PERFORMANCE ANALYSIS OF A SECOND ORDER STATISTICS BASED SOLUTION FOR THE MIMO CHANNEL IDENTIFICATION PROBLEM

João Xavier and Victor Barroso

Instituto Superior Técnico – Instituto de Sistemas e Robótica
Av. Rovisco Pais, 1049–001 Lisboa, Portugal
{jxavier,vab}@isr.ist.utl.pt

ABSTRACT

The CFC$_2$ algorithm is a recently introduced analytical solution for the blind MIMO channel identification problem, provided a certain spectral diversity holds for the stochastic inputs of the MIMO system. Here, we develop a theoretical study to derive the asymptotic performance of the CFC$_2$ algorithm, in terms of mean-square error. Asymptotic normality of the MIMO channel estimate is proved, and the asymptotic error covariance matrix derived. Computer simulation results are included to validate the theoretical expressions.

1. PROBLEM STATEMENT

The closed-form correlative coding (hereafter, CFC$_2$) was introduced in [1, 2]. It provides an analytical solution for the blind MIMO channel identification problem, under an appropriate prefiltering correlative framework which does not increase the transmitted power nor decreases the original data rate. Also, only 2nd order statistics are used. We refer the reader to [1, 2] for details. Here, we analyse the performance of the CFC$_2$ algorithm, in order to assess the asymptotic mean-square error (MSE) of the MIMO channel estimate. For simplicity, we consider here the case of binary sources and zero-mean spatio-temporal white Gaussian noise. The extension to the general case follows easily. The set $\mathcal{T}$ in (5) of [1] is given here by $\mathcal{T} = \{0, 1, \ldots, d\}$ for some integer d.

Notation. We maintain the notation in [1]. In addition, we let $\boldsymbol{I}_{m,n,p} = [\mathbf{0}_{n\times m}\boldsymbol{I}_n\mathbf{0}_{n\times p}]$, $\boldsymbol{S}_{p,n,i} = \boldsymbol{I}_p \otimes \boldsymbol{I}_{0,n,i}$, $\boldsymbol{T}_{p,n,i} = \boldsymbol{I}_p \otimes \boldsymbol{I}_{i,n,0}$ and $\mathbf{1}_{m\times n}$ is the $m \times n$ matrix with all 1's. The symbol $\boxtimes$ denotes the Khatri-Rao product: for $\boldsymbol{A} : m \times n$, $\boldsymbol{B} : p \times n$, $\boldsymbol{A} \boxtimes \boldsymbol{B} = [\boldsymbol{a}_1 \otimes \boldsymbol{b}_1 \cdots \boldsymbol{a}_n \otimes \boldsymbol{b}_n] : mp \times n$, where $\boldsymbol{a}_i, \boldsymbol{b}_i$ denote the ith column of $\boldsymbol{A}$, $\boldsymbol{B}$, respectively. $\boldsymbol{A}^{[k]}$ is the k-fold Kronecker product of $\boldsymbol{A}$: $\boldsymbol{A}^{[k]} = \boldsymbol{A} \otimes \cdots \otimes \boldsymbol{A}$ (k times). For $\boldsymbol{A} : m \times n$, $\mathrm{vec}(\boldsymbol{A}), \mathrm{tr}\{\boldsymbol{A}\}, \|\boldsymbol{A}\|$ represent the vectorization operator, the trace and the Frobenius norm of $\boldsymbol{A}$, respectively. Also, $\boldsymbol{i}_n = \mathrm{vec}(\boldsymbol{I}_n)$ and $\boldsymbol{K}_{n,m}$ is the commutation matrix [6]: for $\boldsymbol{A} : n \times m$, $\boldsymbol{K}_{n,m}\mathrm{vec}(\boldsymbol{A}) = \mathrm{vec}(\boldsymbol{A}^\top)$. For random vectors $\boldsymbol{x}, \boldsymbol{y}$, $\mathtt{E}\{\boldsymbol{x}\}$ is the statistical expectation operator, $\mathtt{corr}\{\boldsymbol{x}\} = \mathtt{E}\{\boldsymbol{x}\boldsymbol{x}^\top\}$ and $\mathtt{cov}\{\boldsymbol{x}, \boldsymbol{y}\} = \mathtt{E}\{\boldsymbol{x}\boldsymbol{y}^\top\} - \mathtt{E}\{\boldsymbol{x}\}\mathtt{E}\{\boldsymbol{y}^\top\}$; $\xrightarrow{\mathrm{d}}, \perp$ stand for convergence in distribution and statistical independency, respectively, and $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Random matrices are viewed as random vectors through the $\mathrm{vec}(\cdot)$ operator. For matrix-to-matrix differentiable mappings $\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X})$, $\boldsymbol{Y} : m \times n$, $\boldsymbol{X} : k \times l$, $\dot{\boldsymbol{F}} : mn \times kl$ denotes the usual derivative by interpreting both $\boldsymbol{Y}$ and $\boldsymbol{X}$ as vectors, through $\mathrm{vec}(\cdot)$.

2. PERFORMANCE ANALYSIS: OUTLINE

We view the CFC$_2$ algorithm as a differentiable mapping $\boldsymbol{\Phi}$ which maps the set of sample correlation matrices

$$\widehat{\boldsymbol{R}}_{\boldsymbol{x}} \equiv \left(\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[0], \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[1], \ldots, \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[d]\right) \tag{1}$$

into the corresponding MIMO channel estimate $\widehat{\boldsymbol{H}}$: $\widehat{\boldsymbol{H}} = \boldsymbol{\Phi}\left(\widehat{\boldsymbol{R}}_{\boldsymbol{x}}\right)$. Keep in mind that

$$\widehat{\boldsymbol{R}}_x[\tau] = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}[t]\,\boldsymbol{x}[t-\tau]^\top$$

denote a sequence of random matrices indexed by T (the number of observed data samples), although this dependency is not explicited here (to avoid cumbersome notation). The same applies to the remaning random variables ($\widehat{\boldsymbol{R}}_{\boldsymbol{x}}, \widehat{\boldsymbol{H}}$, etc). Using the concept of m-dependent sequences [3], we start by establishing the asymptotic normality of $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$:

