COMMON AND DEDICATED PILOT-BASED CHANNEL ESTIMATES COMBINING AND KALMAN FILTERING FOR WCDMA TERMINALS

Ahmet Baştuğ[*,†], *Giuseppe Montalbano*[*], *and Dirk Slock*[†]

[*]Philips Semiconductors
06560, Sophia Antipolis, FRANCE
tel: {+33 4 92944130, +33 4 92962381}, fax: +33 4 92961280
{ahmet.bastug, giuseppe.montalbano}@philips.com
http://www.semiconductors.philips.com

[†]Eurecom Institute
06904, Sophia Antipolis, FRANCE
tel: +33 4 93002606 fax: +33 4 93002627
{slock, bastug}@eurecom.fr
http://www.eurecom.fr

ABSTRACT

We consider a family of user dedicated downlink channel estimation methods in WCDMA receivers which are particularly suited for the presence of dedicated channel transmit beamforming and which assume no a priori knowledge of the path delays and the beamforming parameters. They exploit all the transmitted pilot sequences as well as the structured dynamics of the channel. First we build slot-wise least squares (LS) estimates of the channels associated with dedicated and common pilots. Then we optimally improve the dedicated channel estimate quality by jointly Kalman filtering the two LS estimates or alternatively (suboptimally) Kalman filtering them separately and combining via weighted LS. In the suboptimal case, the order of Kalman filtering and weighted LS combining results in differing performance and complexity in different conditions.

1. INTRODUCTION

The UMTS standard [1] user dedicated downlink physical channel (DPCH) consists of dedicated physical control channel (DPCCH), carrying user dedicated pilots, time multiplexed with the dedicated physical data channel (DPDCH) carrying dedicated data. In addition, common pilots are continuously provided over the common pilot channel (CPICH). Most channel estimation techniques proposed for WCDMA receivers are based on either the DPCCH (see e.g. [3, 4] and references therein), or on the CPICH (see e.g [5]). However, on the one hand, the accuracy of channel estimation approaches relying only on the DPCCH is limited by the reduced number of dedicated pilots per slot and by the lack of pilots during the DPDCH period that prevents effective tracking of fast fading channels. On the other hand, classical channel estimation approaches based on the CPICH can better adapt to fast fading conditions, but they are not suited for dedicated channel estimation in the presence of dedicated transmit beamforming. Both approaches remain suboptimal though, due to the fact that they neglect the shared structure by the common and the dedicated propagation channels. There already exist some works for path-wise dedicated channel estimation which make use of both dedicated and common pilots [6], [7], under the assumption of perfect a priori knowledge of the path delays. Moreover they implicitly assume the channel associated with the DPCH to be identical to the one associated with the CPICH. However, as envisaged in the Release 5 of the UMTS standard, this assumption does not hold in the case when beamforming is employed for DPCH transmission. Indeed user dedicated transmit beamforming affects only the DPCH transmission while the CPICH is evenly broadcasted to all users in the cell. Hence, when dedicated beamforming is present one would be tempted to conclude that CPICH can no longer be used for dedicated channel estimation, while the dedicated pilots can still be exploited yet with all the previously described limitations. Actually in order to exploit the common pilots as well, the knowledge of the transmit beamforming parameters, i.e. the beamforming weight vector, antenna array responses corresponding to the excited angles and their related statistics should be known at the receiver. Furthermore, even in the absence of transmit beamforming, the offset between the transmit powers assigned to the DPCCH and CPICH needs to be estimated in order to properly form a combined estimate of the actual dedicated channel.

In general, even in the presence of dedicated beamforming the DPCH and CPICH associated propagation channels are correlated to a certain extent, as it has been shown by field test measurements. A general dedicated channel estimation technique which optimally combines the channel estimates from common and dedicated pilots via a generic CPICH-DPCH channel correlation model was introduced in [2]. In addition to the correlation between dedicated and common channels, there is also the channel temporal correlation governed by the Doppler spread, which can be exploited to improve the channel estimation accuracy. To this end, by fitting the channel dynamics to an autoregressive model of sufficient order, Wiener filtering or Kalman filtering can be applied to refine the previously block-wise obtained estimates.

In this paper we elaborate on one optimal and two suboptimal spatio-temporal Kalman filtering and Kalman smoothing methods that benefit from all the known sources of information, i.e the temporal and cross-correlations of common and dedicated pilots. Their performances are quantified via simulations in terms of the dedicated channel estimate normalized mean square error (NMSE).

