in Proc. IEEE SPAWC-2005, New York (NY), June 2005, pp. 595–599, Copyright IEEE 2005

2005 IEEE 6th Workshop on Signal Processing Advances in Wireless Communications

IMPROVED MMSE ESTIMATION OF CORRELATED MIMO CHANNELS USING A STRUCTURED CORRELATION ESTIMATOR*

Nicolai Czink[†], *Gerald Matz*[‡], *Dominik Seethaler*[‡], *Franz Hlawatsch*[‡]

[†]Telecommunications Research Center Vienna (ftw.)
Donau-City-Strasse 1, A-1220 Vienna, Austria. Email: czink@ftw.at

[‡]Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
Email: {gerald.matz, dominik.seethaler, franz.hlawatsch}@tuwien.ac.at

ABSTRACT

Channel estimation is an important and challenging task in MIMO communications. The minimum mean-square-error (MMSE) channel estimator is able to exploit spatial correlation of the MIMO channel but requires prior estimation of the channel correlation matrix. In this paper, we investigate pilot-based MMSE channel estimation *including channel correlation estimation*. We propose an MMSE channel estimator using a *structured* correlation estimator and demonstrate its advantages over conventional MMSE estimators. Simulation results show that the proposed channel estimator outperforms conventional channel estimators in the case of strong spatial correlation and at low SNR.

1. INTRODUCTION

1.1. Background and Contribution

Multi-input multi-output (MIMO) communication systems promise strongly improved spectral efficiency and reliability [1]. Accurate estimation of the MIMO channel is a prerequisite for fully realizing this performance potential. The spatial correlation of real MIMO channels [2] can facilitate channel estimation if the correlation characteristics are known with sufficient accuracy.

Two widely used channel estimators are the least-squares (LS) estimator and the minimum mean-square error (MMSE) estimator [3, 4]. The MMSE estimator exploits knowledge of the noise variance and—depending on its formulation—of either the correlation matrix of the received vector or the channel correlation matrix. If this prior knowledge can be estimated with sufficient accuracy, the MMSE estimator outperforms the LS estimator. The receive correlation matrix is usually estimated by the sample correlation of the received vectors. The channel correlation matrix is usually estimated by the sample correlation of "preliminary" channel estimates—typically LS estimates—since channel realizations are not directly observed (e.g., [3, 5]).

In this paper, we study pilot-based MMSE channel estimation *including estimation of the receive correlation matrix or channel correlation matrix*. Two conventional MMSE channel estimators using estimated correlation matrices are reviewed in Section 2. In Section 3, we propose an improved MMSE channel estimator based on *structured correlation estimation* [6] and demonstrate its advantages over conventional channel estimators. Simulation results for channels with synthetic and measured correlation matrices are presented in Section 4. It is shown that for channels with strong spatial correlation and low SNR, the proposed channel estimator generally outperforms the conventional estimators.

*This work was supported by EU Network of Excellence NEWCOM and by FWF grant J-2302.

1.2. System Model

We consider a flat-fading MIMO system with N_t transmit antennas and N_r receive antennas. The channel is assumed to stay constant during a signal block. Channel estimation will be based on the transmission of a block of L pilot symbols per transmit antenna. The transmission of the nth pilot block can be modeled as

$$\mathbf{Y}[n] = \mathbf{S}_0 \mathbf{H}[n] + \mathbf{W}[n], \qquad (1)$$

where $\mathbf{S}_0$ is an $L \times N_t$ pilot symbol matrix, $\mathbf{H}[n]$ is an $N_t \times N_r$ Gaussian channel matrix, $\mathbf{W}[n]$ is an $L \times N_r$ white Gaussian noise matrix, and $\mathbf{Y}[n]$ is the $L \times N_r$ received matrix.

It is convenient to vectorize the matrix equation (1) as

$$\mathbf{y}[n] = \mathbf{S}\mathbf{h}[n] + \mathbf{w}[n]. \qquad (2)$$

Here, $\mathbf{S} = \mathbf{I}_{N_r} \otimes \mathbf{S}_0$, $\mathbf{h}[n] = \mathrm{vec}\{\mathbf{H}[n]\}$, $\mathbf{w}[n] = \mathrm{vec}\{\mathbf{W}[n]\}$, and $\mathbf{y}[n] = \mathrm{vec}\{\mathbf{Y}[n]\}$, where $\otimes$ denotes the Kronecker product and $\mathrm{vec}\{\cdot\}$ denotes columnwise stacking of a matrix into a vector. We assume $L \geq N_t$ (usually $L > N_t$ since this results in an increased effective channel SNR). The pilot symbol matrix $\mathbf{S}$ of size $LN_r \times N_tN_r$ is a tall or square matrix; we assume that it has full rank so that $(\mathbf{S}^H\mathbf{S})^{-1}$ exists. The noise vector $\mathbf{w}[n]$ is Gaussian with correlation matrix $\sigma_w^2\mathbf{I}$. The channel is assumed wide-sense stationary and ergodic. The full correlation matrix of the channel—describing the correlation between any two elements of the channel matrix—is given by

$$\mathbf{R}_{\mathbf{h}} = \mathrm{E}\{\mathbf{h}[n]\mathbf{h}^H[n]\}.$$

2. CONVENTIONAL MMSE CHANNEL ESTIMATORS

In this section, we discuss MMSE estimators based on conventional estimators of the receive or channel correlation matrix.

