve performance much over the non-iterative scheme (i.e., the MMSE estimation scheme) for small N_d. With large N_d (e.g., slow fading), however, the iterative scheme outperforms the non-iterative scheme; e.g., with 1P+8D training, the performance of the iterative detection/estimation improves by about 1 dB compared to the MMSE estimation.

Next, in Fig. 4, we present the rate-3/4 turbo coded BER of LAS detector using estimated CSIR for the cases of 1P+8D and 1P+1D training. From Fig. 4, it can be seen that, compared to that of perfect CSIR, the estimated CSIR performance is worse by about 3 dB in terms of coded BER for 1P+8D training. With MMSE estimation scheme, 10^{-4} coded BER occurs at about $12 - 7.7 = 4.3$ dB away from the capacity bound for 1P+1D and 1P+8D training. This nearness to capacity bound improves by about 0.6 dB for the iterative detection/estimation scheme. We note that for the system in Fig. 4 with parameters 16×16 STBC, 4-QAM, rate-3/4 turbo code, and 1P+8D training with $T = 144, \tau = 16$, we achieve a high spectral efficiency of $16 \times 2 \times \frac{3}{4} \times \frac{8}{9} = 21.3$ bps/Hz even after accounting for the overheads involved in channel estimation (i.e., pilot matrix) and channel coding, while achieving good near-capacity performance at low complexity. This points to the suitability of the proposed approach of using LAS detection along with iterative detection/estimation in practical implementation of large STBC MIMO systems.

Finally, in Fig. 5, we illustrate the coded BER performance of LAS detection and iterative detection/estimation scheme for different coherence times, T, for a fixed $N_t = N_r = 16$, 16×16 STBC, 4-QAM, and rate-3/4 turbo code. The various values of T considered and the corresponding spectral efficiencies are: $i)$ $T = 32$, 1P+1D, 12 bps/Hz, $ii)$ $T = 144$, 1P+8D, 21.3 bps/Hz, $iii)$ $T = 400$, 1P+24D, 23.1 bps/Hz, and $iv)$ $T = 784$, 1P+48D, 23.5 bps/Hz. In all these cases, optimum pilot/data power allocation in [11] is used. From Fig. 5, it can be seen that for these four cases, 10^{-4} coded BER occurs at around 12 dB, 10.6 dB, 9.7 dB, and 9.4 dB, respectively. The 10^{-4} coded BER for perfect CSIR happens at around 8.5 dB. This indicates that the performance with estimated CSIR improves as T is increased, and that performance loss less than 1 dB compared to perfect CSIR can be achieved with large T (i.e., slow fading). For example, with 1P+48D training ($T = 784$), the performance with estimated CSIR gets close to that with perfect CSIR both in terms of spectral efficiency (23.5 vs 24 bps/Hz) as well as SNR at which 10^{-4} coded BER occurs (8.5 vs 9.4 dB). This is expected, since the channel estimation becomes increasingly accurate in slow fading (large coherent times) while incurring only

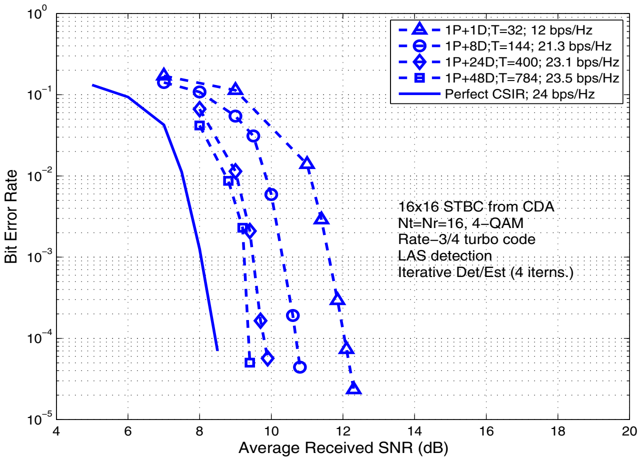

Fig. 5. Turbo coded BER performance of LAS detection and iterative estimation/detection as a function of coherence time, $T = 32, 144, 400, 784$, for a given $N_t = N_r = 16$. 16×16 STBC, 4-QAM, rate-3/4 turbo code. *Spectral efficiency and BER performance with estimated CSIR approaches to those with perfect CSIR in slow fading (i.e., large T).*

a small loss in spectral efficiency due to pilot matrix overhead. This result is significant because T is typically large in fixed/low-mobility wireless applications, and the proposed system can effectively achieve high spectral efficiencies as well as good performance in such applications.

