TRAINED SPACE-TIME BLOCK DECODING FOR FADING CHANNELS WITH FREQUENCY OFFSETS

Petre Stoica and Girish Ganesan

Department of Systems and Control,
Information Technology, Uppsala University,
P. O. Box 27, SE-751 03, Uppsala, Sweden.

ABSTRACT

Space-time block coding (STBC) is a recent appealing solution to the problem of exploiting transmit diversity in multi-antenna systems for communications over flat fading channels. In a standard STBC scheme the receiver requires Channel State Information (CSI), which can be acquired via training at the expense of a reduced information rate. Alternatively, the requirement of CSI can be avoided altogether by using differential encoding. The existing trained or differential schemes for STBC assume that the channel is time-invariant during the transmission of at least two data blocks. However, wireless channels may be often time varying owing to frequency offsets induced by either Doppler shifts or carrier frequency mismatches. In this paper we present a simple trained STBC scheme for fading channels with frequency offsets.

1. INTRODUCTION, PRELIMINARIES AND PROBLEM STATEMENT

1.1. Channel Model

We consider a wireless communication system with m receive and n transmit antennas. Let $t = 1, 2, \ldots$ be the discrete-time index expressed in units of the sampling interval. Also, let A_{kp} denote the fading coefficient from the p-th transmit antenna to the k-th receive antenna, and let ω_k be the (angular) frequency offset between the k-th receive antenna and the transmit antennas. We assume that ω_k is the same (or nearly so) for all transmit antennas. This should be true for the part of ω_k induced by a possible carrier frequency mismatch between the transmit and receive oscillators. The previous assumption should also be valid for the Doppler shift part of ω_k provided that the multipath components arriving at the receive antennas have similar angles of arrival [7]. Making use of the notation and assumptions introduced above, we can model the output of the receive array as (see [7] for a detailed derivation of the equation below):

$$y(t) = \begin{bmatrix} e^{i\omega_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{i\omega_m t} \end{bmatrix} A\, z(t) + e(t) \quad (1)$$

where $z(t)$ is the transmitted (baseband) signal, $A = \{A_{kp}\}$, and $e(t)$ denotes a noise term. We assume that $\{e_k(t);$ for $k = 1, \ldots m; t = 1, 2, \ldots\}$ is a sequence of i.i.d. Gaussian random variables with mean zero and common variance σ^2.

1.2. Brief Review of STBC

Let $\{s_k\}_{k=1}^{P}$ denote the symbols that we want to transmit in the block b (we omit the dependence of s_k on b to simplify notation). In a STBC scheme this is done by transmitting a space-time block (for $b = 0, 1, 2, \ldots$):

$$Z_b = [z(Nb+1) \ldots z(Nb+N)] \quad (n \times N) \quad (2)$$

that depends on $\{s_k\}_{k=1}^{P}$. Following [3, 10] we let Z_b depend linearly on $\{s_k\}$:

$$Z_b = \sum_{k=1}^{P}[B_k\, Re(s_k) + i\, C_k\, Im(s_k)] \quad (3)$$

The matrices $\{B_k, C_k\}$ in (3) are chosen such that ([2, 3, 4, 5, 10])

$$Z_b Z_b^* = \left(\sum_{k=1}^{P} |s_k|^2\right) I \quad (4)$$

where the superscript $*$ denotes the conjugate transpose. The property in (4) of STBC plays a key role in the simplification of the maximum likelihood detector (MLD) for the transmitted symbols (see [3, 10] for details).

However the MLD relies on the assumption that the receiver knows A and $\{\omega_k\}$ or at least can estimate them accurately. This aspect is discussed next.