2.1. Fundamentals of MMSE Channel Estimation

We desire to estimate the channel $\mathbf{h}[n]$ from the received vector $\mathbf{y}[n]$. The *MMSE channel estimator* is given by [4]

$$\hat{\mathbf{h}}[n] = \mathbf{R}_{\mathbf{yh}}^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{y}[n], \qquad (3)$$

with $\mathbf{R}_{\mathbf{yh}} = \mathrm{E}\{\mathbf{y}[n]\mathbf{h}^H[n]\}$ and $\mathbf{R}_{\mathbf{y}} = \mathrm{E}\{\mathbf{y}[n]\mathbf{y}^H[n]\}$. The mean-square error (MSE) of this estimate depends on $\mathbf{S}$; it is smallest when $\mathbf{S}^H\mathbf{S} = c\mathbf{I}$ with $c \in \mathbb{R}$, which is possible for certain modulation schemes and pilot block lengths L [7].

0-7803-8867-4/05/$20.00©2005 IEEE

We can develop two different but equivalent formulations of the MMSE estimator (3). With (2), we have

$$\mathbf{R}_{\mathbf{yh}} = \mathbf{S}\mathbf{R}_{\mathbf{h}}\,, \qquad \mathbf{R}_{\mathbf{y}} = \mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H + \sigma_w^2\mathbf{I}\,. \qquad (4)$$

By inserting these relations into (3), we obtain the conventional formulation of the MMSE estimator [4] as

$$\hat{\mathbf{h}}[n] = \mathbf{R}_{\mathbf{h}}\mathbf{S}^H(\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H + \sigma_w^2\mathbf{I})^{-1}\mathbf{y}[n]\,. \qquad (5)$$

On the other hand, inserting the first relation of (4) into the second relation and using the full rank of $\mathbf{S}^H\mathbf{S}$ yields $\mathbf{R}_{\mathbf{yh}} = (\mathbf{R}_{\mathbf{y}} - \sigma_w^2\mathbf{I})\,\mathbf{S}(\mathbf{S}^H\mathbf{S})^{-1}$. Inserting this expression into (3), we obtain the following alternative formulation of the MMSE estimator,

$$\hat{\mathbf{h}}[n] = \mathbf{S}^{\#}(\mathbf{I} - \sigma_w^2\mathbf{R}_{\mathbf{y}}^{-1})\,\mathbf{y}[n]\,, \qquad (6)$$

with the pseudoinverse $\mathbf{S}^{\#} := (\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H$ (note the relation $\mathbf{S}^{\#}\mathbf{S} = \mathbf{I}$). The formulation (5) in terms of the channel correlation $\mathbf{R}_{\mathbf{h}}$ and the formulation (6) in terms of the receive correlation $\mathbf{R}_{\mathbf{y}}$ are fully equivalent. However, in practice $\mathbf{R}_{\mathbf{h}}$ and $\mathbf{R}_{\mathbf{y}}$ are replaced by *estimates* $\hat{\mathbf{R}}_{\mathbf{h}}$ and $\hat{\mathbf{R}}_{\mathbf{y}}$, respectively, which gives

$$\hat{\mathbf{h}}_1[n] := \hat{\mathbf{R}}_{\mathbf{h}}\mathbf{S}^H(\mathbf{S}\hat{\mathbf{R}}_{\mathbf{h}}\mathbf{S}^H + \sigma_w^2\mathbf{I})^{-1}\mathbf{y}[n]\,, \qquad (7)$$

$$\hat{\mathbf{h}}_2[n] := \mathbf{S}^{\#}(\mathbf{I} - \sigma_w^2\hat{\mathbf{R}}_{\mathbf{y}}^{-1})\,\mathbf{y}[n]\,. \qquad (8)$$

Depending on the way the correlation estimates $\hat{\mathbf{R}}_{\mathbf{h}}$ and $\hat{\mathbf{R}}_{\mathbf{y}}$ are calculated, the channel estimates $\hat{\mathbf{h}}_1[n]$ and $\hat{\mathbf{h}}_2[n]$ will generally be different (this explains our notation using different subscripts).

2.2. Conventional Correlation Estimators

We now discuss conventional techniques for estimating the correlation matrices $\mathbf{R}_{\mathbf{h}}$ and $\mathbf{R}_{\mathbf{y}}$ from N observed received vectors $\mathbf{y}[n]$ $(n = 1, \ldots, N)$. A simple and straightforward estimator of the receive correlation matrix $\mathbf{R}_{\mathbf{y}}$ is the *sample correlation*

$$\hat{\mathbf{R}}_{\mathbf{y}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}[n]\mathbf{y}^H[n]\,. \qquad (9)$$

This estimator is consistent (i.e., $\hat{\mathbf{R}}_{\mathbf{y}} \to \mathbf{R}_{\mathbf{y}}$ for $N \to \infty$) due to our ergodicity assumption. A recursive calculation can be used in which $\hat{\mathbf{R}}_{\mathbf{y}}$ is updated for each received vector $\mathbf{y}[n]$.

Estimation of the channel correlation matrix $\mathbf{R}_{\mathbf{h}}$ is more difficult since the channel vectors $\mathbf{h}[n]$ are not directly observed by the receiver. We may use a sample correlation with the $\mathbf{h}[n]$ replaced by "preliminary" channel estimates $\hat{\mathbf{h}}_{\text{prel}}[n]$ [3, 5]:

$$\hat{\mathbf{R}}_{\mathbf{h}} = \frac{1}{N}\sum_{n=1}^{N}\hat{\mathbf{h}}_{\text{prel}}[n]\hat{\mathbf{h}}_{\text{prel}}^H[n]\,. \qquad (10)$$

This can again be implemented recursively. For calculating $\hat{\mathbf{h}}_{\text{prel}}[n]$, the LS estimator is often used, i.e.,

$$\hat{\mathbf{h}}_{\text{prel}}[n] = \hat{\mathbf{h}}_{\text{LS}}[n] = \mathbf{S}^{\#}\mathbf{y}[n]\,. \qquad (11)$$