$$\sqrt{T}\left(\widehat{\boldsymbol{R}}_{\boldsymbol{x}} - \boldsymbol{R}_{\boldsymbol{x}}\right) \xrightarrow{\mathrm{d}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_0), \tag{2}$$

for a certain covariance matrix $\boldsymbol{\Sigma}_0$. Here,

$$\boldsymbol{R}_{\boldsymbol{x}} \equiv (\boldsymbol{R}_{\boldsymbol{x}}[0], \boldsymbol{R}_{\boldsymbol{x}}[1], \ldots, \boldsymbol{R}_{\boldsymbol{x}}[d])$$

denotes the expected value of $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$. By the Delta method [5], we have $\sqrt{T}\left(\widehat{\boldsymbol{H}} - \boldsymbol{H}\right) \xrightarrow{\mathrm{d}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} = \dot{\boldsymbol{\Phi}}(\boldsymbol{R}_{\boldsymbol{x}})\,\boldsymbol{\Sigma}_0\dot{\boldsymbol{\Phi}}(\boldsymbol{R}_{\boldsymbol{x}})^\top$. To obtain the derivative of the mapping $\boldsymbol{\Phi}$, we interpret $\boldsymbol{\Phi}$ as the composition of $n = 4$ (simpler) differentiable mappings, $\boldsymbol{\Phi} = \boldsymbol{\Phi}_4 \circ \boldsymbol{\Phi}_3 \circ \boldsymbol{\Phi}_2 \circ \boldsymbol{\Phi}_1$, each $\boldsymbol{\Phi}_i$ representing an intermediate step of the CFC$_2$ algorithm; $\dot{\boldsymbol{\Phi}}$ is obtained by the chain rule: $\dot{\boldsymbol{\Phi}} = \dot{\boldsymbol{\Phi}}_4 \cdot \dot{\boldsymbol{\Phi}}_3 \cdot \dot{\boldsymbol{\Phi}}_2 \cdot \dot{\boldsymbol{\Phi}}_1$. In the following, we fill in the details of this outline.

3. PERFORMANCE ANALYSIS: DETAILS

First, we prove the asymptotic normality of the sample statistics $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$, and obtain the associated asymptotic covariance matrix $\boldsymbol{\Sigma}_0$ in (2). Then, we focus on each of the differentiable mappings $\boldsymbol{\Phi}_i$, $i = 1, \ldots, n$, and obtain $\dot{\boldsymbol{\Phi}}_i$.

Asymptotic normality of $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$. We say that a random variable $\alpha \sim \mathcal{A}(\kappa)$ if $\mathtt{E}\{\alpha\} = 0$, $\mathtt{E}\{\alpha^2\} = 1$, $\mathtt{E}\{\alpha^3\} = 0$ and $\mathtt{E}\{\alpha^4\} = \kappa$. Also, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^\top \sim \mathcal{A}^n(\kappa)$ if $\alpha_i \perp \alpha_j$, for $i \neq j$, and $\alpha_i \sim \mathcal{A}(\kappa)$. Finally, $\boldsymbol{\alpha} = \left(\boldsymbol{\alpha}_1^\top, \boldsymbol{\alpha}_2^\top, \ldots, \boldsymbol{\alpha}_p^\top\right)^\top \sim \mathcal{A}_p^n(\boldsymbol{\kappa})$, where $\boldsymbol{\kappa} = (\kappa_1, \kappa_2, \ldots, \kappa_p)^\top$ if $\boldsymbol{\alpha}_i \perp \boldsymbol{\alpha}_j$, for $i \neq j$, and $\boldsymbol{\alpha}_i \sim \mathcal{A}^n(\kappa_i)$. We shall need lemma 1 concerning the second and fourth moments of the distribution $\mathcal{A}_p^n(\boldsymbol{\kappa})$. The proof is omitted due to paper length constraints.

Lemma 1. For $\boldsymbol{\alpha} \sim \mathcal{A}_p^n(\boldsymbol{\kappa})$, $\mathtt{E}\{\boldsymbol{\alpha} \otimes \boldsymbol{\alpha}\} = \boldsymbol{P}_p^n(\boldsymbol{i}_p \otimes \boldsymbol{i}_n)$, where $\boldsymbol{P}_p^n = \boldsymbol{I}_p \otimes \boldsymbol{K}_{p,n} \otimes \boldsymbol{I}_n$. Also, $\mathtt{corr}\{\boldsymbol{\alpha} \otimes \boldsymbol{\alpha}\} = \boldsymbol{P}_p^n \boldsymbol{C} \boldsymbol{P}_p^{n\top}$, where

$$\begin{aligned}\boldsymbol{C} = \boldsymbol{I}_{p^2n^2} + \boldsymbol{i}_p\boldsymbol{i}_p^\top \otimes \boldsymbol{i}_n\boldsymbol{i}_n^\top + \boldsymbol{K}_{p,p} \otimes \boldsymbol{K}_{n,n} + \\ \operatorname{diag}\left[(\kappa_1 - 3)\boldsymbol{I}_{pn^2}, (\kappa_2 - 3)\boldsymbol{I}_{pn^2}, \ldots, (\kappa_p - 3)\boldsymbol{I}_{pn^2}\right] \square\end{aligned}$$

For further reference, we let $\boldsymbol{\mu}_p^n = \mathtt{E}\{\boldsymbol{\alpha} \otimes \boldsymbol{\alpha}\}$ and $\boldsymbol{\mathcal{C}}_p^n(\boldsymbol{k}) = \mathtt{corr}\{\boldsymbol{\alpha} \otimes \boldsymbol{\alpha}\}$, for $\boldsymbol{\alpha} \sim \mathcal{A}_p^n(\boldsymbol{\kappa})$. For a scalar signal $\{\alpha[t] : t \in \mathbb{Z}\}$ and given length n, we define the stacking operator

$$\mathcal{S}^n(\alpha[t]) = (\alpha[t], \alpha[t-1], \ldots, \alpha[t-(n-1)])^\top.$$