1.3. Problem Statement and Discussion of Existing Blind and Differential Schemes

Under the assumption that $\omega_k = 0$ ($k = 1, \ldots, m$), the channel fading matrix A can be estimated using one or more pilot blocks. A can also be estimated blindly (using only one pilot symbol) or semi-blindly (combining the trained and blind schemes), see the recent paper [8]. The main goal of the present paper is to introduce a simple training-based scheme for the estimation of *both A and ω_k*, which should be the right thing to do in those applications where the frequency offsets cannot be ignored. Owing to the fact that usually transmit diversity is employed when the receiver has rather limited computational capabilities, *we aim to keep the complexity of our trained STBC scheme as low as possible.*

We end this section with a brief discussion of differential encoding as an alternative to acquiring CSI at the receiver via training (in this case the acquired CSI is used in the MLD of a standard STBC scheme. Let us first consider the case in which the frequency offsets are negligible. In a differential scheme (for $\omega_k = 0$), the information-bearing block Z_b is not transmitted directly; instead, it is differentially encoded in such a way that the *MLD based on two consecutive blocks*

does not require CSI. Differential STBC schemes achieving this desideratum were recently suggested in [6, 9]. Compared with a trained scheme, a differential scheme does not transmit pilot symbols and hence does not sacrifice information rate. On the other hand, if trading-off information rate for BER performance is what we would like, a trained scheme will do whereas a differential one will not. Additionally, a differential scheme requires unitary symbols (see [6, 9]), while a training-based one has no such restriction. Hence we believe that a differential STBC scheme (such as the ones in [6, 9]) may be useful in fast fading scenarios; otherwise a trained scheme may be preferred .

In the perhaps more practically relevant case in which the frequency offsets cannot be ignored, we can think of using *double differential encoding* in an attempt to avoid the need for CSI at the receiver site. A double differential scheme was recently suggested in [7] for a class of group codes that does *not* include the STBC considered in this paper. The cited paper trades-off BER performance for computational simplicity. More precisely, the code matrices considered in [7] are diagonal, which limits the achievable BER performance. This resulting performance loss adds to the loss induced by the fact that the detector used is not the MLD (the MLD for double differential encoding is computationally involved even in the single antenna case [11]). Furthermore, our attempts to extend the double differential scheme of [7] to STBC (with non-diagonal code matrices $\{Z_b\}$) failed to provide a detector with a low complexity. Consequently a trained scheme (such as the one proposed in what follows) may be the way to go whenever the frequency offsets are deemed to be non-negligible.

2. TRAINED STBC SCHEME

Let $\mathbf{a}_k^*$ denote the k-th row of A, and let

$$
\begin{aligned}
\mathbf{y}_k^*(b) &= [y_k(Nb+1) \ldots y_k(Nb+N)] \\
\mathbf{e}_k^*(b) &= [e_k(Nb+1) \ldots e_k(Nb+N)]
\end{aligned} \tag{5}
$$

Using this notation we can write the data equation (1) (for $t = Nb+1, \ldots, Nb+N$) as follows:

$$
\mathbf{y}_k^*(b) = e^{i\omega_k Nb} \mathbf{a}_k^* Z_b \Omega_k + \mathbf{e}_k^*(b) \quad (k = 1, \ldots, m) \tag{6}
$$

where:

$$
\Omega_k = \begin{bmatrix} e^{i\omega_k} & & 0 \\ & \ddots & \\ 0 & & e^{iN\omega_k} \end{bmatrix} \tag{7}
$$

Assume that $Z_{b+1}, \ldots, Z_{b+L}$ are *pilot blocks*, which hence are known to the receiver (the receiver also knows the time when the transmitter launches the pilot blocks into the channel). We would like to use the pilot blocks and the corresponding received signals, $\{\mathbf{y}_k(b+1), \ldots, \mathbf{y}_k(b+L)\}_{k=1}^m$, to estimate A and $\{\omega_k\}$. Note that in general we need $L \geq 2$ to guarantee the identifiability of both $\{\mathbf{a}_k\}$ and $\{\omega_k\}$ in (6).

Owing to the assumption made on the noise term in (6), the *ML estimates* of $\mathbf{a}_k$ and ω_k, for given $Z_{b+1}, \ldots, Z_{b+L}$ are obtained by minimizing the following criterion:

$$
\sum_{l=1}^{L} \|\mathbf{y}_k(b+l) - \Omega_k^* Z_{b+l}^* e^{-i\omega_k N(b+l)} \mathbf{a}_k\|^2 \tag{8}
$$

As we will see shortly the minimization of (8) becomes simpler if

$$
Z_{b+1} = \ldots = Z_{b+L} \triangleq Z \tag{9}
$$

and

$$
Z \text{ is a square unitary matrix} \tag{10}
$$

To simplify the computations involved, and for lack of a better choice of the pilot blocks, we choose $\{Z_{b+l}\}_{l=1}^L$ to satisfy (9). Note that if the pilot blocks were chosen as STBC matrices, then satisfying (10) would be possible only for very few values of n. To satisfy (10) for any value of n we do not choose the pilot blocks as STBC matrices but let the pilot block Z to be any unitary $n \times n$ matrix. That is:

$$
Z^* Z = Z Z^* = I_{n \times n} \tag{11}
$$

(Z can possibly be scaled to satisfy a transmit power constraint).