(An alternative not discussed here is provided by the MMSE channel estimator using the default correlation $\mathbf{R}_{\mathbf{h}} := \mathbf{I}$.) By inserting (11) in (10), we obtain $\hat{\mathbf{R}}_{\mathbf{h}}$ as a linear function of $\hat{\mathbf{R}}_{\mathbf{y}}$ in (9):

$$\hat{\mathbf{R}}_{\mathbf{h}} = \mathbf{S}^{\#}\hat{\mathbf{R}}_{\mathbf{y}}\mathbf{S}^{\#H}. \qquad (12)$$

2.3. Conventional Channel Estimators

We now use the above correlation estimators in the MMSE channel estimator. Inserting (12) into (7), we obtain an expression of $\hat{\mathbf{h}}_1[n]$ in terms of the receive sample correlation matrix $\hat{\mathbf{R}}_{\mathbf{y}}$ in (9):

$$\hat{\mathbf{h}}_1[n] = \mathbf{S}^{\#}\hat{\mathbf{R}}_{\mathbf{y}}\mathbf{P}\left(\mathbf{P}\hat{\mathbf{R}}_{\mathbf{y}}\mathbf{P} + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{y}[n]\,, \qquad (13)$$

with the orthogonal projector $\mathbf{P} := \mathbf{S}\mathbf{S}^{\#} = \mathbf{S}(\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H$. We note that multiplication of a vector by $\mathbf{P}$ amounts to the orthogonal projection of that vector onto the column span of the pilot matrix $\mathbf{S}$ (i.e., the linear space spanned by the columns of $\mathbf{S}$). The projector $\mathbf{P}$ satisfies the relations $\mathbf{P}^H = \mathbf{P}$, $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{P}\mathbf{S} = \mathbf{S}$, and $\mathbf{S}^{\#}\mathbf{P} = \mathbf{S}^{\#}$. We can rewrite (13) as

$$\hat{\mathbf{h}}_1[n] = \mathbf{S}^{\#}\hat{\mathbf{R}}_{\mathbf{y}_S}\left(\hat{\mathbf{R}}_{\mathbf{y}_S} + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{y}[n]\,,$$

where $\hat{\mathbf{R}}_{\mathbf{y}_S} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}_S[n]\mathbf{y}_S^H[n]$ is the sample correlation of the projected receive vectors $\mathbf{y}_S[n] := (\mathbf{P}\mathbf{y})[n]$. Finally, we recall that a second MMSE channel estimator is given by $\hat{\mathbf{h}}_2[n]$ in (8), with $\hat{\mathbf{R}}_{\mathbf{y}}$ given by the sample correlation in (9).

3. AN IMPROVED MMSE ESTIMATOR

The conventional MMSE channel estimators $\hat{\mathbf{h}}_1[n]$ and $\hat{\mathbf{h}}_2[n]$ discussed above suffer from the fact that the sample correlation estimator $\hat{\mathbf{R}}_{\mathbf{y}}$ in (9) does not account for the specific structure of $\mathbf{R}_{\mathbf{y}}$ as described by (4), and the channel correlation estimator $\hat{\mathbf{R}}_{\mathbf{h}}$ in (10) is an *ad hoc* method involving channel pre-estimates that are rather arbitrary. These drawbacks are avoided by the systematic approach presented in the following. As a by-product, this approach yields an accurate estimate of the noise variance σ_w^2.

3.1. Structured Correlation Estimation

The correlation $\mathbf{R}_{\mathbf{y}}$ in (4) is highly structured: it consists of the low-rank component $\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H$ corresponding to the column span of $\mathbf{S}$ and the scaled identity matrix $\sigma_w^2\mathbf{I}$. Furthermore, $\mathbf{R}_{\mathbf{y}}$ depends *linearly* on the channel correlation $\mathbf{R}_{\mathbf{h}}$ and the noise variance σ_w^2. We thus propose to use a *structured LS estimate* [6] of $\mathbf{R}_{\mathbf{y}}$ that enforces the structure $\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H + \sigma_w^2\mathbf{I}$. First, the "parameters" $\mathbf{R}_{\mathbf{h}}$ and σ_w^2 are estimated in an LS-optimal fashion:

$$(\hat{\mathbf{R}}'_{\mathbf{h}}, \widehat{\sigma_w^2}) := \arg\min_{\mathbf{R}_{\mathbf{h}},\sigma_w^2}\left\|\hat{\mathbf{R}}_{\mathbf{y}} - (\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H + \sigma_w^2\mathbf{I})\right\|^2, \qquad (14)$$

where $\hat{\mathbf{R}}_{\mathbf{y}}$ is the receive sample correlation in (9) and $\|\cdot\|$ denotes the Frobenius norm. The resulting structured receive correlation estimate is then given by

$$\hat{\mathbf{R}}_{\mathbf{y}}^{(s)} = \mathbf{S}\hat{\mathbf{R}}'_{\mathbf{h}}\mathbf{S}^H + \widehat{\sigma_w^2}\mathbf{I}\,. \qquad (15)$$

Following the development in [6], it can be shown that the minimization problem (14) is equivalent to solving the following simultaneous linear equations for $\mathbf{R}_{\mathbf{h}}$ and σ_w^2:

$$\mathbf{S}^H\hat{\mathbf{R}}_{\mathbf{y}}\mathbf{S} = \mathbf{S}^H\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H\mathbf{S} + \sigma_w^2\mathbf{S}^H\mathbf{S} \qquad (16)$$

$$\text{tr}\{\hat{\mathbf{R}}_{\mathbf{y}}\} = \text{tr}\{\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H\} + LN_r\sigma_w^2\,, \qquad (17)$$

where $\text{tr}\{\cdot\}$ denotes trace.