Correspondingly, for p signals $\{\alpha_i[t] : t \in \mathbb{Z}\}$ and given n, we let

$$\mathcal{S}_p^n(\alpha_1[t], \ldots, \alpha_p[t]) = \left(\mathcal{S}^n(\alpha_1[t])^\top, \ldots, \mathcal{S}^n(\alpha_p[t])^\top\right)^\top.$$

We say that the random multivariate time series $\boldsymbol{\alpha}[t]$ has the distribution $\mathcal{Z}_p^n(\boldsymbol{\kappa})$, for $\boldsymbol{\kappa} = (\kappa_1, \kappa_2, \ldots, \kappa_p)^\top$, written $\boldsymbol{\alpha}[\cdot] \sim \mathcal{Z}_p^n(\boldsymbol{\kappa})$, if $\boldsymbol{\alpha}[t] = \mathcal{S}_p^n(\alpha_1[t], \ldots, \alpha_p[t])$, with $\alpha_i[s] \perp \alpha_j[t]$, for $i \neq j$ or $s \neq t$, and $\alpha_i[t] \sim \mathcal{A}(\kappa_i)$, for all t.

Lemma 2. Let $\boldsymbol{\alpha}[\cdot] \sim \mathcal{Z}_p^n(\boldsymbol{\kappa})$. We have

$$\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^T \boldsymbol{\alpha}[t] \otimes \boldsymbol{\alpha}[t] - \boldsymbol{\mu}_p^n\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_p^n(\boldsymbol{\kappa})\right),$$

where

$$\boldsymbol{\Sigma}_p^n(\boldsymbol{\kappa}) = \boldsymbol{K}_0 + 2\sum_{l=1}^{n-1}\boldsymbol{K}_l - (2n-1)\boldsymbol{\mu}_p^n \boldsymbol{\mu}_p^{n\top},$$

with $\boldsymbol{K}_l = \boldsymbol{S}_{p,n,l}^{[2]} \boldsymbol{\mathcal{C}}_p^{n+l}(\boldsymbol{\kappa}) \boldsymbol{T}_{p,n,l}^{[2]\top}$ $\square$

Proof. By definition, there exist p mutually independent random signals $\alpha_i[t]$, $\alpha_i[s] \perp \alpha_i[t]$, for $s \neq t$, and $\alpha_i[t] \sim \mathcal{A}(\kappa_i)$, such that $\boldsymbol{\alpha}[t] = \mathcal{S}_p^n(\alpha_1[t], \ldots, \alpha_p[t])$. Within the context of this proof we let $\boldsymbol{\alpha}^k[t] = \mathcal{S}_p^k(\alpha_1[t], \ldots, \alpha_p[t])$; thus, $\boldsymbol{\alpha}[t] = \boldsymbol{\alpha}^n[t]$. Let $\lambda[t] = \boldsymbol{u}^T(\boldsymbol{\alpha}[t] \otimes \boldsymbol{\alpha}[t])$, for arbitrary (fixed) $\boldsymbol{u}$. From the given assumptions, it is clear that $\lambda[t]$ is a stationary $(n-1)$-dependent sequence [3]. Thus,

$$\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^T \lambda[t] - \mu\right) \xrightarrow{\mathrm{d}} \mathcal{N}(0, \nu^2),$$

$\mu = \mathtt{E}\{\lambda[t]\}$, $\nu^2 = \nu_0 + 2\sum_{l=1}^{n-1}\nu_l$, and $\nu_l = \mathtt{cov}\{\lambda[t], \lambda[t-l]\}$, for $l = 0, 1, \ldots, n-1$ [4]. Since $\boldsymbol{\alpha}[t] \sim \mathcal{A}_p^n(\boldsymbol{\kappa})$, we have $\mu = \boldsymbol{\theta}^\top \boldsymbol{\mu}_p^n$. Using the identities

$$\boldsymbol{\alpha}[t] = \boldsymbol{S}_{p,n,l}\,\boldsymbol{\alpha}^{n+l}[t] \qquad \boldsymbol{\alpha}[t-l] = \boldsymbol{T}_{p,n,l}\,\boldsymbol{\alpha}^{n+l}[t],$$

we may write

$$\begin{aligned}\boldsymbol{\alpha}[t] \otimes \boldsymbol{\alpha}[t] &= \boldsymbol{S}_{p,n,l}^{[2]}\left(\boldsymbol{\alpha}^{n+l}[t] \otimes \boldsymbol{\alpha}^{n+l}[t]\right) \\ \boldsymbol{\alpha}[t-l] \otimes \boldsymbol{\alpha}[t-l] &= \boldsymbol{T}_{p,n,l}^{[2]}\left(\boldsymbol{\alpha}^{n+l}[t] \otimes \boldsymbol{\alpha}^{n+l}[t]\right).\end{aligned}$$

Thus,

$$\begin{aligned}&\mathtt{E}\{\lambda[t]\lambda[t-l]\} \\ &\quad= \boldsymbol{u}^\top \boldsymbol{S}_{n,p,l}^{[2]}\,\mathtt{corr}\left\{\boldsymbol{\alpha}^{n+l}[t] \otimes \boldsymbol{\alpha}^{n+l}[t]\right\}\boldsymbol{T}_{n,p,l}^{[2]\top}\boldsymbol{u} \\ &\quad= \boldsymbol{u}^\top \boldsymbol{K}_l \boldsymbol{u},\end{aligned}$$

for $l = 0, 1, \ldots, n-1$. Finally

$$\begin{aligned}\nu^2 &= \mathtt{E}\left\{\lambda[t]^2\right\} + 2\sum_{l=1}^{n-1}\mathtt{E}\left\{\lambda[t]\lambda[t-l]^\top\right\} - (2n-1)\mu^2 \\ &= \boldsymbol{u}^\top \boldsymbol{\Sigma}_p^n(\boldsymbol{\kappa})\boldsymbol{u}.\end{aligned}$$

Since $\boldsymbol{u}$ was chosen arbitrarily, the Cramér-Wold device [5] concludes the proof $\square$

