Estimation of block-fading channels with reduced-rank correlation matrix

O. Simeone[1,2] and U. Spagnolini[2]

(1) CWCSPR, ECE Dept., NJIT, Newark, NJ 07102, USA

(2) Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milano, Italy

***Abstract*— In block-fading transmission, the discrete-time taps of time-varying channels can be generally modelled as a stationary random vector process with unknown and rank-deficient correlation matrix. The model applies to single/ multi-antenna systems for single/multi-carrier modulation. In this paper, we propose a hybrid estimator that sequentially estimates the (deterministic) rank-deficient channel correlation matrix and the (random) channel taps. In addition, we derive a bound on the corresponding estimation error by adapting the Cramér Rao bound (CRB) to the case of mixed deterministic and random parameters (Hybrid CRB). Finally, we assess that the proposed estimator is statistically efficient through comparison with the Hybrid CRB.**

I. INTRODUCTION

In communication systems the block-transmission is routinely employed in order to deal with time variations of wireless channels (see, e.g., [1]). Known training sequences are periodically transmitted with the payload bits so as to aid the receivers to estimate the channel for data detection within each block. The block length is generally optimized to guarantee that the channel variations within the block are negligible, while inter-block variations of fading depend on the time between transmission of successive blocks.

The scope of this paper is estimation of block-fading channels where the inter-block interval is arbitrary but sufficiently large to render ineffective the use of simple averaging technique across consecutive blocks. Moreover, we assume Rayleigh fading and the channel taps be correlated within each block with an unknown and generally rank-deficient correlation matrix. This model complies with assumptions widely made in the context of block transmission, in either a single or multi-antenna scenario for single or multicarrier modulation. Channel correlation is due to the fact that, in general, the same multipath component affects multiple channel taps in either time and/or spatial domain. On the other hand, rank deficiency arises whenever the temporal and/or spatial spread of the multipath channel is not large enough to require the entire signal space for its description. This feature has been often employed in order to obtain an efficient parametrization of the channel for the purpose of channel estimation [2] [3] [4] for single or multicarrier modulation [5]. A simple example of the model is given by a frequency-flat Multiple Input Multiple Output (MIMO) channel with correlated fading at the transmit/ receive antennas [6].

In the general framework described above, this paper proposes a channel estimator that formalizes previously proposed solutions (see, e.g., [2]) by using the frameworks of Method of Moments (MoM) and MMSE estimation. Moreover, it is shown that a performance bound on the channel estimation error can be obtained by evaluating the Hybrid Cramér-Rao Bound (HCRB), a generalization of the conventional CRB to the case where the parameters to be estimated encompass both deterministic and a random components [7]. Finally, the proposed estimators is shown to be able to reach the Hybrid CRB for a number of blocks large enough.

Basic notation: lowercase (uppercase) bold denotes column vector (matrices), $(\cdot)^T$ is the matrix transpose, $(\cdot)^H$ is the Hermitian transposition, $(\cdot)^\dagger$ is the pseudoinverse, $||\mathbf{X}||^2_{\mathbf{A}} = \mathrm{tr}\{\mathbf{X}^H\mathbf{A}\mathbf{X}\}$ is the norm weighted by a positive definite matrix $\mathbf{A}$, $\mathbf{v} = \mathrm{vec}\{\mathbf{V}\}$ is the stacking operator and $\otimes$ is the Kronecker matrix product.

II. SIGNAL MODEL

In block-fading channels, channel estimation with training sequences (known at the receiver) can be reduced to the following linear regression model over K transmitted blocks:

$$\mathbf{y}_k = \mathbf{X}\mathbf{h}_k + \mathbf{n}_k, \quad k = 1, 2, ..., K. \tag{1}$$

