Channel Estimation and Data Detection for MIMO Systems under Spatially and Temporally Colored Interference

Yi Song

Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario, Canada K7L 3N6
Email: songy@ee.queensu.ca

Steven D. Blostein

Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario, Canada K7L 3N6
Email: sdb@ee.queensu.ca

The impact of interference on multiple-input multiple-output (MIMO) systems has recently attracted interest. Most studies of channel estimation and data detection for MIMO systems consider spatially and temporally white interference at the receiver. In this paper, we address channel estimation, interference correlation estimation, and data detection for MIMO systems under both spatially and temporally colored interference. We examine the case of one dominant interferer in which the data rate of the desired user could be the same as or a multiple of that of the interferer. Assuming known temporal interference correlation as a benchmark, we derive maximum likelihood (ML) estimates of the channel matrix and spatial interference correlation matrix, and apply these estimates to a generalized version of the Bell Labs Layered Space-Time (BLAST) ordered data detection algorithm. We then investigate the performance loss by not exploiting interference correlation. For a $(5, 5)$ MIMO system undergoing independent Rayleigh fading, we observe that exploiting both spatial and temporal interference correlation in channel estimation and data detection results in potential gains of 1.5 dB and 4 dB for an interferer operating at the same data rate and at half the data rate, respectively. Ignoring temporal correlation, it is found that spatial correlation accounts for about 1 dB of this gain.

Keywords and phrases: multiple-input multiple-output, interference, channel estimation, data detection.

1. INTRODUCTION

Wireless systems with multiple transmitting and receiving antennas have been shown to have a large Shannon channel capacity in a rich scattering environment [1, 2]. By transmitting parallel data streams over a multiple-input multi-output (MIMO) channel, it was shown that the Shannon capacity of the MIMO channel increases significantly with the number of transmitting and receiving antennas [2]. Layered space-time architectures were proposed for high-rate transmission in [3, 4]. Space-time coding techniques have also been investigated [5, 6].

While substantial research efforts have focussed on point-to-point MIMO link performance, the impact of interference on MIMO systems has received less interest. In a cellular environment, cochannel interference (CCI) from other cells exists due to channel reuse. In [7], channel capacities in the presence of spatially colored interference were derived under different assumptions of knowledge of the channel matrix and interference statistics at the transmitter. The

impact of spatially colored interference on MIMO channel capacity was studied in [8, 9, 10]. The capacity of MIMO systems with interference in the limiting case of a large number of antennas was studied in [11]. The overall capacity of a group of users, each employing a MIMO link, was investigated in [12]. The output signal-to-interference power ratio (SIR) was analytically calculated in [13], when a single data stream is transmitted over independent Rayleigh MIMO channels. While the majority of the studies deals with channel capacity, in this paper we focus on the achievable symbol error rate performance of a MIMO link with interference.

Prior results on estimation of vector channels and spatial interference statistics for code division multiple access (CDMA) single-input multiple-output systems can be found in [14]. Most studies of channel estimation and data detection for MIMO systems assume spatially and temporally white interference. For example, in [15], maximum likelihood (ML) estimation of the channel matrix using training sequences was presented assuming temporally white interference. Assuming perfect knowledge of the channel matrix at

the receiver, ordered zero-forcing (ZF) and minimum mean-squared error (MMSE) detection were studied for both spatially and temporally white interference in [4, 16], respectively. However, in cellular systems, the interference is, in general, both spatially and temporally colored.

In this paper, we propose and study a new algorithm that jointly estimates the channel matrix and the spatial interference correlation matrix in an ML framework. We develop a multi-vector-symbol MMSE data detector that exploits interference correlation. In the case of a single dominant interferer and large signal-to-noise ratio (SNR), we show that spatial and temporal second-order interference statistics can be decoupled in the form of a matrix Kronecker product. In finite SNR, the decoupling of spatial and temporal statistics of interference-plus-noise is only an approximation. We also determine the conditions where this approximation breaks down.

Although temporal interference correlation is difficult to estimate in practice, our objectives are to determine the performance benchmark achieved if temporal correlation was known. As sources of temporal correlation, we consider cases in which the data rate of the desired user is either the same as or a multiple of that of the interferer. The new ML algorithm serves as a performance benchmark when temporal and spatial interference correlation are exploited in joint channel estimation and data detection. We also assess the performance improvement obtained in more practical cases where only part of the correlation information is exploited, including the performance obtained by assuming temporally white interference, that is, ignoring temporal correlation.

The paper is organized as follows. In Section 2, we present our system model of temporal and spatial interference. In Section 3, we derive ML estimates of channel and spatial interference correlation matrices assuming known temporal interference correlation. In Section 4, one-vector-symbol detection is extended to a multi-vector-symbol version which is used to exploit temporal interference correlation. In Section 5, we consider the case of one interferer and large SNR and assess the benefits of taking temporal and/or spatial interference correlation into account for channel estimation and data detection. We then examine the level of SNR at which the approximation of separate spatial and temporal interference-plus-noise statistics break down. In cases where the spatial and temporal correlation are not separable, the performance improvement obtained by exploiting the spatial correlation is evaluated. For reference, comparisons are made to the well-known direct matrix inversion (DMI) algorithm [17], generalized to multiple input signals, a batch method that does not require estimates of channel and spatial interference correlation matrices.

In this paper, the notation $(\cdot)^T$ refers to transpose, $(\cdot)^*$ refers to conjugate, $(\cdot)^\dagger$ refers to conjugate transpose, and $\mathbf{I}_N$ refers to an $N \times N$ identity matrix.

2. SYSTEM MODEL

We consider a single-user link consisting of N_t transmitting and N_r receiving antennas, denoted as (N_t, N_r). The desired user transmits data frame by frame. Each frame has M data vectors. The first N data vectors are used for training, so that the desired user's channel matrix and interference statistics can be estimated, and the remaining data vectors are for information transmission. In a slow flat fading environment, the received signal vector at time j is expressed as

$$\mathbf{y}_j = \mathbf{H}\mathbf{x}_j + \mathbf{n}_j, \quad j = 0, \dots, M - 1, \tag{1}$$

where $\mathbf{x}_j$ is the transmitted data vector, $\mathbf{H}$ is the $N_r \times N_t$ spatial channel gain matrix, and the interference vector $\mathbf{n}_j$ is zero-mean circularly symmetric complex Gaussian. We assume that the channel matrix $\mathbf{H}$ is fixed during one frame. This is a reasonable assumption since high-speed data services envisioned for MIMO systems are generally intended for low mobility users. By the same argument, it is also assumed that the interference statistics are fixed during one frame.

In practice, the interference may be both spatially and temporally correlated. We assume that the cross correlation between the interference vectors at time i and j is $E\{\mathbf{n}_i\mathbf{n}_j^\dagger\} = \boldsymbol{\Lambda}_M(i, j)\mathbf{R}$, where $\boldsymbol{\Lambda}_M(i, j)$ is the (i, j)th element of an $M \times M$ matrix $\boldsymbol{\Lambda}_M$. The (i, j)th element of matrix $\mathbf{R}$ is the correlation between the ith and jth elements of interference vector $\mathbf{n}_k$, $k \in 0, \dots, M - 1$. As a result, the covariance matrix of the concatenated interference vector $\bar{\mathbf{n}} = [\mathbf{n}_0^T \cdots \mathbf{n}_{M-1}^T]^T$ is

$$
\begin{aligned}
E\{\bar{\mathbf{n}}\bar{\mathbf{n}}^\dagger\} &= \begin{bmatrix} \boldsymbol{\Lambda}_M(0,0)\mathbf{R} & \cdots & \boldsymbol{\Lambda}_M(0, M-1)\mathbf{R} \\ \vdots & & \vdots \\ \boldsymbol{\Lambda}_M(M-1,0)\mathbf{R} & \cdots & \boldsymbol{\Lambda}_M(M-1, M-1)\mathbf{R} \end{bmatrix} \\
&= \boldsymbol{\Lambda}_M \otimes \mathbf{R},
\end{aligned}
\tag{2}
$$

where $\otimes$ denotes Kronecker product, and matrices $\boldsymbol{\Lambda}_M$ and $\mathbf{R}$ capture the temporal and spatial correlation of the interference, respectively. The above model implies that the spatial and temporal interference statistics are separable. The correlation matrices $\boldsymbol{\Lambda}_M$ and $\mathbf{R}$ are determined by the application-specific signal model. In Section 5, we provide an example in which the interference covariance matrix has the above Kronecker product form. When the interference statistics can only be approximated by (2), the conditions where this approximation breaks down are investigated in Section 5.4.3. In addition to interference correlation, we remark that a decoupled temporal and spatial correlation structure arises in the statistics of fading vector channels consisting of a mobile with one antenna and a base station with an antenna array [18].