Rewrite the data model of equation (2) in [1] as

$$\boldsymbol{x}[t] = \underbrace{\left[\widetilde{\boldsymbol{H}}_1\,\widetilde{\boldsymbol{H}}_2 \cdots \widetilde{\boldsymbol{H}}_P\;\sigma\boldsymbol{I}_N\right]}_{\boldsymbol{\mathcal{H}}}\underbrace{\begin{bmatrix}\boldsymbol{a}_1[t] \\ \vdots \\ \boldsymbol{a}_P[t] \\ \boldsymbol{b}[t]\end{bmatrix}}_{\boldsymbol{\theta}[t]}, \qquad (3)$$

where $\boldsymbol{b}[t] = \boldsymbol{n}[t]/\sigma = (b_1[t], \ldots, b_N[t])^\top$, $\widetilde{\boldsymbol{H}}_p = \boldsymbol{H}_p\boldsymbol{G}_p$ and $\boldsymbol{G}_p : \tau_p \times (\tau_p + \gamma_p - 1)$ is a Toeplitz matrix with first line given by

$$\left(g_p[0], \ldots, g_p[\gamma_p - 1], \mathbf{0}_{1\times(\tau_p - 1)}\right).$$

Recall from equation (3) in [1] that $g_p[\cdot]$ denotes the correlative filter associated with the pth user. In (3),

$$\boldsymbol{a}_p[t] = (a_p[t], \ldots, a_p[t - L_p - 1])^\top,$$

$L_p \equiv \tau_p + \gamma_p - 1$, where the scalar signal $a_p[t]$ denotes the (unfiltered) i.i.d. information symbol sequence emitted by the pth source. With $\boldsymbol{x}[t]$ as in (3) and $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$ as in (1), we have $\hat{\boldsymbol{r}}_{\boldsymbol{x}} = \operatorname{vec}\left(\widehat{\boldsymbol{R}}_{\boldsymbol{x}}\right) = \left(\boldsymbol{I}_{d+1} \otimes \boldsymbol{\mathcal{H}}^{[2]}\right)\hat{\boldsymbol{r}}_{\boldsymbol{\theta}}$, where

$$\hat{\boldsymbol{r}}_{\boldsymbol{\theta}} = \frac{1}{T}\sum_{t=1}^T\left[\begin{array}{ccc}(\boldsymbol{\theta}[t] \otimes \boldsymbol{\theta}[t])^\top & \cdots & (\boldsymbol{\theta}[t-d] \otimes \boldsymbol{\theta}[t])^\top\end{array}\right]^\top.$$

But, $\boldsymbol{\theta}[t - \tau] = \boldsymbol{\mathcal{I}}[\tau]\boldsymbol{\alpha}[t]$, for $0 \leq \tau \leq d$,

$$\boldsymbol{\alpha}[t] = \mathcal{S}_{P+N}^{L+d}(a_1[t], \ldots, a_P[t], b_1[t], \ldots, b_N[t]),$$

$L \equiv \max\{L_1, \ldots, L_P\}$, and

$$\begin{aligned}\boldsymbol{\mathcal{I}}[\tau] &= \operatorname{diag}[\boldsymbol{\mathcal{I}}_a[\tau], \boldsymbol{\mathcal{I}}_b[\tau]] \\ \boldsymbol{\mathcal{I}}_a[\tau] &= \operatorname{diag}[\boldsymbol{I}_{\tau,L_1,L+d-L_1-\tau}, \ldots, \boldsymbol{I}_{\tau,L_P,L+d-L_P-\tau}] \\ \boldsymbol{\mathcal{I}}_b[\tau] &= \boldsymbol{I}_N \otimes \boldsymbol{I}_{\tau,1,L+d-1-\tau}.\end{aligned}$$

Thus, $\hat{\boldsymbol{r}}_{\boldsymbol{x}} = \boldsymbol{M}\hat{\boldsymbol{r}}_{\boldsymbol{\alpha}}$, where $\boldsymbol{M} = \left(\boldsymbol{I}_{d+1} \otimes \boldsymbol{\mathcal{H}}^{[2]}\right)\boldsymbol{\mathcal{I}}$,

$$\boldsymbol{\mathcal{I}} = \left[\begin{array}{ccc}(\boldsymbol{\mathcal{I}}[0] \otimes \boldsymbol{\mathcal{I}}[0])^\top & \cdots & (\boldsymbol{\mathcal{I}}[d] \otimes \boldsymbol{\mathcal{I}}[0])^\top\end{array}\right]^\top,$$

and $\hat{\boldsymbol{r}}_{\boldsymbol{\alpha}} = \frac{1}{T}\sum_{t=1}^T \boldsymbol{\alpha}[t] \otimes \boldsymbol{\alpha}[t]$. Applying Lemma 2 to $\boldsymbol{\alpha}[\cdot] \sim \mathcal{Z}_{P+N}^{L+d}(\boldsymbol{\kappa})$, where $\boldsymbol{\kappa} = (\mathbf{1}_{1\times P}, 3\mathbf{1}_{1\times N})^\top$ (due to our assumptions of P binary sources and i.i.d. Gaussian entries in $\boldsymbol{n}[t]$), and by the Delta method [5], we obtain $\sqrt{T}(\widehat{\boldsymbol{r}}_{\boldsymbol{x}} - \boldsymbol{r}_{\boldsymbol{x}}) \xrightarrow{\mathrm{d}} \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{r}_{\boldsymbol{x}} = \mathrm{vec}(\boldsymbol{R}_{\boldsymbol{x}})$, $\boldsymbol{\Sigma}_0 = \boldsymbol{M}\boldsymbol{\Sigma}_{P+N}^{L+d}(\boldsymbol{\kappa})\boldsymbol{M}^\top$.