Noise variance estimator. Pre- and post-multiplying (16) by $(\mathbf{S}^H\mathbf{S})^{-1}$ yields

$$\mathbf{R}_{\mathbf{h}} = \mathbf{S}^{\#}\hat{\mathbf{R}}_{\mathbf{y}}\mathbf{S}^{\#H} - \sigma_w^2(\mathbf{S}^H\mathbf{S})^{-1}. \qquad (18)$$

Plugging this expression into (17) leads to

$$LN_r\sigma_w^2 = \text{tr}\{\hat{\mathbf{R}}_\mathbf{y} - \mathbf{P}\hat{\mathbf{R}}_\mathbf{y}\mathbf{P} + \sigma_w^2\mathbf{P}\}$$
$$= \text{tr}\{\mathbf{P}^\perp\hat{\mathbf{R}}_\mathbf{y}\mathbf{P}^\perp\} + N_tN_r\sigma_w^2, \qquad (19)$$

with $\mathbf{P}^\perp := \mathbf{I} - \mathbf{P}$ being the orthogonal projector on the orthogonal complement of the column span of $\mathbf{S}$. Solving (19) for σ_w^2 and using $\text{tr}\{\mathbf{P}^\perp\hat{\mathbf{R}}_\mathbf{y}\mathbf{P}^\perp\} = \text{tr}\{\hat{\mathbf{R}}_{\mathbf{P}^\perp\mathbf{y}}\} = \frac{1}{N}\sum_{n=1}^{N}\text{tr}\{\mathbf{P}^\perp\mathbf{y}[n]$ $(\mathbf{P}^\perp\mathbf{y}[n])^H\} = \frac{1}{N}\sum_{n=1}^{N}\|\mathbf{P}^\perp\mathbf{y}[n]\|^2$ yields the noise variance estimate

$$\widehat{\sigma_w^2} = \frac{1}{(L-N_t)N_r}\,\frac{1}{N}\sum_{n=1}^{N}\|\mathbf{P}^\perp\mathbf{y}[n]\|^2.$$

Note that this presupposes $L > N_t$, which is usually satisfied in practice. With (2), we have $\mathbf{P}^\perp\mathbf{y}[n] = \mathbf{P}^\perp\mathbf{w}[n]$. Thus, $\mathbf{P}^\perp\mathbf{y}[n]$ can be interpretated as a "noise estimate" within the $(L-N_t)N_r$-dimensional "noise-only space" that is the orthogonal complement of the column span of $\mathbf{S}$. We also note that $\mathbf{P}^\perp\mathbf{y}[n] = \mathbf{y}[n] - \mathbf{S}\hat{\mathbf{h}}_{\text{LS}}[n]$ with the LS channel estimator $\hat{\mathbf{h}}_{\text{LS}}[n] = \mathbf{S}^\#\mathbf{y}[n]$.

Channel correlation estimator. Substituting $\widehat{\sigma_w^2}$ for σ_w^2 in (18) gives the channel correlation estimate

$$\hat{\mathbf{R}}'_\mathbf{h} := \mathbf{S}^\#\hat{\mathbf{R}}_\mathbf{y}\mathbf{S}^{\#H} - \widehat{\sigma_w^2}(\mathbf{S}^H\mathbf{S})^{-1} = \mathbf{S}^\#(\hat{\mathbf{R}}_\mathbf{y} - \widehat{\sigma_w^2}\mathbf{I})\,\mathbf{S}^{\#H}. \qquad (20)$$

This can be expressed in terms of the conventional channel correlation estimate $\hat{\mathbf{R}}_\mathbf{h} = \mathbf{S}^\#\hat{\mathbf{R}}_\mathbf{y}\mathbf{S}^{\#H}$ in (12):

$$\hat{\mathbf{R}}'_\mathbf{h} = \hat{\mathbf{R}}_\mathbf{h} - \widehat{\sigma_w^2}(\mathbf{S}^H\mathbf{S})^{-1}. \qquad (21)$$

This simple relation of $\hat{\mathbf{R}}'_\mathbf{h}$ to $\hat{\mathbf{R}}_\mathbf{h}$ leads to an interesting interpretation. The conventional channel correlation estimate $\hat{\mathbf{R}}_\mathbf{h}$ is easily seen to be *biased*:

$$\text{E}\{\hat{\mathbf{R}}_\mathbf{h}\} = \mathbf{R}_\mathbf{h} + \sigma_w^2(\mathbf{S}^H\mathbf{S})^{-1}.$$

Hence, (21) shows that $\hat{\mathbf{R}}'_\mathbf{h}$ attempts to compensate the bias in $\hat{\mathbf{R}}_\mathbf{h}$ by subtracting $\widehat{\sigma_w^2}(\mathbf{S}^H\mathbf{S})^{-1}$. Note that this bias compensation does not increase the estimation variance.