The received signal $\mathbf{y}_k \in \mathbb{C}^N$ represents the observations at kth block (possibly on multiple receiving antennas), the $N \times M$ $(N \geq M)$ regression matrix $\mathbf{X}$ contains the known training sequence(s) (transmitted by possibly more than one antenna) arranged into a convolutional matrix independent on k and full column rank (i.e., $rank(\mathbf{X}) = M$), the additive noise is $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$ with known covariance $\mathbf{Q}$ and uncorrelated samples across blocks: $E[\mathbf{n}_k\mathbf{n}^H_{k-m}] = \delta_m\mathbf{Q}$ (δ_m is the Kronecker delta function). The unknown and block-varying channel is arranged into the vector $\mathbf{h}_k \in \mathbb{C}^M$, and it can model either a single antenna link or multiple antennas links. As an example, in a frequency flat MIMO system, vector $\mathbf{h}_k$ contains the channel gains between all the pairs of transmit/receive antennas. Specialization of this model to multi-antenna and multi-carrier systems can be found in, e.g., [8].

The channel vector $\mathbf{h}_k$ is assumed to be a zero-mean stationary (within the observation interval $k = 1, ..., K$) Gaussian process with correlation

$$E[\mathbf{h}_k\mathbf{h}_{k-m}] = \mathbf{R}_h\varphi_m. \tag{2}$$

The correlation matrix $\mathbf{R}_h$ is unknown and generally rank-deficient so that $rank(\mathbf{R}_h) = r \leq M$,the correlation function

φ_m across blocks depends on the time-variations of the channel due to Doppler and it is assumed to be known. Notice that here the correlation φ_m is considered to be the same for all entries of vector $\mathbf{h}_k$ while extension to a more general model requires a more complex notation with some minor modifications to the analysis below. According to (2), the channel $\mathbf{h}_k$ can be restated as $\mathbf{h}_k = \mathbf{R}_h^{1/2}\mathbf{b}_k$,where $\mathbf{b}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is a stationary Gaussian process with $E[\mathbf{b}_k\mathbf{b}_{k-m}] = \mathbf{I}_M\varphi_m$ and $\mathbf{R}_h^{1/2}$ as the $M \times M$ square root matrix of $\mathbf{R}_h$ ($\mathbf{R}_h^{1/2}\mathbf{R}_h^{H/2} = \mathbf{R}_h$).

In summary, the problem is that of estimating the set of channel vectors $\{\mathbf{h}_k\}_{k=1}^K$ from the observations $\{\mathbf{y}_k\}_{k=1}^K$ assuming that the rank-deficient correlation matrix $\mathbf{R}_h$ (2) is unknown, whereas the rank r, the temporal correlation φ_m (2), the correlation of noise $\mathbf{Q}$ are assumed to be known. If necessary, these latter quantities can be estimated using standard techniques (see, e.g., [3]). According to the rank-deficiency of $\mathbf{R}_h$ it is convenient to parameterize $\mathbf{R}_h^{1/2}$ as the product of two full rank matrices $\mathbf{A}$ ($M \times r$) and $\mathbf{C}$ ($M \times r$) such that $\mathbf{R}_h^{1/2} = \mathbf{A}\mathbf{C}^H$. It follows that the channel vector can be conveniently expressed as

$$\mathbf{h}_k = \mathbf{A}\mathbf{C}^H\mathbf{b}_k. \tag{3}$$

Notice that the parametrization (3) is not unique as there are many different ways to define the square root of $\mathbf{R}_h$, and the two factors $\mathbf{A}$ and $\mathbf{C}$. However, in model (3) the full-rank matrices $\mathbf{A}$ and $\mathbf{C}$ are now assumed to be deterministic and block-invariant whereas vector $\mathbf{b}_k$ is random and block-varying. Based on the model (3), the block-varying vector $\mathbf{h}_k$ depends on both deterministic ($\mathbf{A}$ and $\mathbf{C}$) and random ($\{\mathbf{b}_k\}_{k=1}^K$) quantities so that the estimator has to be designed accordingly.