Mapping $\boldsymbol{\Phi}_1$. We define $\boldsymbol{\Phi}_1$ as the mapping

$$\left(\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[0], \ldots, \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[d]\right) \overset{\boldsymbol{\Phi}_1}{\mapsto} \left(\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{W}}, \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[1], \ldots, \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[d]\right),$$

where $\widehat{\boldsymbol{\lambda}} = \left(\hat{\lambda}_1, \ldots, \hat{\lambda}_M\right)^\top$, $M = \tau_1 + \cdots + \tau_P$, $\hat{\lambda}_i \geq \hat{\lambda}_j$ for $i < j$, and $\widehat{\boldsymbol{W}} = [\,\widehat{\boldsymbol{w}}_1 \cdots \widehat{\boldsymbol{w}}_M\,]$, denote the M biggest eigenvalues and respective eigenvectors of $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[0]$. Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_M)^\top$ and $\boldsymbol{W} = [\,\boldsymbol{w}_1 \cdots, \boldsymbol{w}_M\,]$ denote the counterparts of $\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{W}}$ in $\boldsymbol{R}_{\boldsymbol{x}}[0]$, and assume (for simplicity) that $\lambda_i > \lambda_j$, for $i < j$, i.e., the M biggest eigenvalues of $\boldsymbol{R}_{\boldsymbol{x}}[0]$ are distinct. Then, the functions $\lambda_i(\cdot)$ and $\boldsymbol{w}_i(\cdot)$ which extract the ith biggest eigenvalue of $\boldsymbol{X}$ and associated eigenvector, respectively, are differentiable in a neighborhood of $\boldsymbol{R}_{\boldsymbol{x}}[0]$, for $i = 1, \ldots, M$ [6]; the derivatives at $\boldsymbol{R}_{\boldsymbol{x}}[0]$ are given by

$$\begin{aligned} \dot{\lambda}_i &= \boldsymbol{w}_i^\top \otimes \boldsymbol{w}_i \\ \dot{\boldsymbol{w}}_i &= \boldsymbol{w}_i^\top \otimes (\lambda_i \boldsymbol{I}_N - \boldsymbol{R}_{\boldsymbol{x}}[0])^\dagger. \end{aligned}$$

Using these results we can write the derivative of $\boldsymbol{\Phi}_1$ at $\boldsymbol{R}_{\boldsymbol{x}}[0]$ as

$$\dot{\boldsymbol{\Phi}}_1 = \begin{bmatrix} (\boldsymbol{W} \boxtimes \boldsymbol{W})^\top & \boldsymbol{0} \\ \boldsymbol{w}_1^\top \otimes (\lambda_1 \boldsymbol{I}_N - \boldsymbol{R}_{\boldsymbol{x}}[0])^\dagger & \boldsymbol{0} \\ \vdots & \vdots \\ \boldsymbol{w}_M^\top \otimes (\lambda_M \boldsymbol{I}_N - \boldsymbol{R}_{\boldsymbol{x}}[0])^\dagger & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{dN^2} \end{bmatrix}.$$

Remark that, with the definition of $\widehat{\boldsymbol{\lambda}}$ and $\widehat{\boldsymbol{W}}$ as above, the matrix

$$\widehat{\boldsymbol{G}} = \widehat{\boldsymbol{W}}\left(\widehat{\boldsymbol{\Lambda}} - \sigma^2 \boldsymbol{I}_N\right)^{1/2}, \tag{4}$$

where $\widehat{\boldsymbol{\Lambda}} = \mathrm{diag}\left[\hat{\lambda}_1, \ldots, \hat{\lambda}_M\right]$, corresponds to the estimate of $\boldsymbol{G}$ in equation (6) of [1].

Mapping $\boldsymbol{\Phi}_2$. We let $\boldsymbol{\Phi}_2$ be the mapping defined by

$$\left(\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{W}}, \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[1], \ldots, \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[d]\right) \overset{\boldsymbol{\Phi}_2}{\mapsto} \left(\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{W}}, \widehat{\boldsymbol{B}}[1], \ldots, \widehat{\boldsymbol{B}}[d]\right),$$

where $\widehat{\boldsymbol{B}}[\tau] = \widehat{\boldsymbol{G}}^\dagger \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[\tau]\widehat{\boldsymbol{G}}^{\dagger\top}$, corresponds to the estimate of $\boldsymbol{B}[\tau]$ (7) of [1]; $\widehat{\boldsymbol{B}}[\tau]$ depends on $\widehat{\boldsymbol{\lambda}}$, $\widehat{\boldsymbol{W}}$ and $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[\tau]$, as $\widehat{\boldsymbol{B}}[\tau] = \boldsymbol{F}(\widehat{\boldsymbol{\lambda}})\widehat{\boldsymbol{W}}^\top \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[\tau]\widehat{\boldsymbol{W}}\boldsymbol{F}(\widehat{\boldsymbol{\lambda}})$, $\boldsymbol{F}(\boldsymbol{\lambda}) \equiv \mathrm{diag}[g(\lambda_1), \ldots, g(\lambda_M)]$ and the function $g(x) = 1/\sqrt{(x - \sigma^2)}$. The derivative of $\boldsymbol{F}$ at $\boldsymbol{\lambda}$ is easily seen to be given by

$$\boldsymbol{\mathcal{G}}(\boldsymbol{\lambda}) \equiv \dot{\boldsymbol{F}} = (\boldsymbol{I}_M \boxtimes \boldsymbol{I}_M)\,\mathrm{diag}[\dot{g}(\lambda_1), \ldots, \dot{g}(\lambda_M)],$$

where $\dot{g}(x) = -\frac{1}{2}(x - \sigma^2)^{-3/2}$. On the other hand, holding $\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[\tau]$ fixed, and defining $\boldsymbol{C}[\tau] = \widehat{\boldsymbol{W}}^\top \widehat{\boldsymbol{R}}_{\boldsymbol{x}}[0]\widehat{\boldsymbol{W}}$, we can compute the derivative of $\widehat{\boldsymbol{B}}[\tau]$ with respect to $\boldsymbol{F}$:

$$d\widehat{\boldsymbol{B}}[\tau] = d\boldsymbol{F}\boldsymbol{C}[\tau]\boldsymbol{F} + \boldsymbol{F}\boldsymbol{C}[\tau]d\boldsymbol{F};$$

thus, letting $\widehat{\boldsymbol{b}}[\tau] = \mathrm{vec}\left(\widehat{\boldsymbol{B}}[\tau]\right)$, we have

$$d\widehat{\boldsymbol{b}}[\tau] = \left(\boldsymbol{F}\boldsymbol{C}[\tau]^\top \otimes \boldsymbol{I}_M + \boldsymbol{I}_M \otimes \boldsymbol{F}\boldsymbol{C}[\tau]\right) d\boldsymbol{f},$$