3. JOINT ESTIMATION OF CHANNEL AND SPATIAL INTERFERENCE STATISTICS

During a training period of N vector symbols, we concatenate the received signal vectors, the training signal vectors and the interference vectors as $\bar{\mathbf{y}} = [\mathbf{y}_0^T \cdots \mathbf{y}_{N-1}^T]^T$, $\bar{\mathbf{x}} = [\mathbf{x}_0^T \cdots \mathbf{x}_{N-1}^T]^T$, and $\bar{\mathbf{n}} = [\mathbf{n}_0^T \cdots \mathbf{n}_{N-1}^T]^T$, respectively. The

received signal in (1) is rewritten as the vector

$$\bar{\mathbf{y}} = (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}} + \bar{\mathbf{n}}, \tag{3}$$

where $\bar{\mathbf{n}}$ is circularly symmetric complex Gaussian with zero-mean and covariance matrix $\mathbf{\Lambda}_N \otimes \mathbf{R}$. Assuming prior knowledge of temporal interference correlation matrix $\mathbf{\Lambda}_N$, we need to estimate channel matrix $\mathbf{H}$ and spatial interference correlation matrix $\mathbf{R}$. If $\mathbf{R}$ and $\mathbf{\Lambda}_N$ are nonsingular, the conditional probability density function (pdf) is

$$\begin{aligned}\Pr(\bar{\mathbf{y}}|\mathbf{H}, \mathbf{R}) = &\frac{1}{\pi^{N \cdot N_r} \det(\mathbf{\Lambda}_N \otimes \mathbf{R})} \\ &\times \exp\Big\{-[\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]^\dagger \\ &\times (\mathbf{\Lambda}_N \otimes \mathbf{R})^{-1}[\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]\Big\}.\end{aligned} \tag{4}$$

3.1. ML solution

The ML estimate of the pair of matrices $(\mathbf{H}, \mathbf{R})$ is the value of $(\mathbf{H}, \mathbf{R})$ that maximizes the conditional pdf in (4), which is equivalent to maximizing $\ln \Pr(\bar{\mathbf{y}}|\mathbf{H}, \mathbf{R})$.

Letting $\mathbf{A}$ and $\mathbf{B}$ denote $m \times m$ and $n \times n$ square matrices, and using identities [19]

$$\begin{aligned}\det(\mathbf{A} \otimes \mathbf{B}) &= \det(\mathbf{A})^n \det(\mathbf{B})^m, \\ (\mathbf{A} \otimes \mathbf{B})^{-1} &= \mathbf{A}^{-1} \otimes \mathbf{B}^{-1},\end{aligned} \tag{5}$$

where $\mathbf{A}, \mathbf{B}$ are nonsingular, it can be shown that maximizing (4) is equivalent to minimizing

$$\begin{aligned}f(\mathbf{H}, \mathbf{R}) = &\ln \det(\mathbf{R}) \\ &+ \frac{1}{N}[\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]^\dagger \\ &\times (\mathbf{\Lambda}_N^{-1} \otimes \mathbf{R}^{-1})[\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}].\end{aligned} \tag{6}$$

Denoting the elements of $\mathbf{\Lambda}_N^{-1}$ as

$$\mathbf{\Lambda}_N^{-1} = \begin{bmatrix} \alpha_{0,0} & \cdots & \alpha_{0,N-1} \\ \vdots & & \vdots \\ \alpha_{N-1,0} & \cdots & \alpha_{N-1,N-1} \end{bmatrix}, \tag{7}$$

we rewrite (6) as

$$\begin{aligned}&f(\mathbf{H}, \mathbf{R}) \\ &= \ln \det(\mathbf{R}) \\ &+ \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha_{i,j}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^\dagger \mathbf{R}^{-1}(\mathbf{y}_j - \mathbf{H}\mathbf{x}_j) \\ &= \ln \det(\mathbf{R}) \\ &+ \text{trace}\Big\{\mathbf{R}^{-1}\frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha_{i,j}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)(\mathbf{y}_j - \mathbf{H}\mathbf{x}_j)^\dagger\Big\}.\end{aligned} \tag{8}$$

To find the value of $(\mathbf{H}, \mathbf{R})$ that minimizes $f(\mathbf{H}, \mathbf{R})$ in (8),

we set $\partial f(\mathbf{H}, \mathbf{R})/\partial \mathbf{H} = \mathbf{0}$. Define the weighted sample correlation matrices[1] as

$$\begin{aligned}\tilde{\mathbf{R}}_{yy} &= \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha_{i,j}\mathbf{y}_i\mathbf{y}_j^\dagger, \\ \tilde{\mathbf{R}}_{xy} &= \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha_{i,j}\mathbf{x}_i\mathbf{y}_j^\dagger, \\ \tilde{\mathbf{R}}_{xx} &= \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha_{i,j}\mathbf{x}_i\mathbf{x}_j^\dagger.\end{aligned} \tag{9}$$

Using the identities of matrix derivative [19], it can be shown [20] that (8) is minimized by

$$\hat{\mathbf{H}} = \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1}. \tag{10}$$

Setting $\partial f(\hat{\mathbf{H}}, \mathbf{R})/\partial \mathbf{R} = \mathbf{0}$, it can also be shown that the estimate of spatial interference correlation matrix is given by

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha_{i,j}(\mathbf{y}_i - \hat{\mathbf{H}}\mathbf{x}_i)(\mathbf{y}_j - \hat{\mathbf{H}}\mathbf{x}_j)^\dagger \tag{11}$$

$$= \tilde{\mathbf{R}}_{yy} - \hat{\mathbf{H}}\tilde{\mathbf{R}}_{xy}. \tag{12}$$

We remark that if $\tilde{\mathbf{R}}_{xy}$ and $\tilde{\mathbf{R}}_{xx}$ in (10) were known cross- and auto-correlation matrices, the estimate for $\mathbf{H}$ would represent the Wiener solution.

3.2. Special case: temporally white interference

If an interference is temporally white, with loss of generality, we may substitute $\mathbf{\Lambda}_N = \mathbf{I}_N$ into (9), (10), (11), and (12), and obtain estimates

$$\hat{\mathbf{H}}_w = \mathbf{R}_{xy}^\dagger \mathbf{R}_{xx}^{-1}, \tag{13}$$

$$\hat{\mathbf{R}}_w = \mathbf{R}_{yy} - \hat{\mathbf{H}}_w \mathbf{R}_{xy}, \tag{14}$$

where the subscript w indicates temporally white interference, and the sample correlation matrices are

$$\mathbf{R}_{yy} = \frac{1}{N}\sum_{i=0}^{N-1}\mathbf{y}_i\mathbf{y}_i^\dagger, \tag{15}$$

$$\mathbf{R}_{xy} = \frac{1}{N}\sum_{i=0}^{N-1}\mathbf{x}_i\mathbf{y}_i^\dagger, \tag{16}$$

$$\mathbf{R}_{xx} = \frac{1}{N}\sum_{i=0}^{N-1}\mathbf{x}_i\mathbf{x}_i^\dagger. \tag{17}$$

Note that $\hat{\mathbf{H}}_w$ in (13) is the same as the channel estimate used in [15].