Receive correlation estimator. Inserting (20) into (15), the structured estimate of $\mathbf{R}_\mathbf{y}$ finally results as

$$\hat{\mathbf{R}}_\mathbf{y}^{(s)} = \mathbf{P}\hat{\mathbf{R}}_\mathbf{y}\mathbf{P} + \widehat{\sigma_w^2}\mathbf{P}^\perp = \hat{\mathbf{R}}_{\mathbf{y}_S} + \widehat{\sigma_w^2}\mathbf{P}^\perp. \qquad (22)$$

In view of the subspace decomposition (cf. (4))

$$\mathbf{R}_\mathbf{y} = \mathbf{S}\mathbf{R}_\mathbf{h}\mathbf{S}^H + \sigma_w^2\mathbf{I} = (\mathbf{S}\mathbf{R}_\mathbf{h}\mathbf{S}^H + \sigma_w^2\mathbf{P}) + \sigma_w^2\mathbf{P}^\perp,$$

the form (22) is quite intuitive: the first term, $\mathbf{P}\hat{\mathbf{R}}_\mathbf{y}\mathbf{P} = \hat{\mathbf{R}}_{\mathbf{y}_S}$, is the projection of the sample correlation $\hat{\mathbf{R}}_\mathbf{y}$ onto the column span of $\mathbf{S}$ and accounts for $\mathbf{S}\mathbf{R}_\mathbf{h}\mathbf{S}^H + \sigma_w^2\mathbf{P}$, i.e., the low-rank component $\mathbf{S}\mathbf{R}_\mathbf{h}\mathbf{S}^H$ plus the noise correlation within the column span of $\mathbf{S}$; the second term, $\widehat{\sigma_w^2}\mathbf{P}^\perp$, estimates the noise correlation $\sigma_w^2\mathbf{P}^\perp$ in the orthogonal complement of the column span of $\mathbf{S}$.

It is interesting to compare $\hat{\mathbf{R}}_\mathbf{y}^{(s)}$ with the conventional sample correlation $\hat{\mathbf{R}}_\mathbf{y}$. Using $\mathbf{I} = \mathbf{P} + \mathbf{P}^\perp$, we can decompose $\hat{\mathbf{R}}_\mathbf{y}$ as

$$\hat{\mathbf{R}}_\mathbf{y} = \mathbf{P}\hat{\mathbf{R}}_\mathbf{y}\mathbf{P} + \mathbf{P}\hat{\mathbf{R}}_\mathbf{y}\mathbf{P}^\perp + \mathbf{P}^\perp\hat{\mathbf{R}}_\mathbf{y}\mathbf{P} + \mathbf{P}^\perp\hat{\mathbf{R}}_\mathbf{y}\mathbf{P}^\perp \qquad (23)$$
$$= \hat{\mathbf{R}}_{\mathbf{y}_S} + \mathbf{\Delta}, \qquad (24)$$

where $\mathbf{\Delta}$ subsumes the last three terms of (23). Comparing (24) with (22), we see that the difference lies exclusively in the second component, which is $\widehat{\sigma_w^2}\mathbf{P}^\perp$ for $\hat{\mathbf{R}}_\mathbf{y}^{(s)}$ and $\mathbf{\Delta}$ for $\hat{\mathbf{R}}_\mathbf{y}$. Of course, $\widehat{\sigma_w^2}\mathbf{P}^\perp$ is a much more accurate estimate of the true component $\sigma_w^2\mathbf{P}^\perp$ than is $\mathbf{\Delta}$, and thus the overall accuracy of $\hat{\mathbf{R}}_\mathbf{y}^{(s)}$ is better than that of $\hat{\mathbf{R}}_\mathbf{y}$. This advantage becomes stronger when L is increased relative to N_t, because then the dimension $(L-N_t)N_r$ of the orthogonal complement space is larger.

In our experiments, we always observed $\widehat{\sigma_w^2}$ to be highly accurate (its MSE is typically about 30 dB below the MSE of the preliminary LS channel estimates). Therefore, for our further analysis and for our simulations in Section 4, we will use the true value of σ_w^2 instead of the estimate $\widehat{\sigma_w^2}$ for simplicity.

3.2. The Proposed Channel Estimator

Inserting the improved channel correlation estimator $\hat{\mathbf{R}}'_\mathbf{h}$ into (7) yields the MMSE channel estimator

$$\hat{\mathbf{h}}'[n] := \hat{\mathbf{R}}'_\mathbf{h}\mathbf{S}^H(\mathbf{S}\hat{\mathbf{R}}'_\mathbf{h}\mathbf{S}^H + \sigma_w^2\mathbf{I})^{-1}\mathbf{y}[n].$$

Using (20) gives an expression of $\hat{\mathbf{h}}'[n]$ in terms of $\hat{\mathbf{R}}_\mathbf{y}$:

$$\hat{\mathbf{h}}'[n] = \mathbf{S}^\#(\hat{\mathbf{R}}_\mathbf{y} - \sigma_w^2\mathbf{I})\mathbf{P}\left[\mathbf{P}(\hat{\mathbf{R}}_\mathbf{y} - \sigma_w^2\mathbf{I})\mathbf{P} + \sigma_w^2\mathbf{I}\right]^{-1}\mathbf{y}[n].$$

With (22), we readily obtain

$$\hat{\mathbf{h}}'[n] = \mathbf{S}^\#(\mathbf{I} - \sigma_w^2\hat{\mathbf{R}}_\mathbf{y}^{(s)-1})\mathbf{y}[n],$$

which is seen to be identical to $\hat{\mathbf{h}}_2[n]$ in (8) with $\hat{\mathbf{R}}_\mathbf{y}$ replaced by $\hat{\mathbf{R}}_\mathbf{y}^{(s)}$. Hence, using the channel correlation estimate $\hat{\mathbf{R}}'_\mathbf{h}$ in $\hat{\mathbf{h}}_1[n]$ and the structured receive correlation estimate $\hat{\mathbf{R}}_\mathbf{y}^{(s)}$ in $\hat{\mathbf{h}}_2[n]$ results in the *same* channel estimator $\hat{\mathbf{h}}'[n]$.