where $\boldsymbol{f} = \mathrm{vec}(\boldsymbol{F})$. By the chain rule, the derivative of $\widehat{\boldsymbol{B}}[\tau]$ with respect to $\widehat{\boldsymbol{\lambda}}$ at

$$(\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{B}[1], \ldots, \boldsymbol{B}[d]) = \boldsymbol{\Phi}_1(\boldsymbol{R}_{\boldsymbol{x}}[0], \ldots, \boldsymbol{R}_{\boldsymbol{x}}[d]) \tag{5}$$

is given by

$$\boldsymbol{\partial}_{\widehat{\boldsymbol{\lambda}}}\widehat{\boldsymbol{B}}[\tau] = \left(\boldsymbol{F}(\boldsymbol{\lambda})\boldsymbol{C}[\tau]^\top \otimes \boldsymbol{I}_M + \boldsymbol{I}_M \otimes \boldsymbol{F}(\boldsymbol{\lambda})\boldsymbol{C}[\tau]\right)\boldsymbol{\mathcal{G}}(\boldsymbol{\lambda}),$$

where $\boldsymbol{C}[\tau] = \boldsymbol{W}^\top \boldsymbol{R}_{\boldsymbol{x}}[\tau]\boldsymbol{W}$. Using the same strategy, i.e., holding the pairs $\widehat{\boldsymbol{\lambda}}$, $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[\tau]$ and $\widehat{\boldsymbol{\lambda}}$, $\widehat{\boldsymbol{W}}$ fixed, and applying standard calculus rules, we obtain the derivatives of $\widehat{\boldsymbol{B}}[\tau]$ with respect to $\widehat{\boldsymbol{W}}$ and $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[\tau]$, at the point (5), respectively:

$$\begin{aligned} \boldsymbol{\partial}_{\widehat{\boldsymbol{W}}}\widehat{\boldsymbol{B}}[\tau] &= \left(\boldsymbol{F}(\boldsymbol{\lambda})\boldsymbol{W}^\top \boldsymbol{R}_{\boldsymbol{x}}[\tau]^\top \otimes \boldsymbol{F}(\boldsymbol{\lambda})\right)\boldsymbol{K}_{N,M} + \\ &\quad \boldsymbol{F}(\boldsymbol{\lambda}) \otimes \boldsymbol{F}(\boldsymbol{\lambda})\boldsymbol{W}^\top \boldsymbol{R}_{\boldsymbol{x}}[\tau] \\ \boldsymbol{\partial}_{\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[\tau]}\widehat{\boldsymbol{B}}[\tau] &= \left(\boldsymbol{F}(\boldsymbol{\lambda})\boldsymbol{W}^\top\right)^{[2]}. \end{aligned}$$

From all these considerations, the derivative of $\boldsymbol{\Phi}_2$ at (5) is

$$\dot{\boldsymbol{\Phi}}_2 = \begin{bmatrix} \boldsymbol{I}_M & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{NM} & \boldsymbol{0} \\ \boldsymbol{\partial}_{\widehat{\boldsymbol{\lambda}}}\widehat{\boldsymbol{B}}[1] & \boldsymbol{\partial}_{\widehat{\boldsymbol{W}}}\widehat{\boldsymbol{B}}[1] & \boldsymbol{e}_1^\top \otimes \boldsymbol{\partial}_{\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[1]}\widehat{\boldsymbol{B}}[1] \\ \vdots & \vdots & \vdots \\ \boldsymbol{\partial}_{\widehat{\boldsymbol{\lambda}}}\widehat{\boldsymbol{B}}[d] & \boldsymbol{\partial}_{\widehat{\boldsymbol{W}}}\widehat{\boldsymbol{B}}[d] & \boldsymbol{e}_d^\top \otimes \boldsymbol{\partial}_{\widehat{\boldsymbol{R}}_{\boldsymbol{x}}[d]}\widehat{\boldsymbol{B}}[d] \end{bmatrix},$$

where $\boldsymbol{e}_i$ denotes the ith column of $\boldsymbol{I}_d$.

Mapping $\boldsymbol{\Phi}_3$. The mapping $\boldsymbol{\Phi}_3$ performs the operation

$$\left(\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{W}}, \widehat{\boldsymbol{B}}[1], \ldots, \widehat{\boldsymbol{B}}[d]\right) \overset{\boldsymbol{\Phi}_3}{\mapsto} \left(\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{W}}, \widehat{\boldsymbol{U}}\right),$$

where $\widehat{\boldsymbol{U}}$ corresponds to the estimate of the matrix $\boldsymbol{U}$ at the end of step 2 of the CFC_2 algorithm in [1]. Thus, in $\widehat{\boldsymbol{U}} = \left[\widehat{\boldsymbol{U}}_1 \cdots \widehat{\boldsymbol{U}}_P\right]$, the submatrix $\widehat{\boldsymbol{U}}_p : M \times \tau_p$ is the non-zero solution of the homogeneous linear system in the unknwown $\boldsymbol{X} : M \times \tau_p$

$$\left\{\begin{aligned} \widehat{\boldsymbol{B}}[1]\boldsymbol{X} - \boldsymbol{X}\boldsymbol{A}_p[1] &= \boldsymbol{0} \\ \widehat{\boldsymbol{B}}[1]^\top \boldsymbol{X} - \boldsymbol{X}\boldsymbol{A}_p[1]^\top &= \boldsymbol{0} \\ &\vdots \\ \widehat{\boldsymbol{B}}[d]\boldsymbol{X} - \boldsymbol{X}\boldsymbol{A}_p[d] &= \boldsymbol{0} \\ \widehat{\boldsymbol{B}}[d]^\top \boldsymbol{X} - \boldsymbol{X}\boldsymbol{A}_p[d]^\top &= \boldsymbol{0} \end{aligned}\right.,$$