---

[1]To distinguish weighted sample correlation matrices from conventional sample correlation matrices in Section 3.2, we denote the former by a tilde and the latter without a tilde.

3.3. *Whitening filter interpretation*

To obtain insight on the estimates in (10) and (12), we let the received signal vectors during the training period undergo a linear transformation where the transformed received signal vectors are

$$[\mathbf{y}_0' \cdots \mathbf{y}_{N-1}'] = [\mathbf{y}_0 \cdots \mathbf{y}_{N-1}]\boldsymbol{\Lambda}_N^{-1/2}. \tag{18}$$

At the output of the transformation, we have

$$\mathbf{y}_i' = \mathbf{H}\mathbf{x}_i' + \mathbf{n}_i', \quad i = 0, \ldots, N-1, \tag{19}$$

where the transformed training signal vectors and interference vectors are

$$\begin{aligned}
[\mathbf{x}_0' \cdots \mathbf{x}_{N-1}'] &= [\mathbf{x}_0 \cdots \mathbf{x}_{N-1}]\boldsymbol{\Lambda}_N^{-1/2}, \\
[\mathbf{n}_0' \cdots \mathbf{n}_{N-1}'] &= [\mathbf{n}_0 \cdots \mathbf{n}_{N-1}]\boldsymbol{\Lambda}_N^{-1/2},
\end{aligned} \tag{20}$$

respectively. Concatenating the transformed interference vectors as $\bar{\mathbf{n}}' = [\mathbf{n}_0'^T \cdots \mathbf{n}_{N-1}'^T]^T$, it can be shown that

$$\bar{\mathbf{n}}' = (\boldsymbol{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r})\bar{\mathbf{n}}, \tag{21}$$

where $\bar{\mathbf{n}} = [\mathbf{n}_0^T \cdots \mathbf{n}_{N-1}^T]^T$. Since the covariance matrix of $\bar{\mathbf{n}}$ is $\boldsymbol{\Lambda}_N \otimes \mathbf{R}$, the covariance matrix of $\bar{\mathbf{n}}'$ is

$$\begin{aligned}
\text{cov}(\bar{\mathbf{n}}') &= (\boldsymbol{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r}) \text{cov}(\bar{\mathbf{n}}) (\boldsymbol{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r})^{\dagger} \\
&= (\boldsymbol{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r})(\boldsymbol{\Lambda}_N \otimes \mathbf{R})(\boldsymbol{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r}) \\
&= \mathbf{I}_N \otimes \mathbf{R},
\end{aligned} \tag{22}$$

where we used $(\mathbf{A} \otimes \mathbf{B})^{\dagger} = \mathbf{A}^{\dagger} \otimes \mathbf{B}^{\dagger}$ and $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ [19]. We also used the fact that the temporal correlation matrix $\boldsymbol{\Lambda}_N$ is symmetric, as well as $\boldsymbol{\Lambda}_N^{-1/2}$. From (22), it is obvious that the transformed interference vectors $\{\mathbf{n}_0' \cdots \mathbf{n}_{N-1}'\}$ are temporally white with spatial correlation matrix $\mathbf{R}$.

As a result, we can estimate $\mathbf{H}$ and $\mathbf{R}$ from the sample correlation matrices of transformed signal vectors as in Section 3.2. The sample correlation matrix

$$\begin{aligned}
\mathbf{R}_{y'y'} &= \frac{1}{N}\sum_{i=0}^{N-1} \mathbf{y}_i' \mathbf{y}_i'^{\dagger} \\
&= \frac{1}{N}[\mathbf{y}_0' \cdots \mathbf{y}_{N-1}'][\mathbf{y}_0' \cdots \mathbf{y}_{N-1}']^{\dagger} \\
&= \frac{1}{N}[\mathbf{y}_0 \cdots \mathbf{y}_{N-1}]\boldsymbol{\Lambda}_N^{-1/2}\boldsymbol{\Lambda}_N^{-\dagger/2}[\mathbf{y}_0 \cdots \mathbf{y}_{N-1}]^{\dagger} \\
&= \frac{1}{N}[\mathbf{y}_0 \cdots \mathbf{y}_{N-1}]\boldsymbol{\Lambda}_N^{-1}[\mathbf{y}_0 \cdots \mathbf{y}_{N-1}]^{\dagger} = \tilde{\mathbf{R}}_{yy},
\end{aligned} \tag{23}$$

which shows that the weighted sample correlation matrix of $\{\mathbf{y}_0 \cdots \mathbf{y}_{N-1}\}$ is equivalent to the sample correlation matrix of $\{\mathbf{y}_0' \cdots \mathbf{y}_{N-1}'\}$. Similarly, the weighted sample correlation matrices $\tilde{\mathbf{R}}_{xy}$ and $\tilde{\mathbf{R}}_{xx}$ are equivalent to the sample correlation matrices $\mathbf{R}_{x'y'}$ and $\mathbf{R}_{x'x'}$, respectively. Therefore, the estimates in (10) and (12) can also be realized by first temporally

whitening the interference, and then forming the estimates from the sample correlation matrices of the transformed signal vectors.

4. DATA DETECTION

We focus on ordered MMSE detection due to the better performance of MMSE compared to ZF detection [21]. For received signal vector $\mathbf{y}_i = \mathbf{H}\mathbf{x}_i + \mathbf{n}_i$, modifying the BLAST algorithm in [16], the steps of ordered MMSE detection of $\mathbf{x}_i$ from $\mathbf{y}_i$ with estimated channel and interference spatial correlation matrices are as follows:

Step 1. Initialization: set $k = 1$, $\mathbf{H}_k = \hat{\mathbf{H}}$, $\tilde{\mathbf{x}}_k = \mathbf{x}_i$, $\tilde{\mathbf{y}}_k = \mathbf{y}_i$.

Step 2. Calculate the estimation error covariance matrix $\mathbf{P}_k = (\mathbf{I}_{N_t+1-k} + \mathbf{H}_k^{\dagger}\hat{\mathbf{R}}^{-1}\mathbf{H}_k)^{-1}$. Find $m = \arg\min_j \mathbf{P}_k(j, j)$, where $\mathbf{P}_k(j, j)$ denotes the jth diagonal element of $\mathbf{P}_k$. Hence, the mth signal component of $\tilde{\mathbf{x}}_k$ has the smallest estimation error variance.

Step 3. Calculate the weighting matrix $\mathbf{A}_k = (\mathbf{I}_{N_t+1-k} + \mathbf{H}_k^{\dagger}\hat{\mathbf{R}}^{-1}\mathbf{H}_k)^{-1}\mathbf{H}_k^{\dagger}\hat{\mathbf{R}}^{-1}$. The mth element of $\tilde{\mathbf{x}}_k$ is estimated by $\hat{x}_k^m = Q(\mathbf{A}_k(m, :)\tilde{\mathbf{y}}_k)$, where $\mathbf{A}_k(m, :)$ denotes the mth row of matrix $\mathbf{A}_k$, and $Q(\cdot)$ denotes the slicing operation appropriate to the signal constellation.