In general, the proposed estimator $\hat{\mathbf{h}}'[n]$ will outperform $\hat{\mathbf{h}}_1[n]$ since it uses the bias-compensated channel correlation estimate $\hat{\mathbf{R}}'_\mathbf{h}$ instead of the *ad hoc* estimate $\hat{\mathbf{R}}_\mathbf{h}$ in (10), and it will outperform $\hat{\mathbf{h}}_2[n]$ since it uses the structured receive correlation estimate $\hat{\mathbf{R}}_\mathbf{y}^{(s)}$ instead of the sample correlation $\hat{\mathbf{R}}_\mathbf{y}$. The latter effect again becomes stronger when L is increased relative to N_t.

In Section 4, we will verify through simulation results that these theoretical advantages of the proposed estimator $\hat{\mathbf{h}}'[n]$ result in improved estimation accuracy.

3.3. Efficient Implementation

We will now describe an efficient implementation of the proposed estimator $\hat{\mathbf{h}}'[n]$ in which all calculations are performed in the N_tN_r-dimensional column span of $\mathbf{S}$ (cf. [3]). Let $\mathbf{U}$ be an $LN_r \times N_tN_r$ matrix whose columns are an orthonormal basis of the column span of $\mathbf{S}$. It can then be shown that

$$\hat{\mathbf{h}}'[n] = \mathbf{S}'^{-1}(\mathbf{I} - \widehat{\sigma_w^2}\hat{\mathbf{R}}_{\mathbf{y}'}^{-1})\mathbf{y}'[n],$$

with the $N_tN_r \times 1$ vector $\mathbf{y}'[n] := \mathbf{U}^H\mathbf{y}[n]$ and the $N_tN_r \times N_tN_r$ matrices $\mathbf{S}' := \mathbf{U}^H\mathbf{S}$ and $\hat{\mathbf{R}}_{\mathbf{y}'} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}'[n]\mathbf{y}'^H[n]$. Note that $\mathbf{U}$, $\mathbf{S}'$, and $\mathbf{S}'^{-1}$ can be precomputed; furthermore, the matrices $\mathbf{S}'$ and $\mathbf{S}'^{-1}$ are typically sparse. A block diagram for this implementation is depicted in Fig. 1. Similar subspace implementations exist for the estimators $\hat{\mathbf{h}}_1[n]$ and $\hat{\mathbf{h}}_2[n]$.

For the individual steps of this implementation, we obtain the following complexity estimates per pilot block (i.e., for each n):

- Calculation of $\mathbf{y}'[n]$: $\mathcal{O}(LN_tN_r^2)$;

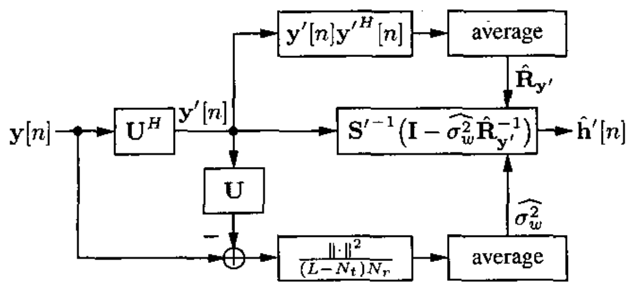

Fig. 1. *Efficient implementation of the proposed estimator* $\hat{\mathbf{h}}'[n]$.

- Recursive calculation of $\hat{\mathbf{R}}_{\mathbf{y}'}$: $\mathcal{O}(N_t^2 N_r^2)$ per matrix update;
- Recursive calculation of $\widehat{\sigma_w^2}$: $\mathcal{O}(LN_tN_r^2)$ per update;
- Computation of the "estimator matrix" $\mathbf{S}'^{-1}(\mathbf{I}-\widehat{\sigma_w^2}\hat{\mathbf{R}}_{\mathbf{y}'}^{-1})$: using Woodbury's identity [6], this can be done recursively with complexity $\mathcal{O}(N_t^2N_r^2)$ per matrix update [8];
- Multiplication of $\mathbf{y}'[n]$ by the estimator matrix: $\mathcal{O}(N_t^2N_r^2)$.

Thus, the dominant complexity is $\mathcal{O}(LN_tN_r^2)$ (note that $L > N_t$). This is similar to the complexity of analogous subspace implementations of the conventional estimators $\hat{\mathbf{h}}_1[n]$ and $\hat{\mathbf{h}}_2[n]$ [3].

4. SIMULATION RESULTS

To assess and compare the performance of the various channel estimators, we performed simulations for a MIMO system with $N_t = N_r = 4$ transmit/receive antennas. Channel realizations were generated according to the full-correlation model

$$\mathbf{h}[n] = \mathbf{R}_{\mathbf{h}}^{1/2}\mathbf{g}[n].$$

Here $\mathbf{g}[n] \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$, i.e., the entries of the vector $\mathbf{g}[n]$ are iid complex Gaussian with zero mean and unit variance, and $\mathbf{R}_{\mathbf{h}}^{1/2}$ is the positive semidefinite square root of $\mathbf{R}_{\mathbf{h}}$. Channel correlation matrices were constructed according to the Kronecker model [9] using exponential correlation coefficients:

$$\mathbf{R}_{\mathbf{h}} = \mathbf{R}_0 \otimes \mathbf{R}_0 \quad \text{with } (\mathbf{R}_0)_{ij} = \rho^{|i-j|}, \quad |\rho| < 1. \quad (25)$$

Alternatively, channel correlation matrices were derived from measurements [10] taken in a quasi-line-of-sight (QLOS) and a non-line-of-sight (NLOS) scenario. For well-defined SNR levels, the channel correlation matrices $\mathbf{R}_{\mathbf{h}}$—both synthetic and measured—were normalized as $\text{tr}\{\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H\} = N_t$. Unless indicated otherwise, we used a pilot block length of $L = 4N_t = 16$, a pilot matrix $\mathbf{S}$ with BPSK symbols and satisfying $\mathbf{S}^H\mathbf{S} = c\,\mathbf{I}$, and a correlation training length of $N = 1000$ (the number of blocks used for estimating $\mathbf{R}_{\mathbf{y}}$ according to (9) and $\mathbf{R}_{\mathbf{h}}$ according to (10)).