scaled to norm $\sqrt{\tau_p}$, see [1]; $\boldsymbol{A}_p[\tau]$ was defined in equation (4) of [1], and depends solely on the correlative filter $g_p[\cdot]$. It is straighforward to see that $\widehat{\boldsymbol{u}}_p = \mathrm{vec}\left(\widehat{\boldsymbol{U}}_p\right)$ can be obtained as $\widehat{\boldsymbol{u}}_p = \sqrt{\tau_p}\widehat{\boldsymbol{v}}_p$, where $\widehat{\boldsymbol{v}}_p$ denotes the eigenvector associated with the smallest eigenvalue of $\widehat{\boldsymbol{S}}_p = \widehat{\boldsymbol{T}}_p^\top \widehat{\boldsymbol{T}}_p$, where

$$\widehat{\boldsymbol{T}}_p = \left[\ \widehat{\boldsymbol{T}}_p[1]^\top \quad \widehat{\boldsymbol{T}}_p[2]^\top \quad \cdots \quad \widehat{\boldsymbol{T}}_p[d]^\top\ \right]^\top,$$

and

$$\begin{aligned} \widehat{\boldsymbol{T}}_p[\tau] &= \begin{bmatrix} \widehat{\boldsymbol{T}}_p^{(1)}[\tau] \\ \widehat{\boldsymbol{T}}_p^{(2)}[\tau] \end{bmatrix} \\ \widehat{\boldsymbol{T}}_p^{(1)}[\tau] &= \boldsymbol{I}_{\tau_p} \otimes \widehat{\boldsymbol{B}}[\tau] - \boldsymbol{A}_p[\tau]^\top \otimes \boldsymbol{I}_M \\ \widehat{\boldsymbol{T}}_p^{(2)}[\tau] &= \boldsymbol{I}_{\tau_p} \otimes \widehat{\boldsymbol{B}}[\tau]^\top - \boldsymbol{A}_p[\tau] \otimes \boldsymbol{I}_M. \end{aligned}$$

Let

$$(\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{B}[1], \ldots, \boldsymbol{B}[d]) = (\boldsymbol{\Phi}_2 \circ \boldsymbol{\Phi}_1)(\boldsymbol{R_x}[0], \ldots, \boldsymbol{R_x}[d]), \tag{6}$$

and define $\boldsymbol{T}_p^{(1)}[\tau], \boldsymbol{T}_p^{(2)}[\tau], \boldsymbol{T}_p[\tau], \boldsymbol{T}_p, \boldsymbol{S}_p, \boldsymbol{v}_p$ as the direct counterparts of $\widehat{\boldsymbol{T}}_p^{(1)}[\tau], \widehat{\boldsymbol{T}}_p^{(2)}[\tau], \widehat{\boldsymbol{T}}_p[\tau], \widehat{\boldsymbol{T}}_p, \widehat{\boldsymbol{S}}_p, \widehat{\boldsymbol{v}}_p$, i.e., the same definitions but without the $\widehat{(\cdot)}$. It can be seen [1, 2] that the smallest eigenvalue of $\boldsymbol{S}_p$ is 0 with multiplicity 1. Thus, from [6] again, $\boldsymbol{v}_p(\cdot)$ is a differentiable function in a neighborhood of $\boldsymbol{S}_p$ and, at the point (6), we have

$$\boldsymbol{\partial}_{\widehat{\boldsymbol{S}}_p} \widehat{\boldsymbol{v}}_p = \boldsymbol{v}_p^\top \otimes (-\boldsymbol{S}_p)^\dagger.$$

Also, trivially

$$\boldsymbol{\partial}_{\widehat{\boldsymbol{T}}_p} \widehat{\boldsymbol{S}}_p = \left(\boldsymbol{T}_p^\top \otimes \boldsymbol{I}_{M\tau_p}\right) \boldsymbol{K}_{2dM\tau_p, M\tau_p} + \boldsymbol{I}_{M\tau_p} \otimes \boldsymbol{T}_p^\top.$$

Letting $\widehat{\boldsymbol{B}} = \left(\widehat{\boldsymbol{B}}[1], \ldots, \widehat{\boldsymbol{B}}[d]\right)$ we have, after some calculus,

$$\boldsymbol{\partial}_{\widehat{\boldsymbol{B}}} \widehat{\boldsymbol{T}}_p = \left[\sum_{i=1}^{2d} \left(\boldsymbol{e}_i \otimes \boldsymbol{I}_{M\tau_p}\right)^\top \otimes \left(\boldsymbol{e}_i \otimes \boldsymbol{I}_{M\tau_p}\right)\right] \boldsymbol{I}_d \otimes \boldsymbol{\nabla},$$

where $\boldsymbol{e}_i$ is the ith column of $\boldsymbol{I}_{2d}$, $\boldsymbol{\nabla} = \left[\boldsymbol{\mathcal{P}}^\top \; (\boldsymbol{\mathcal{P}} \boldsymbol{K}_{M,M})^\top\right]^\top$, and $\boldsymbol{\mathcal{P}} = \left[\left(\boldsymbol{I}_M \otimes \boldsymbol{K}_{M,\tau_p}\right) \operatorname{vec}\left(\boldsymbol{I}_{\tau_p} \otimes \boldsymbol{I}_M\right)\right] \otimes \boldsymbol{I}_M$. All these partial results can be collected by the chain rule to yield the derivative of $\widehat{\boldsymbol{u}}_p$ with respect to $\widehat{\boldsymbol{B}}$,

$$\boldsymbol{\nabla}_p \equiv \boldsymbol{\partial}_{\widehat{\boldsymbol{B}}} \widehat{\boldsymbol{u}}_p = \sqrt{\tau_p}\, \boldsymbol{\partial}_{\widehat{\boldsymbol{B}}} \widehat{\boldsymbol{T}}_p \cdot \boldsymbol{\partial}_{\widehat{\boldsymbol{T}}_p} \widehat{\boldsymbol{S}}_p \cdot \boldsymbol{\partial}_{\widehat{\boldsymbol{S}}_p} \widehat{\boldsymbol{v}}_p.$$

Finally, the derivative of the mapping $\boldsymbol{\Phi}_3$ at (6) is

$$\dot{\boldsymbol{\Phi}}_3 = \begin{bmatrix} \boldsymbol{I}_M & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{NM} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\nabla}_1 \\ \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\nabla}_P \end{bmatrix}.$$