Step 4. Assuming that the detected signal is correct, remove the detected signal from the received signal $\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{y}}_k - \hat{x}_k^m\mathbf{H}_k(:, m)$, where $\mathbf{H}_k(:, m)$ denotes the mth column of $\mathbf{H}_k$.

Step 5. $\mathbf{H}_{k+1}$ is obtained by eliminating the mth column of matrix $\mathbf{H}_k$ and $\tilde{\mathbf{x}}_{k+1}$ is obtained by eliminating the mth component of vector $\tilde{\mathbf{x}}_k$.

Step 6. If $k < N_t$, increment k and go to Step 2.

We refer to this scheme as *one-vector-symbol detection*, as we detect $\mathbf{x}_i$ using $\mathbf{y}_i$ only.

When an interference is temporally colored, there may exist a performance to be gained by taking the temporal interference correlation into account. That is, we may use $\mathbf{y}_{N+1}, \ldots, \mathbf{y}_M$ to detect $\mathbf{x}_{N+1}, \ldots, \mathbf{x}_M$ jointly where N is the training length and M is the frame length. Due to the complexity of using all the received signal vectors and for simplicity of presentation, we consider a *two-vector-symbol detection* in which $(\mathbf{y}_i, \mathbf{y}_{i+1})$ is used to detect $(\mathbf{x}_i, \mathbf{x}_{i+1})$ jointly. The one-vector-symbol algorithm can be easily extended to the two-vector-symbol version by writing

$$\underbrace{\begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_{i+1} \end{bmatrix}}_{\check{\mathbf{y}}_i} = \underbrace{\begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}}_{\check{\mathbf{H}}} \underbrace{\begin{bmatrix} \mathbf{x}_i \\ \mathbf{x}_{i+1} \end{bmatrix}}_{\check{\mathbf{x}}_i} + \underbrace{\begin{bmatrix} \mathbf{n}_i \\ \mathbf{n}_{i+1} \end{bmatrix}}_{\check{\mathbf{n}}_i}. \tag{24}$$

With the estimated channel, an estimate of $\check{\mathbf{H}}$, denoted as $\hat{\check{\mathbf{H}}}$, can be obtained. Using the estimated spatial interference correlation and the known temporal interference correlation, we are able to estimate the covariance matrix of $\check{\mathbf{n}}_i$, denoted as $\hat{\check{\mathbf{R}}}$. Replacing $\mathbf{x}_i$, $\mathbf{y}_i$, $\hat{\mathbf{H}}$, and $\hat{\mathbf{R}}$ in the one-vector-symbol algorithm by $\check{\mathbf{x}}_i$, $\check{\mathbf{y}}_i$, $\hat{\check{\mathbf{H}}}$, and $\hat{\check{\mathbf{R}}}$, respectively, we obtain the two-vector-symbol detection algorithm.

5. APPLICATIONS

In this section, we apply the channel estimation in Section 3 and data detection in Section 4 to the case of a single-user link with one dominant cochannel interferer operating at different data rates.

5.1. System model

Consider a desired user with one dominant cochannel interferer. The assumption of one cochannel interferer can apply to cellular TDMA or FDMA systems when sectoring is used. For example, in 7-cell reuse systems, with 60 degree sectors, the number of cochannel interfering cells would be reduced to one [22]. We assume that the desired and interfering users have N_t and L transmitting antennas, respectively, and that there are N_r receiving antennas. Assuming that the thermal noise is small relative to the interference, we ignore the thermal noise in the problem formulation. An investigation of this assumption in channels with noise appears in Section 5.4.3. We also assume that over the duration of a transmitted frame, a randomly delayed replica of the interfering signal is transmitted continuously, and that the interference statistics do not change. This assumption may not hold for asynchronous packet transmission systems. In a slow flat fading environment, the vector signal at the receiving antennas is

$$\mathbf{y}(t) = \sqrt{\frac{P_s T}{N_t}} \mathbf{H} \sum_{k=0}^{M-1} \mathbf{x}_k \tilde{g}(t - kT) \\ + \sqrt{\frac{P_I T_I}{L}} \mathbf{H}_I \sum_{k=-\infty}^{\infty} \mathbf{b}_k \tilde{g}_I(t - kT_I - \tau),$$

(25)

where M is the frame length, and $\mathbf{H}$ ($N_r \times N_t$) and $\mathbf{H}_I$ ($N_r \times L$) are the channel matrices of the desired and interfering users, respectively. The channel matrices are also assumed fixed over a frame and have independent realizations from frame to frame. The data transmission rates of the desired and interfering users are $1/T$ and $1/T_I$, respectively. The spectra of transmit impulse responses $\tilde{g}(t)$ and $\tilde{g}_I(t)$ are square root raised cosines with parameters T and T_I, respectively. The same roll-off factor, β, is assumed for both $\tilde{g}(t)$ and $\tilde{g}_I(t)$. The data vectors of the desired and interfering users are $\mathbf{x}_k$ ($N_t \times 1$) and $\mathbf{b}_k$ ($L \times 1$), respectively. We assume that the data symbols in $\mathbf{x}_k$'s and $\mathbf{b}_k$'s are mutually independent, zero mean, and with unit variance. We denote P_s and P_I as the transmit powers of the desired and interfering users, respectively. The delay of the interfering user relative to the desired user is τ, assumed to lie in $0 \le \tau < T_I$.

Passing $\mathbf{y}(t)$ in (25) through a filter matched to the transmit impulse response of the desired user, $\tilde{g}(t)$, the vector signal at the output of the matched filter is

$$\mathbf{y}_{\text{MF}}(t) = \sqrt{\frac{P_s T}{N_t}} \mathbf{H} \sum_{k=0}^{M-1} \mathbf{x}_k g(t - kT) \\ + \sqrt{\frac{P_I T_I}{L}} \mathbf{H}_I \sum_{k=-\infty}^{\infty} \mathbf{b}_k g_I(t - kT_I - \tau),$$

(26)

where $g(t) = \tilde{g}(t) * \tilde{g}(t)$, $g_I(t) = \tilde{g}_I(t) * \tilde{g}(t)$, and $*$ denotes convolution. As a result, $g(t)$ has a raised cosine spectrum and satisfies the Nyquist condition for zero intersymbol interference.

Assuming perfect synchronization for the desired user, as we sample the output of the matched filter (26) at time $t = jT$, we obtain

$$\mathbf{y}_j = \sqrt{\frac{P_s T}{N_t}} \mathbf{H} \mathbf{x}_j + \underbrace{\sqrt{\frac{P_I T_I}{L}} \mathbf{H}_I \sum_{k=-\infty}^{\infty} \mathbf{b}_k g_I(jT - kT_I - \tau)}_{\mathbf{n}_j}.$$

(27)

The interference vector $\mathbf{n}_j$ is zero mean since the data vector of interferer $\mathbf{b}_k$ is zero mean. Note that there is no intersymbol interference for the desired user. However, due to the interferer's delay and/or mismatch between the transmit and receive impulse responses, intersymbol interference exists for the interferer.

5.2. Interference statistics

The cross correlation between the interference vectors in (27) at time jT and qT is

$$E\{\mathbf{n}_j \mathbf{n}_q^\dagger\} = \frac{P_I T_I}{L} \mathbf{H}_I \\ \cdot E\left\{ \left(\sum_{k_1=-\infty}^{\infty} \mathbf{b}_{k_1} g_I(jT - k_1 T_I - \tau) \right) \\ \times \left(\sum_{k_2=-\infty}^{\infty} \mathbf{b}_{k_2}^\dagger g_I(qT - k_2 T_I - \tau) \right) \right\} \mathbf{H}_I^\dagger \\ = \frac{P_I T_I}{L} \mathbf{H}_I \mathbf{H}_I^\dagger \\ \cdot \sum_{k=-\infty}^{\infty} \{ g_I(jT - kT_I - \tau) g_I(qT - kT_I - \tau) \},$$

(28)

where the last equality is due to the facts that $E\{\mathbf{b}_{k_1} \mathbf{b}_{k_2}^\dagger\} = \mathbf{0}$ for $k_1 \ne k_2$ and $E\{\mathbf{b}_k \mathbf{b}_k^\dagger\} = \mathbf{I}_L$.