2. CHANNEL MODELS

We assume the time-varying continuous time channels associated with dedicated and common pilots, $h_d(t,\tau)$ and $h_c(t,\tau)$ respectively, to obey the wide sense stationary uncorrelated scattering (WSS-US) model [8]

$$\begin{aligned} h_d(t,\tau) &= \sum_{p=0}^{P-1} c_{d,p}(t)\psi(\tau-\tau_p) \\ h_c(t,\tau) &= \sum_{p=0}^{P-1} c_{c,p}(t)\psi(\tau-\tau_p) \end{aligned} \qquad (1)$$

where $\psi(\tau)$ represents the pulse-shape filter, P denotes the number of significant paths, τ_p represents the p-th path delay, $c_{d,p}(t)$ and $c_{c,p}(t)$ are time-varying complex channel coefficients associated with the p-th path of the dedicated and common channel respectively. In many practical circumstances, the two coefficients $c_{d,p}(t)$ and $c_{c,p}(t)$ result to be fairly highly correlated even in the presence of dedicated downlink beamforming. Notice that in (1) the coefficients $c_{d,p}(t)$ for $p = 0, \ldots P-1$ account also for the complete cascade of the beamforming weight vector, the antenna array response on the excited angles, as well as for the actual propagation channel between the transmitter and the receiver. The receiver is assumed to sample M times per chip period the low-pass filtered received baseband signal. Stacking the M samples per chip period in vectors, the discrete time finite impulse response (FIR) representation of both common and dedicated channels at chip rate takes the form $\boldsymbol{h}_l = [h_{1,l} \ldots h_{M,l}]^T$, which represents the vector of the samples of the overall channel, including the pulse shape, the propagation channel, the anti-aliasing receiver filter and, when applicable, the beamforming weighting. The superscript $(\cdot)^T$ denotes the transpose operator. Assuming the overall channel to have a delay spread of N chip periods, the dedicated and common channel impulse responses take the form $\boldsymbol{h}(n) = \boldsymbol{\Psi c}(n)$ where $\boldsymbol{h} = [\boldsymbol{h}_1^T, \ldots, \boldsymbol{h}_N^T]^T \in \mathscr{C}^{MN\times 1}$, $\boldsymbol{c}(n) = [c_1(n) \ldots c_P(n)]^T \in \mathscr{C}^{P\times 1}$ are the complex path amplitudes and the temporal index n relates to the time instant at which the time-varying channel is observed. The assumption of fixed delays τ_p's over the observation window, yields to a constant pulse-shape convolution matrix $\boldsymbol{\Psi} \in \mathscr{R}^{MN\times P}$ given by

$$\boldsymbol{\Psi} = \boldsymbol{\Psi}(\tau_1, \cdots, \tau_P) = [\boldsymbol{\psi}(\tau_1), \ldots, \boldsymbol{\psi}(\tau_P)]$$

where $\boldsymbol{\psi}(\tau_p)$ represents the sampled version of the pulse shape filter impulse response delayed by τ_p. The complex path amplitudes variations are modeled as an autoregressive (AR) processes of order sufficiently high to characterize the Doppler spectrum. Matching only the channel bandwidth with the Doppler spread leads to a first-order AR(1) model of the form

$$\boldsymbol{c}(n) = \rho\, \boldsymbol{c}(n-1) + \sqrt{1-|\rho|^2}\Delta\boldsymbol{c}(n) = \frac{\sqrt{1-|\rho|^2}}{1-\rho q^{-1}}\Delta\boldsymbol{c}(n)$$

so that, $\boldsymbol{\Psi}$ being constant over the observation time interval, we obtain

$$\boldsymbol{h}(n) = \rho\boldsymbol{h}(n-1) + \sqrt{1-|\rho|^2}\Delta\boldsymbol{h}(n) = \frac{\sqrt{1-|\rho|^2}}{1-\rho q^{-1}}\Delta\boldsymbol{h}(n) \qquad (2)$$

where q^{-1} denotes the delay operator such that $q^{-1}y(n) = y(n-1)$ and ρ represents the AR process temporal coherence correlation coefficient. Since the Doppler spread is assumed to be the same for both channels (1), the model (2) applies to both $\boldsymbol{h}_d(n)$ and $\boldsymbol{h}_c(n)$. The variance of k-th component $h_{c,k}(n)$ of $\boldsymbol{h}_c(n)$ is $\sigma^2_{h_{c,k}} = \sigma^2_{\Delta h_{c,k}} = \boldsymbol{\psi}_k \boldsymbol{D}_c \boldsymbol{\psi}_k^H$ where $\boldsymbol{\psi}_k$ denotes the k-th line of $\boldsymbol{\Psi}$ and $\boldsymbol{D}_c = \mathrm{diag}(\sigma^2_{\Delta c_1}, \ldots, \sigma^2_{\Delta c_{c,P}})$. Notice that $\sigma^2_{c_{c,p}} = \sigma^2_{\Delta c_{c,p}}$. Similarly the variance of k-th component $h_{d,k}(n)$ of $\boldsymbol{h}_d(n)$, is $\sigma^2_{h_{d,k}} = \sigma^2_{\Delta h_{d,k}} = \boldsymbol{\psi}_k \boldsymbol{D}_d \boldsymbol{\psi}_k^H$ where $\boldsymbol{D}_d = \mathrm{diag}(\sigma^2_{\Delta c_{d,1}}, \ldots, \sigma^2_{\Delta c_{d,P}})$.