4.1. MSE versus SNR

We first simulated a synthetic scenario according to (25) with correlation parameter $\rho = 0.6$. Fig. 2 shows the (normalized) MSE $\text{E}\{\|\hat{\mathbf{h}}-\mathbf{h}\|^2\}/\text{E}\{\|\mathbf{h}\|^2\}$ versus the average SNR at the receive antennas for the conventional channel correlation based MMSE estimator $\hat{\mathbf{h}}_1[n]$ (denoted by "R_h-based" in Fig. 2), the conventional receive correlation based MMSE estimator $\hat{\mathbf{h}}_2[n]$ ("R_y-based"), the proposed MMSE estimator $\hat{\mathbf{h}}'[n]$ ("proposed"), and the LS estimator ("LS"). It is seen that $\hat{\mathbf{h}}'[n]$ outperforms all other estimators, especially at low SNR. All MMSE estimators outperform the LS estimator but approach its performance for high SNR.

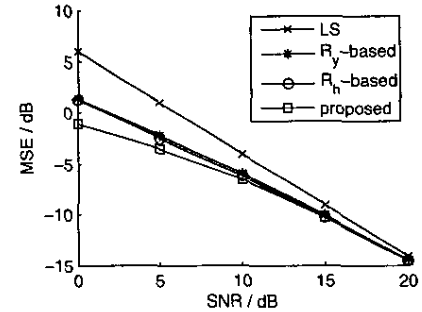

Fig. 2. *MSE of various channel estimators versus the average receive SNR for a synthetic scenario with* $\rho = 0.6$.

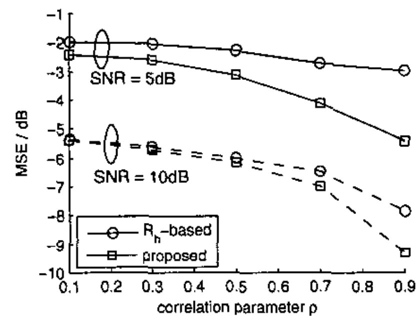

Fig. 3. *MSE of the conventional estimator* $\hat{\mathbf{h}}_1[n]$ *and the proposed estimator* $\hat{\mathbf{h}}'[n]$ *versus the correlation parameter* ρ.

4.2. Influence of the Spatial Channel Correlation

Fig. 3 shows the MSE of the estimators $\hat{\mathbf{h}}_1[n]$ and $\hat{\mathbf{h}}'[n]$ for synthetic scenarios with various correlation parameters ρ, at SNRs of 5 dB and 10 dB. It is seen that the performance advantage of $\hat{\mathbf{h}}'[n]$ over $\hat{\mathbf{h}}_1[n]$ is more pronounced for channels with stronger spatial correlation, and at lower SNRs. Fig. 4 shows the MSE of $\hat{\mathbf{h}}_1[n]$ and $\hat{\mathbf{h}}'[n]$ for QLOS and NLOS scenarios based on measured channel correlations. We see that the performance advantage of $\hat{\mathbf{h}}'[n]$ is larger for the (more correlated) QLOS channel than for the (less correlated) NLOS channel, and it is again largest in the low-SNR regime. For example, in the QLOS environment at an SNR level of 5 dB, $\hat{\mathbf{h}}'[n]$ has about 3 dB less MSE than $\hat{\mathbf{h}}_1[n]$.

4.3. Influence of the Pilot Block Length L

Fig. 5 shows how the performance of the estimators $\hat{\mathbf{h}}_1[n]$, $\hat{\mathbf{h}}_2[n]$, and $\hat{\mathbf{h}}'[n]$ depends on the pilot block length L. We used a synthetic scenario with $\rho = 0.2$, an SNR of 5 dB, and a correlation training length of $N = 500$. The channel correlation matrix $\mathbf{R}_{\mathbf{h}}$ was normalized (i.e., $\text{tr}\{\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H\} = N_t$) for $L = N_t$ but held constant for $L > N_t$ because otherwise the gain in effective SNR obtained for $L > N_t$ would have been canceled by the normalization.

We see from Fig. 5 that all estimators perform better for larger L. The performance of $\hat{\mathbf{h}}_1[n]$ is poor for small L but improves

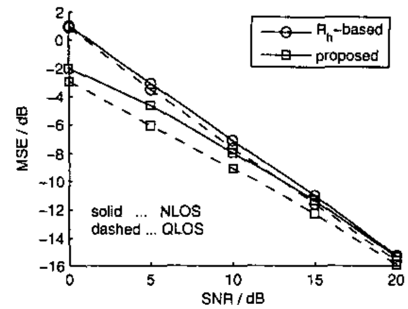

Fig. 4. *MSE of the estimators $\hat{\mathbf{h}}_1[n]$ and $\hat{\mathbf{h}}'[n]$ for real-world NLOS and QLOS scenarios.*

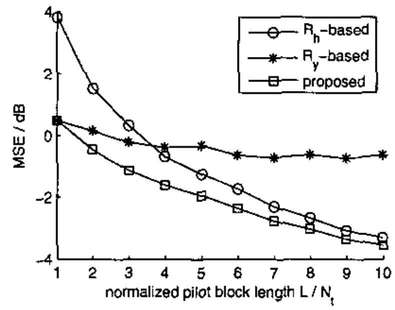

Fig. 5. *MSE of the estimators $\hat{\mathbf{h}}_1[n]$, $\hat{\mathbf{h}}_2[n]$, and $\hat{\mathbf{h}}'[n]$ versus the normalized pilot block length L/N_t.*