During a training period of N vector symbols, the covariance matrix of the concatenated interference vector $\bar{\mathbf{n}} = [\mathbf{n}_0^T \cdots \mathbf{n}_{N-1}^T]^T$ has the form of (2), where

$$\mathbf{\Lambda}_N(j,q) = \sum_{k=-\infty}^{\infty} \{ g_I(jT - kT_I - \tau) g_I(qT - kT_I - \tau) \}, \\ 0 \le j, q \le N - 1,$$

(29)

$$\mathbf{R} = \frac{P_I T_I}{L} \mathbf{H}_I \mathbf{H}_I^\dagger.$$

(30)

The $N_r \times N_r$ spatial correlation matrix $\mathbf{R}$ is determined by the interferer's channel matrix. The $N \times N$ temporal correlation matrix $\mathbf{\Lambda}_N$ depends on parameters T and T_I, delay τ, and pulse $g_I(t)$; it can be calculated a priori if these parameters

are known. The temporal correlation is due to intersymbol interference in the sampled interfering signal. We remark that for the case of multiple interferers with the same delay, the covariance matrix of interference also has the form of (2).

We study temporal interference correlation in the cases where (1) the interferer has the same data rate as that of the desired signal ($T = T_I$) and (2) the data rate of the desired user is an integer multiple of that of the interferer ($T_I = mT$, $m > 1$).

5.2.1. Interferer at the same data rate as the desired signal

With $T = T_I$, $g_I(t)$ has a raised cosine spectrum and is given by [23]

$$g_I(t) = \text{sinc}\left(\frac{\pi t}{T}\right)\frac{\cos(\pi\beta t/T)}{1 - 4\beta^2 t^2/T^2}. \tag{31}$$

We note that $\boldsymbol{\Lambda}_N(j, q)$ depends on $j - q$. This indicates that the sequence consisting of interference vectors is stationary. Hence, the temporal correlation matrix is a symmetric Toeplitz. By appropriate truncation of the infinite series in (29), we can numerically calculate the temporal correlation matrix. For the case of $\beta = 1$, $T = 1$, and $\tau = 0.5$, the elements of the temporal correlation matrix are

$$\boldsymbol{\Lambda}_N(j, q) = \begin{cases} 0.5 & j = q, \\ 0.25 & |j - q| = 1 \\ 0 & \text{otherwise}. \end{cases} \text{ for } 0 \le j, q \le N - 1, \tag{32}$$

5.2.2. Interferer at a lower data rate than the desired signal

It can be shown that $g_I(t)$ is given by

$$g_I(t) = \mathcal{F}^{-1}\left\{\sqrt{G_{rc,T_I}(f)}\sqrt{G_{rc,T}(f)}\right\}, \tag{33}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform and $G_{rc,T}(f)$ is the raised cosine Fourier spectrum with parameter T and roll-off factor β. Unlike the case of the same data rate interferer where $\boldsymbol{\Lambda}_N(j, q)$ depends on $j - q$, in the case of lower data rate interferer, $\boldsymbol{\Lambda}_N(j, q)$ depends on the values of j and q. This indicates that the sequence consisting of interference vectors is cyclostationary [23, 24]. With $T_I = mT$, it can be shown that $\boldsymbol{\Lambda}_N(j, q)$ is periodic with period m, that is, $\boldsymbol{\Lambda}_N(j, q) = \boldsymbol{\Lambda}_N(j+m, q+m)$. As a result, the temporal correlation matrix $\boldsymbol{\Lambda}_N$ is symmetric, but not Toeplitz. Furthermore, for $N \ge m$, the number of nontrivial eigenvalues of $\boldsymbol{\Lambda}_N$ is $\lceil N/m \rceil$, where $\lceil \cdot \rceil$ rounds the argument to the nearest integer towards infinity [25]. For the case of $T_I = 2T$, $T = 1$, $\beta = 1$, $\tau = 0.25$, and training length $N = 6$, by numerical calculation of (29) with appropriate series truncation, the temporal

correlation matrix is

$$\boldsymbol{\Lambda}_6 = \begin{bmatrix} 0.648 & 0.400 & -0.048 & -0.006 & -0.010 & -0.001 \\ 0.400 & 0.277 & 0.105 & 0.084 & 0.002 & 0.011 \\ -0.048 & 0.105 & 0.648 & 0.400 & -0.048 & -0.006 \\ -0.006 & 0.084 & 0.400 & 0.277 & 0.105 & 0.084 \\ -0.010 & 0.002 & -0.048 & 0.105 & 0.648 & 0.400 \\ -0.001 & 0.011 & -0.006 & 0.084 & 0.400 & 0.277 \end{bmatrix}. \tag{34}$$

Note that $\boldsymbol{\Lambda}_6$ in (34) is singular as the number of nontrivial eigenvalues is 3.

5.3. Data detection without estimating channel and interference

During a training period of N symbol vectors, instead of estimating the channel matrix and interference statistics, one can alternatively employ a least squares (LS) estimate of matrix $\mathbf{M}$ which minimizes the average estimation error

$$f_2(\mathbf{M}) = \text{trace}\left\{\frac{1}{N}\sum_{i=0}^{N-1}(\mathbf{x}_i - \mathbf{M}\mathbf{y}_i)(\mathbf{x}_i - \mathbf{M}\mathbf{y}_i)^{\dagger}\right\}. \tag{35}$$

By setting $\partial f_2(\mathbf{M})/\partial\mathbf{M} = \mathbf{0}$, we obtain

$$\mathbf{M} = \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}, \tag{36}$$

where the sample correlation matrices $\mathbf{R}_{xy}$ and $\mathbf{R}_{yy}$ are defined in (16) and (15), respectively. The transmitted signal vector $\mathbf{x}_i$ is detected as $Q(\mathbf{M}\mathbf{y}_i)$, where $Q(\cdot)$ is the slicing operation appropriate to the signal constellation. We remark that (36) is the well-known DMI algorithm [17], generalized for multiple input signals. A significant loss in performance is expected for this LS detector, since without estimates of channel and spatial interference correlation matrices, iterative MMSE detection cannot be performed.

5.4. Simulation results

Monte Carlo simulations are used to assess the benefits of taking temporal and spatial interference correlation into account, for channel estimation and data detection in the case of one interferer. Although temporal interference correlation may be difficult to estimate in practice, we examine this as a benchmark and determine the performance loss due to ignoring this correlation. We evaluate average symbol error rates (SERs) in independent Rayleigh fading channels of rich scattering, that is, the elements in channel matrices $\mathbf{H}$ and $\mathbf{H}_I$ are independent, identically distributed (i.i.d.) zero-mean complex Gaussian with unit variance. We assume that the desired user has 5 transmitting and 5 receiving antennas, and the interfering user has 6 transmitting antennas.[2] Both the desired and interfering users employ uncoded quadrature phase shift keying (QPSK) modulation. The training signal vectors are columns of a fast Fourier transform (FFT) matrix

---

[2]For a nonsingular spatial interference correlation matrix, we set $N_r \le L$.

[16] to guarantee orthogonal training sequences from different transmitting antennas. We define $\text{SIR(dB)} = 10 \log P_s/P_I$. Without loss of generality, we set $P_I = 1$ in the simulation. The SERs of two cases are simulated: (1) interferer at the same data rate as the desired signal and (2) the data rate of the desired user is twice that of the interferer.