Following the previous discussion, we assume that $\{Z_{b+l}\}_{l=1}^L$ in (8) satisfy (9) and (10). Consequently, we can write the criterion in (8) as:

$$
\sum_{l=1}^{L} \|Z\Omega_k \mathbf{y}_k(b+l) - e^{-i\omega_k Nl}(e^{-i\omega_k Nb} \mathbf{a}_k)\|^2 \tag{12}
$$

The minimization of (12) with respect to $\mathbf{a}_k$ yields:

$$
\hat{\mathbf{a}}_k = \frac{e^{i\omega_k Nb}}{L} Z\Omega_k \sum_{l=1}^{L} e^{i\omega_k Nl} \mathbf{y}_k(b+l) \tag{13}
$$

Inserting (13) in (12) gives the following function which is to be minimized with respect to ω_k:

$$
\left\| \left\{ I - \frac{1}{L} \begin{bmatrix} e^{-iN\omega_k}I \\ \vdots \\ e^{-iNL\omega_k}I \end{bmatrix} [e^{iN\omega_k}I \ldots e^{iNL\omega_k}I] \right\} \right.
$$
$$
\left. \times \begin{bmatrix} Z\Omega_k \mathbf{y}_k(b+1) \\ \vdots \\ Z\Omega_k \mathbf{y}_k(b+L) \end{bmatrix} \right\|^2
$$
$$
= const. - \frac{1}{L} \left\| \sum_{l=1}^{L} e^{iNl\omega_k} Z\Omega_k \mathbf{y}_k(b+l) \right\|^2 \tag{14}
$$
$$
= const. - \frac{1}{L} \left\| \sum_{l=1}^{L} e^{iNl\omega_k} \mathbf{y}_k(b+l) \right\|^2 \tag{15}
$$

Note that if either (9) or (10) were not true then we would have to work with a criterion which would be a more complicated function of ω_k than (15) [1].

It follows from (15) that the estimate of ω_k is to be obtained by solving the maximization problem:

$$
\max_{\omega_k} \left\| \sum_{l=1}^{L} e^{i\omega_k Nl} \mathbf{y}_k(b+l) \right\|^2 \tag{16}
$$

In general estimation of ω_k by maximizing (16) is still a bit complicated as it requires a (one-dimensional) nonlinear search. However for $L = 2$, a *closed form* expression for ω_k can be obtained (see below). For $L > 2$, to avoid the nonlinear search needed to maximize (16) we can choose L to be even ($L = 2K$) and *approximate* the maximizer of (16) as described below.

---

[1]To arrive at (15) from (14) it is not necessary that Z is a STBC. All that is necessary is that Z is unitary.

2.1. The case of $L = 2$

The criterion (16) with $L = 2$ has the following simple form:

$$\|\mathbf{y}_k(b+1) + e^{iN\omega_k}\mathbf{y}_k(b+2)\|^2$$
$$= const. + 2\,Re\,[e^{-iN\omega_k}\mathbf{y}_k^*(b+2)\mathbf{y}_k(b+1)] \quad (17)$$

The value of ω_k that maximizes (17) is readily seen to be

$$\hat{\omega}_k = \frac{arg\,[\mathbf{y}_k^*(b+2)\mathbf{y}_k(b+1)]}{N} \quad (18)$$

The corresponding channel estimate is obtained from (13):

$$\hat{\mathbf{a}}_k = \frac{e^{i\hat{\omega}_k Nb}}{2}\,Z\hat{\Omega}_k\,[e^{iN\hat{\omega}_k}\mathbf{y}_k(b+1) + e^{i2N\hat{\omega}_k}\mathbf{y}_k(b+2)] \quad (19)$$

Note that in most cases ω_k is sufficiently small such that $\omega_k N \leq 2\pi$. However, if this condition deemed to be invalid in a particular situation, then the simple estimator in (18) should not be used since it will be affected by aliasing. In the latter case, we may let Z_{b+l} depend on l and use the frequency estimate obtained by minimizing the (1D) function of (14) in lieu of that in (15).