III. Estimation of block-varying stationary channel

From (1) and (3), the log-likelihood function for the estimation of parameters $\mathbf{A}$, $\mathbf{C}$ and $\mathbf{B} = [\mathbf{b}_1, \cdots, \mathbf{b}_K]$ can be proved to be (neglecting uninteresting constants)

$$\mathcal{L} = \sum_{k=1}^{K} \left\| \mathbf{y}_k - \mathbf{X}\mathbf{A}\mathbf{C}^H\mathbf{b}_k \right\|^2_{\mathbf{Q}^{-1}} = \left\| \breve{\mathbf{H}} - \mathbf{A}\mathbf{C}^H\mathbf{B} \right\|^2_{\mathbf{R}^{-1}}. \tag{4}$$

The second equality is crucial for the following reasoning and it can be easily proved by substitution recalling that the pseudoinverse is $\mathbf{X}^\dagger = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$, and by using the definition of the $M \times K$ matrix $\breve{\mathbf{H}} = [\breve{\mathbf{h}}_1 \cdots \breve{\mathbf{h}}_K] = \mathbf{X}^\dagger[\mathbf{y}_1 \cdots \mathbf{y}_K]$ and $\mathbf{R} = \mathbf{X}^\dagger\mathbf{Q}(\mathbf{X}^\dagger)^H$. In other words, according to (4) the K vectors $\{\breve{\mathbf{h}}_k\}_{k=1}^K$ are sufficient statistics for the estimation of $\{\mathbf{h}_k\}_{k=1}^K$. It is interesting to notice that vectors $\{\breve{\mathbf{h}}_k\}_{k=1}^K$ are the least squares estimates of the unknown vectors $\{\mathbf{h}_k\}_{k=1}^K$. Maximum Likelihood estimation could be obtained by minimizing (4) by considering the entries in $\mathbf{B}$ as deterministic variables. Since the set of parameters $\{\mathbf{b}_k\}_{k=1}^K$ are random with a known block-correlation function φ_m, the hybrid estimation method needs to take into account that part of the unknowns are deterministic (entries of $\mathbf{A}$ and $\mathbf{C}$, or equivalently $\mathbf{R}_h$, see Sec. 2) and part are random.

A. Hybrid MoM/Bayesian Estimation

The decoupled structure of the unknowns in (3) suggests that the estimation of the (deterministic and stationary) matrices $(\mathbf{A}, \mathbf{C})$ and the (random) parameters $\mathbf{b}_k$ can be performed separately. Here we propose an hybrid estimator that will be shown in Sec. 5 through numerical simulations to be able to reach the Hybrid CRB for large K.

For estimation of the rank-deficient covariance $\mathbf{R}_h$ we consider the method of moments (MOM) estimator. The correlation matrix of the sufficient statistics $\{\breve{\mathbf{h}}_k\}_{k=1}^K$ reads $\mathbf{R}_{\breve{h}} = E[\breve{\mathbf{h}}_k\breve{\mathbf{h}}_k^H] = \mathbf{R}_h + \mathbf{R}$ so that considering as observations the following $\widetilde{\mathbf{h}} = \mathbf{R}^{-H/2}\breve{\mathbf{h}}_k$ we get $\mathbf{R}_{\widetilde{h}} = E[\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^H] = \mathbf{R}^{-H/2}\mathbf{R}_h\mathbf{R}^{-1/2} + \mathbf{I}_M$. Let the eigenvalue decomposition of $\mathbf{R}_{\widetilde{h}}$ be $\mathbf{R}_{\widetilde{h}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is a $M \times M$ orthonormal matrix and $\mathbf{\Lambda}$ is diagonal with dimension $M \times M$ (it contains the eigenvalues in non-increasing order), we can obtain the correlation matrix $\mathbf{R}_h$ as $\mathbf{R}_h = \mathbf{R}^{H/2}\mathbf{U}_r(\mathbf{\Lambda}_r - \mathbf{I})\mathbf{U}_r^H\mathbf{R}^{1/2}$ (where $\mathbf{U}_r$ is the $M \times r$ matrix collecting the first r columns of $\mathbf{U}$ and $\mathbf{\Lambda}_r$ the $r \times r$ diagonal matrix gathering the first r eigenvalues of $\mathbf{R}_h$). Estimating the second order moment and the corresponding eigenvalue decomposition as