In Figures 1, 2, 3, and 4, with solid and dashed lines representing one- and two-vector-symbol data detection, respectively, we plot average SERs for the following cases:

(a) perfectly known channel parameters and interference statistics, with one-vector-symbol (curve 1) and two-vector-symbol (curve 2) detection;

(b) channel and spatial interference correlation matrices are estimated assuming known temporal interference correlation, with one-vector-symbol (curve 3) and two-vector-symbol (curve 4) detection;

(c) channel and spatial interference correlation matrices are estimated assuming temporally white interference, with one-vector-symbol detection (curve 5);

(d) only the channel matrix $\mathbf{H}$ is estimated assuming temporally white interference; an identity spatial interference correlation matrix is used in one-vector-symbol data detection (curve 6);

(e) LS estimate of the transmitted signal vector without ordered detection (Section 5.3) (curve 7).

We remark that cases (a) and (b) are benchmarks presented for reference, while case (d) corresponds to the well-known BLAST system in [4, 16].

5.4.1. Interferer at the same data rate as the desired signal

We examine the case of $T = 1$, $\beta = 1$, and $\tau = 1/2$, and the nonsingular temporal interference correlation matrix shown in (32). Figures 1 and 2 show the average SERs for training lengths $2N_t$ and $4N_t$, respectively. Comparing the LS detection (curve 7) with other methods, much lower SERs can be achieved by using ordered MMSE detection as expected.

Comparing curves 5 and 6, we observe that for a training length of $4N_t$ symbols, gains can be obtained by estimating spatial interference correlation. However, shorter training lengths such as $2N_t$ produce inaccurate estimates of spatial interference correlation which in turn do not yield any benefit over assuming spatially white interference. As expected, we observe that the improvement by taking into account estimated spatial correlation increases with longer training lengths.

Examining curves 3 and 5 in Figure 2, we observe that the improvement in taking temporal interference correlation into account in *channel estimation* is not significant. Moreover, this rate of improvement rapidly diminishes as the training length increases. This can be explained by noting that in estimating channel and spatial interference correlation matrices for temporally colored interference, the received signal vectors first undergo a transformation which temporally whitens the interference vectors as discussed in Section 3.3. Since the temporal correlation in (32) drops

quickly to zero after one time lag, the benefit in temporal whitening of interference vectors is not significant, especially for long training lengths.

By comparing curves 3 and 4 in Figure 2, there is a slight improvement in using two-vector-symbol over one-vector-symbol detection. This implies that not much gain can be achieved by taking temporal interference correlation into account in *data detection*, owing to the low temporal correlation. Due to better estimates of channel and interference spatial correlation matrices obtained with a longer training length, the performance gap between curves 3 and 4 should increase as the training length increases.

By comparing curves 4 and 6 in Figure 2, we observe a 1.5 dB gain in SIR obtained by estimating spatial interference correlation and taking explicit advantage of known temporal interference correlation in channel estimation and data detection using a training length of $4N_t$. About 1 dB of that gain is due to the estimation of spatial interference correlation, and the remaining 0.5 dB gain is due to exploiting temporal interference correlation in channel estimation and data detection.

5.4.2. Interferer at a lower data rate than the desired signal

We examine the case of $T_I = 2T$, $T = 1$, $\beta = 1$, $\tau = 0.25$ and the temporal interference correlation matrix for training length $N = 6$ shown in (34). Recall that the temporal correlation matrix for the lower-data-rate-interferer case is singular. To avoid the singularity, the diagonal elements of $\mathbf{\Lambda}_N$ are increased by a small amount; hence, the temporal correlation matrix used for channel estimation may be modified to $\mathbf{\Lambda}_N + \delta\mathbf{I}_N$ within the proposed framework. In our simulation, we chose $\delta = 0.01$.

The same set of average SER curves as in the same-data-rate-interferer case are simulated. Figures 3 to 4 show the SERs for different training lengths. As in the case of the same-data-rate interferer, curve 7 illustrates the poor performance without ordered detection. Curves 5 and 6 suggest that for short training lengths it is better to estimate only the channel matrix and assume spatially white interference in data detection; however, for moderately long training lengths, gains can be obtained by estimating spatial interference correlation.

By examining curves 3 and 5 in Figure 4, we observe that the improvement in taking temporal interference correlation into account in *channel estimation*, although larger than that in the same-data-rate-interferer case due to the high temporal correlation in the lower-data-rate-interferer case, is still not that significant.

In contrast to the same-data-rate-interferer case, curves 3 and 4 in Figure 4 show that the improvement of two-vector-symbol over one-vector-symbol detection is significant due to the higher temporal interference correlation. This implies that a significant gain can be achieved by taking the known temporal interference correlation into account in *data detection* for the lower-data-rate-interferer case.