Mapping $\boldsymbol{\Phi}_4$. The mapping $\boldsymbol{\Phi}_4$ is defined by

$$\left(\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{W}}, \widehat{\boldsymbol{U}}\right) \stackrel{\boldsymbol{\Phi}_4}{\to} \widehat{\boldsymbol{H}}.$$

Here, $\widehat{\boldsymbol{H}} = \widehat{\boldsymbol{G}} \widehat{\boldsymbol{U}} \boldsymbol{R_s}[0]^{-1/2}$, where $\boldsymbol{R_s}[0]$ depends only on the correlative fiters, see [1]; $\widehat{\boldsymbol{G}}$ was defined in (4). The derivatives of $\widehat{\boldsymbol{H}}$ at the point

$$(\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{U}) = (\boldsymbol{\Phi}_3 \circ \boldsymbol{\Phi}_2 \circ \boldsymbol{\Phi}_1)(\boldsymbol{R_x}[0], \ldots, \boldsymbol{R_x}[d]) \tag{7}$$

are easily obtained as

$$\begin{aligned} \boldsymbol{\partial}_{\widehat{\boldsymbol{\lambda}}} \widehat{\boldsymbol{H}} &= (\boldsymbol{I}_M \boxtimes \boldsymbol{I}_M) \operatorname{diag}[h(\lambda_1), \ldots, h(\lambda_M)] \\ \boldsymbol{\partial}_{\widehat{\boldsymbol{W}}} \widehat{\boldsymbol{H}} &= \boldsymbol{R_s}[0]^{-1/2} \boldsymbol{U}^\top \otimes \boldsymbol{O}(\boldsymbol{\lambda}) \\ \boldsymbol{\partial}_{\widehat{\boldsymbol{U}}} \widehat{\boldsymbol{H}} &= \boldsymbol{R_s}[0]^{-1/2} \otimes \boldsymbol{U} \boldsymbol{O}(\boldsymbol{\lambda}), \end{aligned}$$

where $h(x) = 1/(2\sqrt{x - \sigma^2})$, and $\boldsymbol{O}(\boldsymbol{\lambda}) = \operatorname{diag}[\lambda_1, \ldots, \lambda_m] - \sigma^2 \boldsymbol{I}_M$. The derivative of $\boldsymbol{\Phi}_4$ at (7) is

$$\dot{\boldsymbol{\Phi}}_4 = \begin{bmatrix} \boldsymbol{\partial}_{\widehat{\boldsymbol{\lambda}}} \widehat{\boldsymbol{H}} & \boldsymbol{\partial}_{\widehat{\boldsymbol{W}}} \widehat{\boldsymbol{H}} & \boldsymbol{\partial}_{\widehat{\boldsymbol{U}}} \widehat{\boldsymbol{H}} \end{bmatrix}.$$

4. COMPUTER SIMULATIONS

Numerical experiments were conducted to validate the theoretical expressions. Respecting the notation in [1], we considered a scenario with $P = 2$ binary users. The entries of a MIMO channel matrix $\boldsymbol{H} : N \times (\tau_1 + \tau_2)$ $(N = 4, \tau_1 = 2, \tau_2 = 1)$ were randomly generated as i.i.d. samples from a zero-mean unit-power Gaussian distribution. User 1 uses no correlative filter, i.e., $\gamma_1 = 1$ $(g_1[0] = 1)$; user 2 uses a correlative filter with $\gamma_2 = 2$ taps $(g_2[0] = 1/\sqrt{2}, g_2[1] = -1/\sqrt{2})$; $\boldsymbol{n}[t]$ is taken as a zero-mean spatio-temporal white Gaussian noise with variance σ^2. The SNR $= \operatorname{tr}\left(\boldsymbol{H}^\top \boldsymbol{H}\right)/N\sigma^2$ was fixed at 10 dB. The number of samples T was varied between $T_{\min} = 200$ and $T_{\max} = 1000$ in steps of $T_{\text{step}} = 50$ samples. For each T, $K = 1000$ independent Monte-Carlo runs were simulated; for the kth run, CFC_2 algorithm produced $\widehat{\boldsymbol{H}}_T^{(k)}$ and the square-error $\epsilon_T^{(k)} = \left\|\widehat{\boldsymbol{H}}_T^{(k)} - \widehat{\boldsymbol{H}}\right\|^2$ was recorded. The mean of these $K = 1000$ errors is denoted by $\bar{\epsilon}_T$ and is the estimate of MSE of the CFC_2's MIMO channel estimate, for T samples. Figure 1 shows the results obtained numerically for $\bar{\epsilon}_T$ (dashed line) against the asymptotic theoretical expressions derived in section 3 (solid line). The curves show a good agreement.

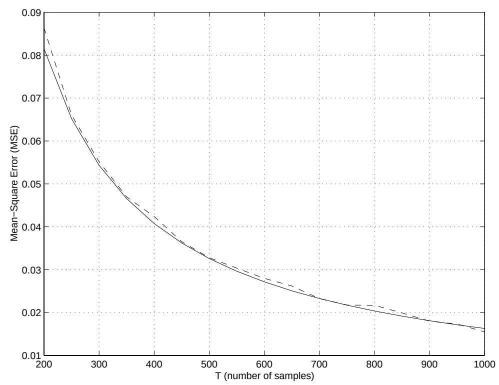

Fig. 1. MSE for the MIMO channel estimate: theoretical (solid) and observed (dashed)

5. REFERENCES

[1] J. Xavier and V. Barroso, "Correlative coding approach for blind identification of MIMO systems," in *Proc. 2nd IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC'99)*, pp. 263–266, May 1999.

[2] J. Xavier, V. Barroso and J. M. F. Moura, "Closed-form blind identification of MIMO channels by correlative coding: an isometry fitting to second order statistics," submitted to *IEEE Transactions on Signal Processing*.

[3] W. Fuller, *Introduction to Statistical Time Series*, John Wiley & Sons.

[4] T. W. Anderson, *The Statistical Analysis of Time Series*, John Wiley & Sons.

[5] P. Sen and J. Singer, *Large Sample Methods in Statistics: An Introduction with Applications*, Chapman & Hall.

[6] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications to Statistics and Econometrics*, Wiley Series in Probability and Statistics.