$$\hat{\mathbf{R}}_{\widetilde{h}} = \frac{1}{K}\sum_{k=1}^{K}\widetilde{\mathbf{h}}_k\widetilde{\mathbf{h}}_k^H = \hat{\mathbf{U}}\hat{\mathbf{\Lambda}}\hat{\mathbf{V}}^H, \tag{5}$$

the MOM estimator for rank-r covariance (r is known) reads

$$\hat{\mathbf{R}}_h = \mathbf{R}^{H/2}\hat{\mathbf{U}}(\hat{\mathbf{\Lambda}}_r - \mathbf{I})^+\hat{\mathbf{U}}^H\mathbf{R}^{1/2}, \tag{6}$$

where notation $()^+$ indicates that negative values on the diagonal of $\hat{\mathbf{\Lambda}}_r - \mathbf{I}$ are set to zero in order to preserve the positive-definiteness of $\hat{\mathbf{R}}_h$. The MoM estimator is consistent, i.e., the estimate $\hat{\mathbf{R}}_h$ converges to the real value $\mathbf{R}_h$ for $K \to \infty$.

The estimation of the ensemble $\{\mathbf{h}_k\}_{k=1}^K$ can be obtained by assuming now the knowledge of $\mathbf{R}_h$ (i.e., as if the estimate $\hat{\mathbf{R}}_h$ in (6) was exact as for $K \to \infty$). The sufficient statistics $\{\breve{\mathbf{h}}_k\}_{k=1}^K$ are collected into the $KM \times 1$ vector $\breve{\mathbf{h}} = [\breve{\mathbf{h}}_1^T, ..., \breve{\mathbf{h}}_K^T]^T$ that can be written as

$$\breve{\mathbf{h}} = \mathbf{h} + \mathbf{w} \tag{7}$$

with $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K \otimes \mathbf{R})$. Based on (7), the MMSE estimator of $\mathbf{h}$ is $\hat{\mathbf{h}}_{MMSE} = E[\mathbf{h}\breve{\mathbf{h}}^H](E[\breve{\mathbf{h}}\breve{\mathbf{h}}^H])^{-1}\breve{\mathbf{h}}$. Since $E[\mathbf{h}\breve{\mathbf{h}}^H] = \mathbf{R}_t \otimes \mathbf{R}_h$ and $E[\breve{\mathbf{h}}\breve{\mathbf{h}}^H] = \mathbf{R}_t \otimes \mathbf{R}_h + \mathbf{I}_K \otimes \mathbf{R}$, then substituting to $\mathbf{R}_h$ the corresponding estimates (6), we finally get the channel estimate

$$\hat{\mathbf{h}} = (\mathbf{R}_t \otimes \hat{\mathbf{R}}_h)(\mathbf{R}_t \otimes \hat{\mathbf{R}}_h + \mathbf{I}_K \otimes \mathbf{R})^{-1}\breve{\mathbf{h}} \tag{8}$$

that accounts for temporal correlation of fading across blocks through the temporal-correlation matrix $\mathbf{R}_t$ as $[\mathbf{R}_t]_{i,j} = \varphi_{i-j}$.

The error correlation matrix of the MMSE estimate of $\mathbf{h}$ is $\mathbf{Q}_{\hat{\mathbf{h}}_{MMSE}} = E[(\hat{\mathbf{h}}_{MMSE} - \mathbf{h})(\hat{\mathbf{h}}_{MMSE} - \mathbf{h})^H]$. Assuming that $\mathbf{R}_h$ is known (which is increasingly true for larger K given the consistency of the MOM estimator once the rank

r is known), the error can be proved by substitution of the definitions to be asymptotically (for $K \to \infty$)

$$\mathbf{Q}_{\hat{\mathbf{h}}_{MMSE}} = \mathbf{R}_t \otimes \mathbf{R}_h - (\mathbf{R}_t \otimes \mathbf{R}_h)(\mathbf{R}_t \otimes \mathbf{R}_h + \mathbf{I} \otimes \mathbf{R})^{-1}(\mathbf{R}_t \otimes \mathbf{R}_h). \quad (9)$$

In the next section, we will compare (9) with the Hybrid CRB.