By comparing curves 4 and 6 in Figure 4, for the training length $4N_t$, there is a total of 4 dB gain in SIR by estimating spatial interference correlation and taking advantage of the
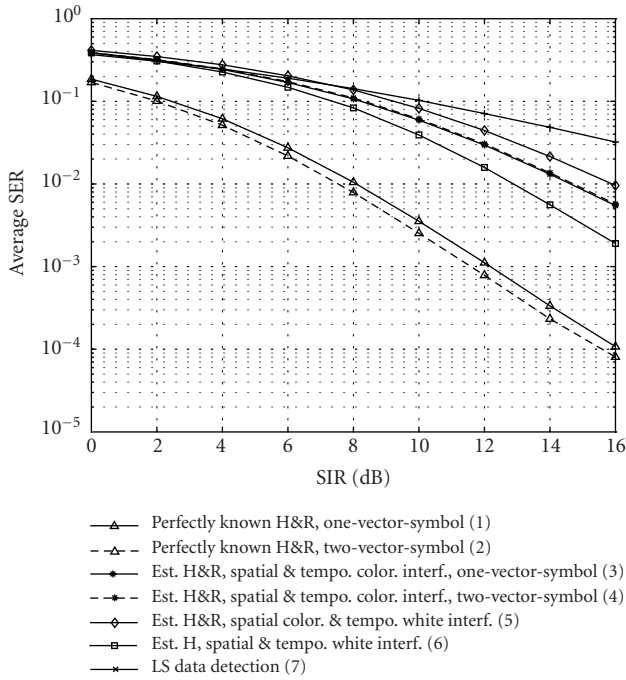
Figure 1: Average SER versus SIR with $N_t = N_r = 5$, $L = 6$, and training length $2N_t$ under independent Rayleigh fading. Both the desired and the interfering users are at the same data rate.
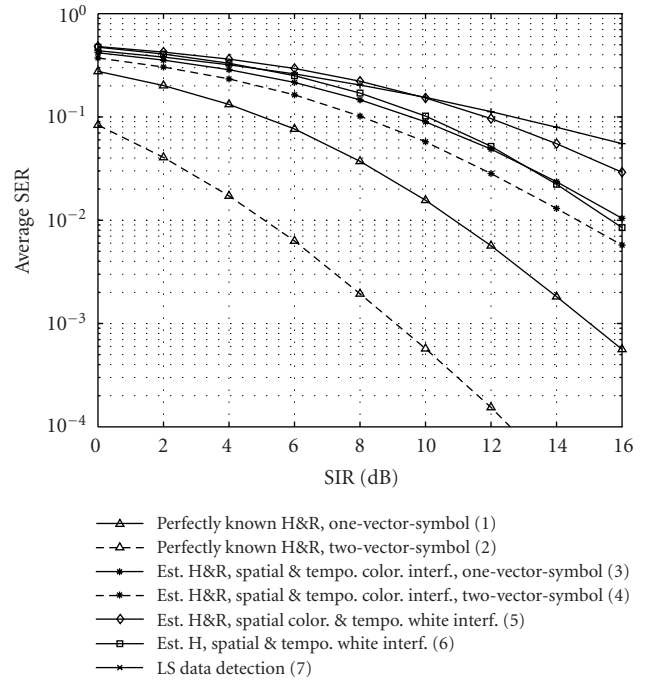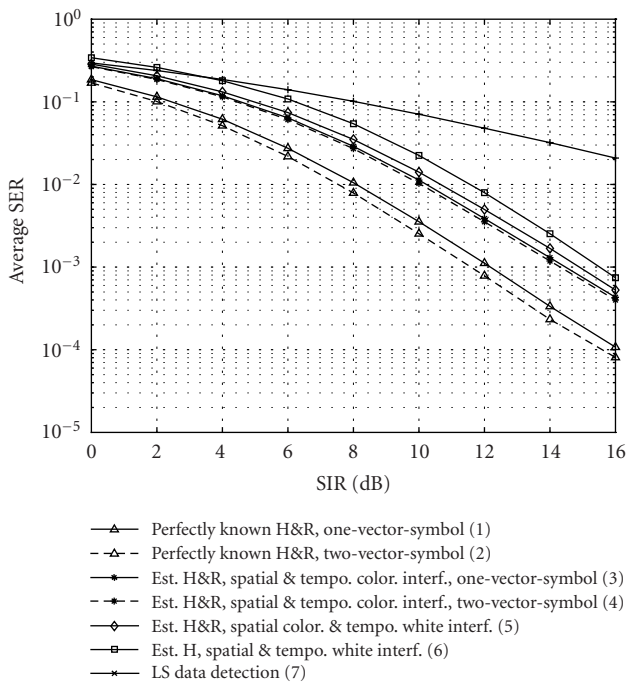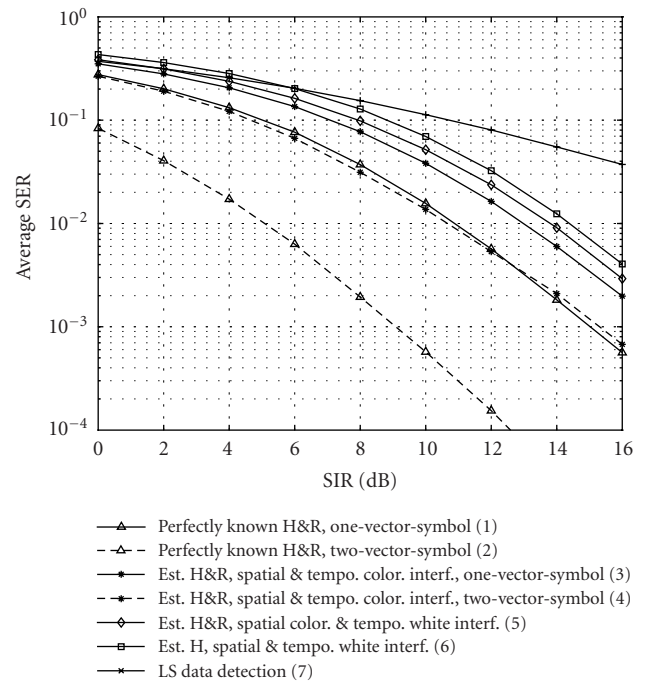

Figure 2: Average SER versus SIR with $N_t = N_r = 5$, $L = 6$, and training length $4N_t$ under independent Rayleigh fading. Both the desired and the interfering users are at the same data rate.

Figure 3: Average SER versus SIR with $N_t = N_r = 5$, $L = 6$, and training length $2N_t$ under independent Rayleigh fading. The data rate of the desired user is twice that of the interfering user.

Figure 4: Average SER versus SIR with $N_t = N_r = 5$, $L = 6$, and training length $4N_t$ under independent Rayleigh fading. The data rate of the desired user is twice that of the interfering user.
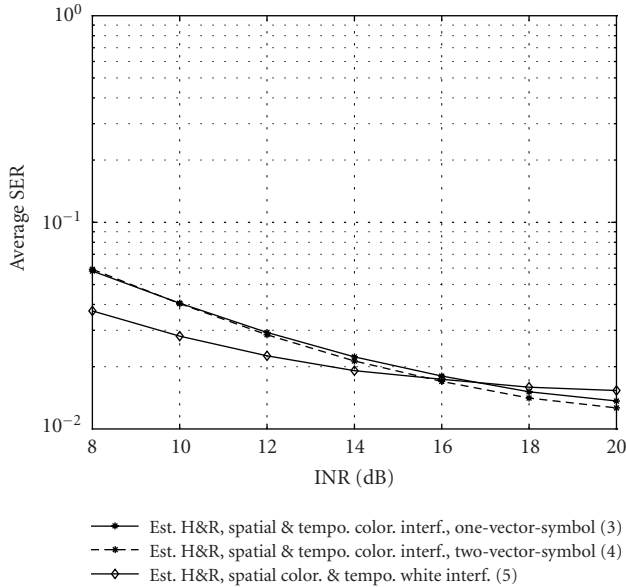
FIGURE 5: Average SER versus INR with $N_t = N_r = 5$, $L = 6$, SIR=10 dB, and training length $4N_t$ under independent Rayleigh fading. Both the desired and the interfering users are at the same data rate.

FIGURE 6: Average SER versus INR with $N_t = N_r = 5$, $L = 6$, SIR=10 dB, and training length $4N_t$ under independent Rayleigh fading. The data rate of the desired user is twice that of the interfering user.

known temporal interference correlation in channel estimation and data detection. About 3.5 dB of the gain is due to exploiting temporal interference correlation in channel estimation and data detection.

5.4.3. Effect of model mismatch

With one interferer and a finite SNR, the interference-plus-noise statistics can only be approximately modelled using a Kronecker product. Here, we investigate when this approximation breaks down. We model thermal noise as a zero-mean circularly symmetric complex Gaussian vector with covariance matrix $\sigma^2 \mathbf{I}_{N_r}$, that is, independent from antenna to antenna, with noise power σ^2 on each antenna. We define (interference-to-noise power ratio) INR $= 10 \log P_I/\sigma^2$, where $P_I = 1$ is used in the simulations. For the case of an interferer at the same data rate and using a training length $4N_t$, we observe in curves 3 and 5 in Figure 5 that, at INRs below 17 dB, taking interference temporal correlation into account appears not to be of benefit. Figure 6 shows the corresponding comparison for the case of the lower-data-rate interferer. In this case, temporal correlation is larger and the decoupled model of interference-plus-noise statistics breaks down at INRs lower than 12 dB.