2.2. The case of $L = 2K$ $(K > 1)$

In this case we can split the training data set in non-overlapping pairs, $\{\mathbf{y}_k(b+1), \mathbf{y}_k(b+2)\}, \ldots, \{\mathbf{y}_k(b+2K-1), \mathbf{y}_k(b+2K)\}$, estimate ω_k and $\mathbf{a}_k$ from each pair separately using (18) and (19) and average the K so-obtained estimates. While in general the estimates of $\{\omega_k$ and $\mathbf{a}_k\}$ derived in this way are not exact ML estimates, they have lower mean squared error (by a factor of K) than the estimates obtained from just one data pair (this follows easily from the assumptions made on the noise in (6)) *and* they are easy to compute.
To summarize:

- In a fast fading scenario we choose $L = 2$ and estimate $\{\omega_k, \mathbf{a}_k\}$ using (18) and (19).

  Note that the two training blocks used for acquiring CSI at the receiver do not have to be adjacent to one another. For instance, they can be separated by a STBC data block as shown in Figure 1.

Training Block	STBC Data Block	Training Block	STBC Data Block	• • •

$\leftarrow\ \text{n}\ \rightarrow|\leftarrow N \succeq \text{n} \rightarrow|$
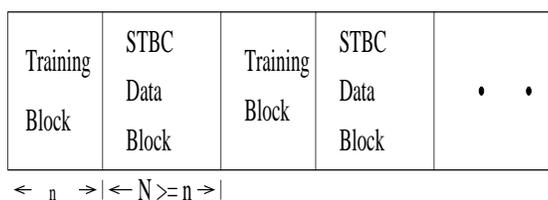
Figure 1: Transmission scheme for fast fading.

  The estimation formulas in (18) and (19) can be easily modified to take into account the fact that the transmission rate of the scheme in Figure 1 is $P/(N + n)$, which for $N = n = P$ becomes $1/2$. Also note that the scheme relies on the assumption that the channel (i.e, A and $\{\omega_k\}$) is nearly constant over these blocks (which is basically what a double differential scheme would also require).

- For slow fading scenarios we can choose $L = 2K$ with $K > 1$ and estimate $\{\omega_k, \mathbf{a}_k\}$ by the averaging technique outlined in Section 2.2.

Alternatively, to avoid reducing the transmission rate too much, we can choose $L = 2$ but unlike the scheme in Figure 1 we insert more than one data block in between the two training blocks. Several other alternative schemes are possible when the fading is not too fast. While we will not dwell into their details, the point we want to make is that a training-based scheme can use the information that the fading is slow to improve the BER performance (possibly at the cost of transmission rate), whereas a (double) differential scheme does not have this flexibility.

The so-obtained estimates $\{\hat{\omega}_k, \hat{\mathbf{a}}_k\}$ are used in the MLD, in lieu of the true unknown values $\{\omega_k, \mathbf{a}_k\}$ to detect the information-bearing symbols transmitted until a new set of pilot blocks is launched into the channel.

3. NUMERICAL EXAMPLES AND CONCLUDING REMARKS

In this section we consider the use of the trained STBC scheme discussed in the previous section in a fast fading channel. We consider a system with two transmit antennas and one receive antenna. We assume that the frequency offsets $\{\omega_k\}$ are random variables with uniform distribution that lie between 0 and 0.5 radians. The complex channel coefficients $\{A_{kp}\}$ are assumed to be i.i.d. Gaussian random variables with mean zero and variance 1. This channel model is the same as the one used in [7].

The transmission scheme is as in Figure 1, that is we transmit one training block $Z = I_{2\times 2}$ followed by a data block. This is followed by another training block and so on. Thus the effective rate is $1/2$. We use the two training blocks adjacent to a particular data block to estimate the channel for that particular data block. In this way we can allow for fast fading channels. The STBC block is an Alamouti code [1] with QPSK constellation. We normalize the STBC such that the corresponding (total) transmitted power is equal to 2, as for the pilot blocks. Since our transmission scheme has rate $1/2$ the effective spectral efficiency is 1 b/s/Hz. Simulations were done for 10000 independent channel realizations. In each realization 10 data blocks were transmitted. In Figure 2 we have plotted the Bit Error Rate (BER) for different values of the bit SNR. Plotted also for comparison is the BEP of the double differential scheme in [7] with the same spectral efficiency and total transmit power.