IV. Hybrid Cramér Rao Bound

As discussed above, channel estimation for rank-deficient block-fading channels can be reduced to the unconstrained estimate of two sets of parameters, namely a deterministic $[\mathbf{a}^T, \mathbf{c}^T]^T$ and random part $\mathbf{b}$ with known probability density function (pdf). In this framework, a lower bound on the MSE of any unbiased estimator can be obtained by computation of the hybrid Cramér Rao Bound (HCRB), a modification of the classical CRB for the case where the unknown parameters depend on both deterministic and random variables. The HCRB is evaluated below by adapting the general derivation from [7].

A. Bayesian Fisher Information Matrix

The MSE matrix $\mathbf{Q}_{\hat{\mathbf{h}}} = E[(\hat{\mathbf{h}} - \mathbf{h})(\hat{\mathbf{h}} - \mathbf{h})^H]$ is bounded by the HCRB for the estimation of $\mathbf{h} = [\mathbf{h}_1^T \cdots \mathbf{h}_K^T]$ as

$$\mathbf{Q}_{\hat{\mathbf{h}}} \geq E_{\mathbf{b}}\left[\frac{\partial \mathbf{h}}{\partial[\mathbf{a}^T\mathbf{c}^T\mathbf{b}^T]}\right] \cdot \mathbf{J}^{-1} \cdot E_{\mathbf{b}}\left[\frac{\partial \mathbf{h}}{\partial[\mathbf{a}^T\mathbf{c}^T\mathbf{b}^T]}\right]^H, \quad (10)$$

where $\mathbf{J}$ is the (Bayesian) Fisher Information Matrix. Notice that matrix $\mathbf{J}$ can be written as the sum of a term accounting for the information due to data $\mathbf{J}_D$ and a term accounting for prior knowledge $\mathbf{J}_P$

$$\mathbf{J} = \mathbf{J}_D + \mathbf{J}_P, \quad (11)$$

that in our framework $\mathbf{J}_P$ consists in the statistical properties of the randomly varying parameters $\mathbf{b} \backsim \mathcal{CN}(\mathbf{0}, \mathbf{R}_t \otimes \mathbf{I}_M)$. The blocks of the information matrix related to prior information $\mathbf{J}_P$ becomes [7]

$$\begin{array}{l} [\mathbf{J}_P]_{bb} = \mathbf{R}_t^{-1} \otimes \mathbf{I}_M \\ [\mathbf{J}_P]_{aa} = [\mathbf{J}_P]_{cc} = [\mathbf{J}_D]_{ab} = [\mathbf{J}_D]_{ac} = \mathbf{0} \end{array}. \quad (12)$$

Since the observation model is Gaussian, we have [7]

$$\mathbf{J}_D = E_{\mathbf{b}}\left[\left(\frac{\partial E[\check{\mathbf{h}}|\mathbf{b}]}{\partial[\mathbf{a}^T\mathbf{c}^T\mathbf{b}^T]}\right)^H (\mathbf{I}_K \otimes \mathbf{R}^{-1}) \left(\frac{\partial E[\check{\mathbf{h}}|\mathbf{b}]}{\partial[\mathbf{a}^T\mathbf{c}^T\mathbf{b}^T]}\right)\right], \quad (13)$$

where $E_{\mathbf{b}}[\cdot]$ denotes the ensemble average with respect to the distribution of $\mathbf{b}$. The least squares estimate is unbiased so that $E[\check{\mathbf{h}}|\mathbf{b}] = \mathbf{h}$**,** this equality is useful to derive the blocks of $\mathbf{J}_D$ as

$$\begin{array}{l} [\mathbf{J}_D]_{aa} = K\mathbf{C}^T\mathbf{C}^* \otimes \mathbf{R}^{-1} \\ [\mathbf{J}_D]_{ac} = K\mathbf{C}^T \otimes \mathbf{R}^{-1}\mathbf{A} \\ [\mathbf{J}_D]_{cc} = K\mathbf{I}_M \otimes \mathbf{A}^H\mathbf{R}^{-1}\mathbf{A} \\ [\mathbf{J}_D]_{bb} = \mathbf{I}_K \otimes \mathbf{C}\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A}\mathbf{C}^H \\ [\mathbf{J}_D]_{ab} = [\mathbf{J}_D]_{cb} = \mathbf{0} \end{array}. \quad (14)$$