D. On Optimum N_t for a Given N_r and T

In [11], through theoretical capacity bounds it has been shown that, for a given N_r, T and SNR, there is an optimum value of N_t that maximizes the capacity bound (refer Figs. 5 and 6 in [11], where the optimum N_t is shown to be greater than N_r in Fig. 5 and less than N_r in Fig. 6). For example, for $N_r = 16$, $T = 48$, and SNR = 10 dB, the capacity bound evaluated using (14) with optimum power allocation for $N_t = 12$ is 19.73 bps/Hz, whereas for $N_t = 16$ the capacity bound reduces to 17.53 bps/Hz showing that the optimum N_t in this case will be less than N_r. We demonstrate this phenomenon in practical systems by comparing the simulated coded BER performance of two systems, referred to as System-I and System-II, using LAS detection and iterative detection/estimation scheme. The parameters of System-I and System-II are as follows: N_r and T are fixed at 16 and 48, respectively, in both systems. System-I uses 16 transmit antennas and 16×16 STBC, whereas System-II uses 12 transmit antennas and 12×12 STBC. Since the pilot matrix is $\sqrt{\mu}\,\mathbf{I}_{N_t}$, the pilot duration τ is 16 and 12, respectively, for System-I and System-II. Optimum pilot/data power allocation and 4-QAM modulation are employed in both systems. System-I uses rate-1/2 turbo code and system-II uses rate-3/4 turbo code. With the above system parameters, the spectral efficiency achieved in System-I is $16 \times 2 \times \frac{1}{2} \times \frac{2}{3} = 10.33$ bps/Hz, whereas System-II achieves a higher spectral efficiency of $12 \times 2 \times \frac{3}{4} \times \frac{3}{4} = 13.5$ bps/Hz. In Fig. 6, we plot the coded BER of both these systems using LAS detection and iterative detection/estimation. From Fig. 6, it can be observed that System-II with a smaller N_t and higher spectral efficiency in fact achieves 10^{-4} coded BER at a lesser SNR than System-I (8.9 dB in Sys-II vs 9.3 dB in Sys-I). This means that because of the training overheads, a larger N_t does not necessarily mean a higher spectral efficiency; this corroborates the theoretical prediction in [11].

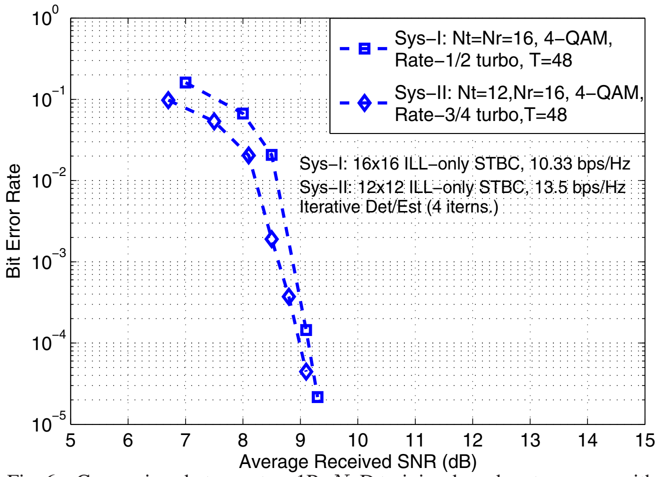

Fig. 6. Comparison between two 1P+N_dD training-based systems, one with a larger N_t than the other for a given N_r and T. *With $N_r = 16$, $T = 48$ and optimum power allocation in both systems, System-II with $N_t = 12$ achieves a higher spectral efficiency (13.5 vs 10.33 bps/Hz) while achieving 10^{-3} coded BER at a lesser SNR (8.6 vs 8.9 dB) than System-I with $N_t = 16$.*

IV. Conclusion

We presented the simulated BER performance of low-complexity detection of large non-orthogonal STBC MIMO systems *in conjunction with a channel estimation scheme*. The very good BER and nearness-to-capacity results reported here at practically affordable low complexities point to a key observation that detection and channel estimation need not be a bottleneck in realizing large STBC MIMO systems. Wireless standards (e.g., IEEE 802.11n/VHT, 802.16) can potentially consider such large MIMO systems as options for the future.

References

[1] F. Oggier, J.-C. Belfiore, and E. Viterbo, *Cyclic Division Algebras: A Tool for Space-Time Coding,* Foundations and Trends in Commun. and Inform. Theory, vol. 4, no. 1, pp. 1-95, Now Publishers, 2007.

[2] B. A. Sethuraman, B. Sundar Rajan, and V. Shashidhar, "Full-diversity high-rate space-time block codes from division algebras," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2596-2616, October 2003.

[3] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: A 2 × 2 full-rate space-time code with non-vanishing determinants," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, April 2005.

[4] K. Vishnu Vardhan, Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *IEEE JSAC Spl. Iss. on Multiuser Detection, for Adv. Commun. Systems and Networks*, vol. 26, no. 3, pp. 473-485, April 2008. Online arXiv:0804.0980v1 [cs.IT] 7 Apr 2008.

[5] Saif K. Mohammed, K. Vishnu Vardhan, A. Chockalingam, B. Sundar Rajan, "Large MIMO systems: A low-complexity detector at high spectral efficiencies," *Proc. IEEE ICC'2008*, Beijing, May 2008.

[6] Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan, "High-rate space-time coded large MIMO systems: Low-complexity detection and performance," *Proc. IEEE GLOBECOM'2008*, Dec. 2008.

[7] Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan, "Low-complexity detection and performance in multi-gigabit high spectral efficiency wireless systems," *Proc. IEEE PIMRC'2008, Sept. 2008.*.

[8] M. Brehler, M. K. Varanasi, "Training-codes for non-coherent multi-antenna block-Rayleigh fading channel," *Proc. CISS'03*, March 2003.

[9] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1097-1119, May 2003.

[10] H. El Gamal, H. Aktas, and M. O. Damen, "Coherent space-time codes for noncoherent channels," *Proc. IEEE GLOBECOM'03*, Dec. 2003.

[11] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory,* vol. 49, no. 4, pp. 951-963, April 2003.

[12] P. Stoica and G. Ganesan, "Space-time block codes: trained, blind and semi-blind detection," *Proc. IEEE ICASSP'02*, pp. 1609-1612, 2002.

[13] Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan, "A low-complexity near-ML performance achieving algorithm for large MIMO detection," *Proc. IEEE ISIT'2008*, Toronto, July 2008.