5.4.4. Effect of exploiting spatial interference-plus-noise correlation

From the above results, temporal interference correlation, even if known, may not result in a performance benefit at lower INRs due to model mismatch. Therefore, we assess the benefit of taking only the *spatial* correlation of interference-plus-noise into account. As a reference, we compare the per-
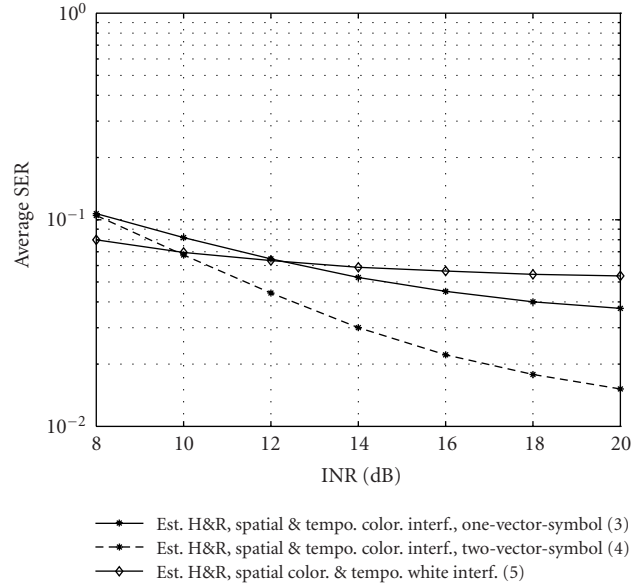
formance to the case of assuming the interference-plus-noise to be spatially white. With total interference power fixed, Figure 7 compares the average SER for one (solid lines) and two (dashed lines) interferers. In the case of two interferers, the interferers have equal power and random relative delays. Both desired and interfering users employ a $(5, 5)$ MIMO link, a total-interference-to-noise ratio of 12 dB, and a training length of $4N_t$. Both the desired and the interfering users operate at the same data rate. Figure 7 shows that for one interferer, there is 1.2 dB gain over a wide range of signal-to-interference-plus-noise ratio (SINRs), by estimating the spatial correlation of interference-plus-noise. For the case of two equal-powered interferers, the corresponding gain in SINR is negligible.

6. CONCLUSIONS

By modelling interference statistics as approximately temporally and spatially separable, we have investigated ML joint estimation of channel parameters and spatial interference correlation matrices. We have assessed the impact of temporal and spatial interference correlation on channel estimation and data detection. For training lengths of at least four times the number of transmitting antennas, gains of around 1 dB are observed by estimating spatial interference correlation. We determine that an additional 0.5 to 3.0 dB in performance gain would result if the known temporal correlation was exploited. For shorter training lengths, however, it is better to estimate only the channel matrix and assume spatially white interference in data detection. One source of temporal correlation occurs where a cochannel interferer operates
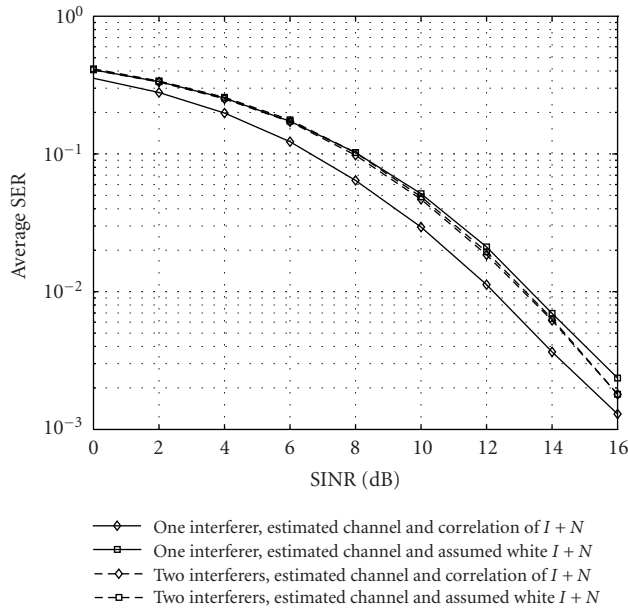
FIGURE 7: The improvement of estimating spatial correlation of interference-plus-noise in practical systems.

at data rate lower than that of the desired user. Exploiting temporal interference correlation in channel estimation was found not to be of benefit. However, if temporal correlation is significant, as in case of lower-data-rate interference, significant performance gains by exploiting temporal interference correlation in data detection are theoretically possible. The minimum INR levels, where separable temporal and interference correlation statistics model was shown to break down and provide no benefit, ranged from 12 or 17 dB, depending on the level of temporal correlation. Of more practical significance, it was shown that at a total INR of 12 dB, 1.2 dB of performance gain can be obtained over a wide range of SINRs by estimating spatial correlation only and neglecting temporal correlation.

ACKNOWLEDGMENT

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[4] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14–16, 1999.

[5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.

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

[7] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Communications Letters*, vol. 5, no. 3, pp. 85–87, 2001.

[8] S. Catreux, P. F. Driessen, and L. J. Greenstein, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Communications Letters*, vol. 4, no. 11, pp. 334–336, 2000.

[9] S. Catreux, P. F. Driessen, and L. J. Greenstein, "Attainable throughput of an interference-limited multiple-input multiple-output (MIMO) cellular system," *IEEE Trans. Communications*, vol. 49, no. 8, pp. 1307–1311, 2001.

[10] Y. Song and S. D. Blostein, "MIMO channel capacity in co-channel interference," in *Proc. 21st Biennial Symposium on Communications*, pp. 220–224, Kingston, Ontario, Canada, June 2002.

[11] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Transactions on Information Theory*, vol. 48, no. 12, pp. 3117–3128, 2002.

[12] R. S. Blum, "MIMO capacity with interference," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 793–801, 2003.

[13] M. Kang and M. S. Alouini, "Performance analysis of MIMO systems with co-channel interference over Rayleigh fading channels," in *Proc. IEEE International Conference on Communications (ICC '02)*, pp. 391–395, New York, NY, USA, April–May 2002.

[14] B. Suard, A. F. Naguib, G. Xu, and A. Paulraj, "Performance of CDMA mobile communication systems using antenna arrays," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '93)*, pp. 153–156, Minneapolis, Minn, USA, April 1993.

[15] T. Marzetta, "BLAST training: estimating channel characteristics for high-capacity space-time wireless," in *Proc. 37th Annual Allerton Conference on Communication, Control, and Computing*, pp. 958–966, Monticello, Ill, USA, September 1999.

[16] B. Hassibi, "A fast square-root implementation for BLAST," in *Proc. 34th Asilomar Conference on Signal, Systems, and Computers*, pp. 1255–1259, Pacific Grove, Calif, USA, October 2000.

[17] J. H. Winters, "Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading," *IEEE Trans. Vehicular Technology*, vol. 42, no. 4, pp. 377–384, 1993.

[18] A. F. Naguib and A. Paulraj, "Performance of wireless CDMA with M-ary orthogonal modulation and cell site antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 9, pp. 1770–1783, 1996.

[19] H. Lütkepohl, *Handbook of Matrices*, John Wiley & Sons, Chichester, 1996.

[20] Y. Song, *Multiple-input multiple-output wireless communication systems with cochannel interference*, Ph.D. thesis, Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario, Canada, 2003.

[21] S. Bäro, G. Bauch, A. Pavlic, and A. Semmler, "Improving BLAST performance using space-time block codes and turbo decoding," in *Proc. Global Telecommunications Conference*, pp. 1067–1071, San Francisco, Calif, USA, November–December 2000.

[22] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice-Hall, Upper Saddle River, NJ, USA, 1996.

[23] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 1995.

[24] W. A. Gardner, *Introduction to Random Processes*, McGraw-Hill, New York, NY, USA, 1990.

[25] G. Long, F. Ling, and J. G. Proakis, "Fractionally-spaced equalizers based on singular value decomposition," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, pp. 1514–1517, New York, NY, USA, April 1988.

Yi Song received her B.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1995 and her M.S. and Ph.D. degrees in electrical engineering from Queen's University, Kingston, Ontario, Canada, in 1998 and 2003, respectively. Her research interests include wireless communications and signal processing for multiple antenna systems.

Steven D. Blostein received his B.S. degree in electrical engineering from Cornell University in 1983, and his M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 1985 and 1988, respectively. He has been on the faculty board of Queen's University in Kingston, Ontario, Canada since 1988, where he currently holds the position of Professor and Head of the Department of Electrical and Computer Engineering. His current interests lie in signal processing for wireless communications as well as detection and estimation theory. He is a Senior Member of the IEEE and a Registered Professional Engineer in Ontario.