Non-singularity of $\mathbf{R}_t$ is assumed. This condition implies that the channel variations across blocks are sufficiently fast to allow matrix $\mathbf{R}_t$ to be full rank. Finally, the (Bayesian) Fisher Information Matrix

$$\mathbf{J}_B = \begin{bmatrix} [\mathbf{J}_D]_{aa} & [\mathbf{J}_D]_{ac} & \mathbf{0} \\ [\mathbf{J}_D]_{ca} & [\mathbf{J}_D]_{cc} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & [\mathbf{J}_D]_{bb} + [\mathbf{J}_P]_{bb} \end{bmatrix} \quad (15)$$

proves that the deterministic covariance matrix and the stocastic amplitudes are decoupled terms. The non-uniqueness of the factorization of $\mathbf{R}_h^{1/2}$ into $\mathbf{A}$ and $\mathbf{C}$ is accounted for by the rank-deficiency of the corresponding Fisher Information Matrix [3]:

$$rank\left\{\begin{bmatrix} [\mathbf{J}]_{aa} & [\mathbf{J}]_{ac} \\ [\mathbf{J}]_{ca} & [\mathbf{J}]_{cc} \end{bmatrix}\right\} = r(N + M - r). \quad (16)$$

B. HCRB for channel estimation

Substituting (15) in (10) and noticing that from (15)

$$E_{\mathbf{b}}\left[\frac{\partial \mathbf{h}}{\partial[\mathbf{a}^T\mathbf{c}^T\mathbf{b}^T]}\right] = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I}_K \otimes \mathbf{A}\mathbf{C}^H \end{bmatrix} \quad (17)$$

the HCRB (10) becomes

$$\begin{array}{c} \mathbf{Q}_{\hat{\mathbf{h}}} \geq (\mathbf{I}_K \otimes \mathbf{A}\mathbf{C}^H)[\mathbf{J}]_{bb}^{-1}(\mathbf{I}_K \otimes \mathbf{C}\mathbf{A}^H) = \\ = \mathbf{R}_t \otimes \mathbf{R}_h - (\mathbf{R}_t \otimes \mathbf{R}_h)(\mathbf{R}_t \otimes \mathbf{R}_h + \mathbf{I}_K \otimes \mathbf{R})^{-1}(\mathbf{R}_t \otimes \mathbf{R}_h), \end{array} \quad (18)$$

where derivation the second equality can be proved by matrix lemmas (see Appendix). The corresponding bound on the mean square error $MSE_{\hat{\mathbf{h}}} = \mathrm{tr}(\mathbf{Q}_{\hat{\mathbf{h}}})$ is obtained from the trace of each of the two forms (18).

Notice that, as expected from the discussion in Sec. 2, the CRB (18) depends only on $\mathbf{R}_h$ (and not on the specific choice of $\mathbf{R}_h^{1/2}$ or on the factors $\mathbf{A}$ and $\mathbf{C}$). In addition, the MSE (9) coincides with the HCRB (18), thus showing that the hybrid estimation algorithm discussed in Sec. 3 is asymptotically (for $K \to \infty$) optimum (i.e., efficient).

Notice that, even if only the HCRB at the kth time instant $\mathbf{Q}_{\hat{\mathbf{h}}_K} = E[(\hat{\mathbf{h}}_k - \mathbf{h}_k)(\hat{\mathbf{h}}_k - \mathbf{h}_k)^H]$ is of interest, in the general case the entire $KM \times KM$ matrix $\mathbf{J}_b$ has to be inverted in (18). In fact, from (18) it is

$$\mathbf{Q}_{\hat{\mathbf{h}}_k} \geq (\mathbf{e}_k^T \otimes \mathbf{R}_h^{1/2})\mathbf{J}_b^{-1}(\mathbf{e}_k \otimes \mathbf{R}_h^{H/2}), \quad (19)$$