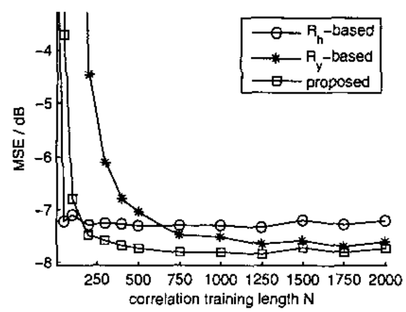

Fig. 6. *MSE of the estimators $\hat{\mathbf{h}}_1[n]$, $\hat{\mathbf{h}}_2[n]$, and $\hat{\mathbf{h}}'[n]$ versus the correlation training length N.*

quickly. The performance of the $\mathbf{R_y}$-based estimator $\hat{\mathbf{h}}_2[n]$ improves only slightly and for large L is approximately constant. This is because the size of the $LN_r \times LN_r$ matrix $\mathbf{R_y}$ increases proportionally with L and thus (for fixed correlation training length N) estimation of $\mathbf{R_y}$ is increasingly difficult. Hence, for long pilot blocks the estimator $\hat{\mathbf{h}}_2[n]$ should not be used. The proposed estimator $\hat{\mathbf{h}}'[n]$ again outperforms all other estimators.

4.4. Influence of the Correlation Training Length N

Finally, we show in Fig. 6 how the performance of the estimators $\hat{\mathbf{h}}_1[n]$, $\hat{\mathbf{h}}_2[n]$, and $\hat{\mathbf{h}}'[n]$ depends on the correlation training length N. We used a synthetic scenario with $\rho = 0.7$ and an SNR of 5 dB. It is seen that the $\mathbf{R_y}$-based estimator $\hat{\mathbf{h}}_2[n]$ performs very poorly for small N. This can be explained by the large size of $\mathbf{R_y}$ (64×64) which calls for a long training phase to allow accurate estimation of $\mathbf{R_y}$. The performance of the $\mathbf{R_h}$-based estimator $\hat{\mathbf{h}}_1[n]$ is very good already for small N but does not further improve for larger N. Finally, the proposed estimator $\hat{\mathbf{h}}'[n]$ outperforms $\hat{\mathbf{h}}_2[n]$ for all N and $\hat{\mathbf{h}}_1[n]$ for N above approximately 170. Below $N \approx 170$, $\hat{\mathbf{h}}_1[n]$ performs better than $\hat{\mathbf{h}}'[n]$ because the matrix $\mathbf{S}\hat{\mathbf{R}}_\mathbf{h}\mathbf{S}^H + \sigma_w^2\mathbf{I}$ inverted in (7) is diagonally dominant due to the bias of $\hat{\mathbf{R}}_\mathbf{h}$ and thus numerical errors in the inversion are reduced. For very large N, the MSE of $\hat{\mathbf{h}}_2[n]$ approaches that of $\hat{\mathbf{h}}'[n]$ whereas the MSE of $\hat{\mathbf{h}}_1[n]$ remains noticeably higher.

5. CONCLUSIONS

We studied MMSE estimation of spatially correlated MIMO channels including pilot symbol based estimation of the channel correlation matrix or receive correlation matrix. A new MMSE channel estimator using a *structured* correlation estimate was proposed and its advantages over conventional MMSE estimators were demonstrated. We also described an efficient subspace implementation of the proposed channel estimator whose complexity is comparable to analogous implementations of conventional estimators.

Simulation results for synthetic and real-world scenarios demonstrated performance gains over conventional MMSE channel estimators that are largest for channels with strong spatial correlation and low SNR. We also studied the dependence of estimator performance on the pilot block length and the correlation training length. Among other results, we found that the conventional receive correlation based estimator performs poorly for large pilot block lengths and/or small correlation training lengths.

6. ACKNOWLEDGMENTS

The authors would like to thank M. Herdin, H. Özcelik, and H. Hofstetter for providing the MIMO channel measurements. Fruitful discussions with W. Weichselberger, M. Jachan, and J. Weinrichter are gratefully acknowledged.

7. REFERENCES

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

[2] E. Bonek, M. Herdin, W. Weichselberger, and H. Özcelik, "MIMO – study propagation first!," in *Proc. IEEE ISSPIT'03*, Darmstadt, Germany, Dec. 2003.

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

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Englewood Cliffs (NJ), 1993.

[5] K. Kopsa, H. Artés, G. Matz, and F. Hlawatsch, "Space-time algorithms for multiuser channel estimation in the downlink of UMTS/TDD," in *Proc. IEEE ICC 2003*, Anchorage, AK, May 2003, pp. 2406–2410.

[6] L. L. Scharf, *Statistical Signal Processing*, Addison Wesley, Reading (MA), 1991.

[7] J. Balakrishnan, M. Rupp, and H. Viswanathan, "Optimal channel training for multiple antenna systems," in *Multiaccess, Mobility and Teletraffic for Wireless Communications: Volume 5*, G. Stüber and B. Jabbari, Eds., pp. 25–36. Kluwer, 2000.

[8] G. Matz, "Recursive MMSE estimation of wireless channels based on training data and structured correlation learning," in *13th IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, July 2005, submitted.

[9] Da-Shan Shiu, G.J. Foschini, M.J. Gans, and J.M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Comm.*, vol. 48, no. 3, pp. 502–513, March 2000.

[10] H. Özcelik, M. Herdin, R. Prestros, and E. Bonek, "How MIMO capacity is linked with single element fading statistics," in *Proc. Int. Conf. Electromagn. Adv. Appl.*, Torino, Italy, Sept. 2003, pp. 775–778.