From the figure it can be seen that the new scheme outperforms the one in [7] by about 3 dB. Moreover the space-time block codes we have used are simpler to decode.

Finally we remark on the fact that the codes we have considered so far are orthogonal space-time block codes. However the algorithm developed in this paper for channel estimation can be applied to any type of space-time code. In particular the code matrices need not be unitary nor square. The available differential/double differential schemes require the code matrix to be square and unitary.

Acknowledgment

4. REFERENCES

[1] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communication," *IEEE Jl. on Select Areas in Comm.*, vol. 16, pp. 1451–1458, October 1998.

[2] G. Ganesan and P. Stoica, "Space-time diversity," in *Signal Processing Advances in Wireless & Mobile Communications-Volume 2* (G. Giannakis, Y. Hua, P. Stoica, and L. Tong, eds.), pp. 59–87, New Jersey: Prentice Hall, 2000.

[3] G. Ganesan and P. Stoica, "Space-Time Diversity Using Orthogonal and Amicable Orthogonal Designs," in *Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Istanbul, Turkey), 2000.

[4] G. Ganesan and P. Stoica, "Space-Time Block Codes: A Maximum SNR Approach," *IEEE Trans. on Info. Theory*, vol. 47, pp. 1650–1656, May 2001.

[5] G. Ganesan and P. Stoica, "Utilizing Space-Time Diversity for Wireless Communications," *Wireless Personal Communications*, vol. 18, pp. 149–163, August 2001.

[6] G. Ganesan and P. Stoica, "Differential Detection Based on Space-Time Block Codes," 2002. To appear in Wireless Personal Communications.

[7] Z. Liu, G. B. Giannakis, and B. L. Hughes, "Double Differential Space-Time Block Coding For Time-Selective Fading Channels," *IEEE Transactions on Communications*, vol. 49, pp. 1529–1539, September 2001.

[8] P. Stoica and G. Ganesan, "Space-Time Block Codes: Trained, Blind and Semi-Blind Detection," 2002. To appear in Digital Signal Processing - A review Journal.

[9] V. Tarokh and H. Jafarkhani, "A Differential Detection Scheme for Transmit Diversity," *IEEE Journal on Selected Areas in Communication*, vol. 18, pp. 1169–1174, July 2000.

[10] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-Time Block Codes from Orthogonal Designs," *IEEE Trans. on Info. Theory*, vol. 45, pp. 1456–1467, July 1999.

[11] D. K. van Alphen and W. C. Lindsey, "Higher-Order Differential Phase Shift Keyed Modulation," *IEEE Transactions on Communications*, vol. 42, pp. 440–448, February/March/April 1994.
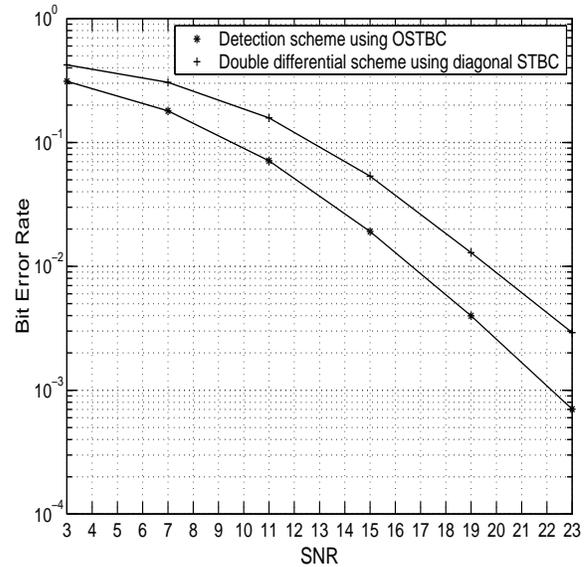
Figure 2: Bit Error Rates for the proposed detection scheme and the scheme in [7]