where the $K \times 1$ vector $\mathbf{e}_k$ is the kth column of the identity matrix $\mathbf{I}_K$. However, if the parameters $\mathbf{b}_k$ are uncorrelated block-to-block (i.e., $\mathbf{R}_t = \mathbf{I}_K$) the Fisher Information Matrix $\mathbf{J}_b$ becomes block diagonal proving that the estimation of vector $\mathbf{b}_k$ for each block can be decoupled from the others (but not the estimate $\hat{\mathbf{R}}_h$) without any performance degradation:

$$\mathbf{Q}_{\hat{\mathbf{h}}_k} \geq \mathbf{R}_h - \mathbf{R}_h(\mathbf{R}_h + \mathbf{R})^{-1}\mathbf{R}_h. \quad (20)$$

V. Numerical example

In this Section we validate the conclusions above by a simple numerical example. We consider $M = 8$ taps channel randomly generated and characterized by rank-3 covariance matrix $\mathbf{R}_h$. Furthermore, noise is assumed to be white, $\mathbf{Q} = \sigma_n^2\mathbf{I}_N$, and the convolution matrix $\mathbf{X}$ for training sequences is orthonormal so that $\mathbf{R}_x = \sigma_x^2\mathbf{I}_M$ (notice that this normalization renders the outcome of this example independent

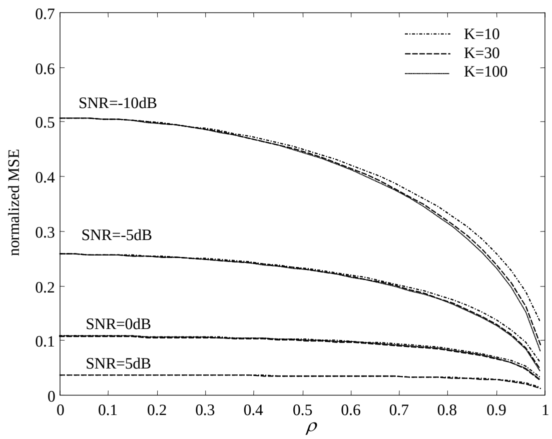

Fig. 1. Hybrid CRB (normalized with respect to the channel norm) versus temporal correlation coefficient ρ for different values of the SNR and K ($M = 8$, $r = 3$).

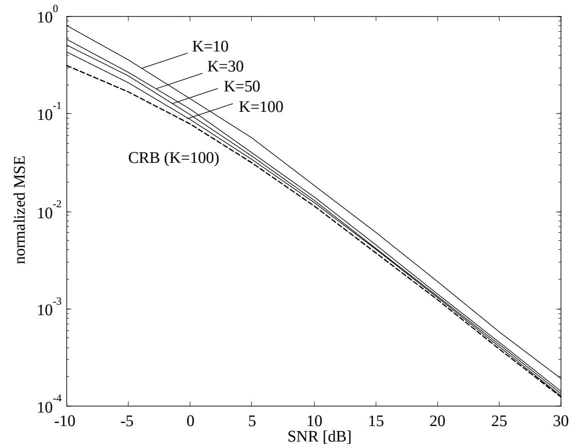

Fig. 2. MSE of the proposed estimator versus SNR for different values of K ($M = 8$, $r = 3, \rho = 0.8$).

on N). It follows that $\mathbf{R} = \sigma_n^2/\sigma_x^2 \mathbf{I}_M$. For simplicity, temporal correlation of the parameters is generated according to an autoregressive model of first order with $\varphi_m = \rho^m$, or equivalently matrix $\mathbf{R}_t$ has a Toeplitz structure with first column $[1\ \rho\ \rho^2 \cdots \rho^{K-1}]^T$, $|\rho| \leq 1$.

The behavior of the lower bound (trace of HCRB in (18)) is shown in fig. 1 (normalized with respect to the norm $E[||\mathbf{h}_k||^2]$) versus the temporal correlation coefficient ρ for different values of the signal to noise ratio $SNR = \sigma_x^2/\sigma_n^2$ and K. Larger values of ρ make temporal filtering of the parameters more effective and accordingly the HCRB decreases. As expected, this is increasingly true for a larger observation interval K, most noticeably for low SNR's.

The MSE of the proposed estimator $MSE_{\hat{\mathbf{h}}} = E[||\hat{\mathbf{h}} - \mathbf{h}||^2]/K$ versus SNR ($\rho = 0.8$) for varying number of blocks (K) is in fig. 2. Notice from fig. 1 that the HCRB for $\rho = 0.8$ is essentially independent on K so that fig. 2 shows only the HCRB (dashed lines) for $K = 100$ (as for others K values it would be similar). According to the consistency of the MoM for the estimate $\hat{\mathbf{R}}_h$, the hybrid estimator reaches the HCRB for large K and SNR.

VI. Conclusion

The considered time-varying channel model for estimation of block-fading channels is quite general as the channel vector is modelled as a stationary Gaussian process with unknown and generally rank-deficient correlation matrix. The model applies to single and multi-antennas with appropriate definitions of the parameters into play. The proposed method approaches the estimate of the mixing deterministic/stochastic terms by decoupling the estimation of stationary rank-deficient correlation with the Method of Moments and the time-varying fading with the MMSE principle. Moreover, a bound on the estimation error has been derived though calculation of the Hybrid CRB. The numerical analysis has proved that the hybrid estimator is asymptotically efficient with the number of blocks.

VII. Appendix: proof of equality (18)

From the matrix inversion lemma:

$$(\mathbf{D}^{-1} + \mathbf{F}\mathbf{E}^{-1}\mathbf{F}^H)^{-1} = \mathbf{D} - \mathbf{D}\mathbf{F}(\mathbf{E} + \mathbf{F}^H\mathbf{D}\mathbf{F})^{-1}\mathbf{F}^H\mathbf{D}^H \tag{21}$$

given invertible matrices $\mathbf{D}$ and $\mathbf{E}$ and matrix $\mathbf{F}$ with appropriate dimensions. The term $[\mathbf{J}]_{bb}^{-1} = (\mathbf{I}_K \otimes \mathbf{C}\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A}\mathbf{C}^H + \mathbf{R}_t^{-1} \otimes \mathbf{I}_M)^{-1}$ in the right hand side of first form of eq.(18) can be cast into a form suitable for the application of (21) by defining $\mathbf{D} = \mathbf{R}_t \otimes \mathbf{I}_M$, $\mathbf{E} = \mathbf{I}_K \otimes \mathbf{R}$ and $\mathbf{F} = \mathbf{I}_K \otimes \mathbf{C}\mathbf{A}^H$. Now, using (21) and usual properties of the Kronecker product, second form of eq.(18) is easily obtained.

References

[1] E. Biglieri, J. Proakis and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.

[2] F.A. Dietrich, W. Utschick, "Pilot-assisted channel estimation based on second-order statistics," *IEEE Trans. Signal Processing*, vol.53, no. 3, pp.1178-1193, March 2005.

[3] P. Stoica and M. Viberg, "Maximum likelihood parameter and rank estimation in reduced-rank multivariate regressions," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 3069-3078, Dec. 1996.

[4] M. Nicoli and U. Spagnolini, "Reduced-rank channel estimation for time-slotted mobile communication systems," *IEEE Trans. Signal Processing*, vol. 53, no. 3, pp. 926-944, March 2005.

[5] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, P .O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931-939, July 1998.

[6] D. Shiu, G. J. Foschini and M. J. Gans, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, March 2000.

[7] H. L. Van Trees, *Optimum array processing*, J. Wiley Ed. 2002.

[8] O. Simeone and U. Spagnolini, "Lower bound on training-based channel estimation error for frequency-selective block-fading Rayleigh MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, no. 11, pp. 3265-3277, Nov. 2004.

[9] P. Stoica and T. L. Marzetta, "Parameter estimation problem with singular information matrices," *IEEE Trans. Signal Processing*, vol. 49, no. 1, pp. 87-90, Jan. 2001.