3. LS ESTIMATIONS OF COMMON AND DEDICATED CHANNELS

All the three proposed approaches start with block-wise dedicated and common channel least squares (LS) estimates $\hat{\boldsymbol{h}}_c(n)$ and $\hat{\boldsymbol{h}}_d(n)$ which are computed based on the a priori knowledge of the common and dedicated pilot chips. For the sake of simplicity, without loss of generality, in this paper, we assume that block-wise corresponds to slot-wise estimates. We assume that dedicated pilot chips are sent in every slot. Let $\boldsymbol{S}_d(n) = \mathsf{S}_d(n) \otimes \mathbf{I}_M$, where $\otimes$ denotes the Kronecker product, represent the block Hankel matrix comprising the dedicated pilot chip sequence intended for the user of interest in slot n. Similarly we refer to $\boldsymbol{S}_c(n) = \mathsf{S}_c(n) \otimes \mathbf{I}_M$ as the block Hankel matrix containing the common pilot chip sequence in slot n. Let $\boldsymbol{Y}(n)$ be the received signal samples vector corresponding to slot n. The LS unstructured FIR common and dedicated channel estimates FIR are given by

$$\begin{aligned}\hat{\boldsymbol{h}}_d(n) &= \arg\min_{\boldsymbol{h}_d} \|\boldsymbol{Y}(n) - \boldsymbol{S}_d(n)\boldsymbol{h}_d(n)\|^2 \\ \hat{\boldsymbol{h}}_c(n) &= \arg\min_{\boldsymbol{h}_c} \|\boldsymbol{Y}(n) - \boldsymbol{S}_c(n)\boldsymbol{h}_c(n)\|^2\end{aligned} \qquad (3)$$

The exact LS solutions of problems (3) are readily given by

$$\begin{aligned}\hat{\boldsymbol{h}}_d(n) &= (\boldsymbol{S}_d^H(n)\boldsymbol{S}_d(n))^{-1}\boldsymbol{S}_d^H(n)\boldsymbol{Y}(n) \\ \hat{\boldsymbol{h}}_c(n) &= (\boldsymbol{S}_c^H(n)\boldsymbol{S}_c(n))^{-1}\boldsymbol{S}_c^H(n)\boldsymbol{Y}(n)\end{aligned} \qquad (4)$$

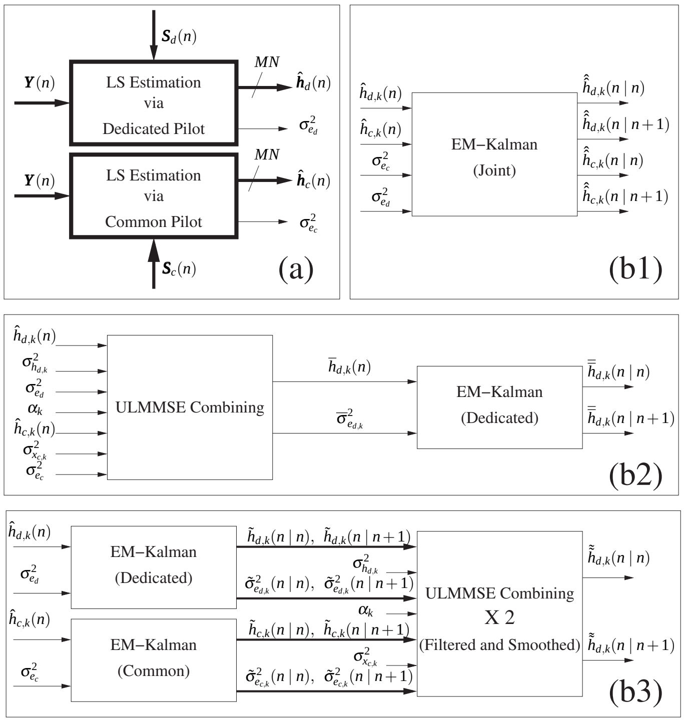

Figure 1: a: LS estimation (joint for all taps), {b1,b2,b3}: {optimal scheme, suboptimal scheme 1, suboptimal scheme 2} for each tap

where $(\cdot)^H$ denotes Hermitian transpose. Note that the equations (4) reduce to

$$\hat{\boldsymbol{h}}_d(n) \approx \beta_d^{-1}\boldsymbol{S}_d^H(n)\boldsymbol{Y}(n); \quad \hat{\boldsymbol{h}}_c(n) \approx \beta_c^{-1}\boldsymbol{S}_c^H(n)\boldsymbol{Y}(n)$$

if the pilot chips can be modeled as i.i.d. random variables, where β_d and β_c represent the dedicated and common pilot chip sequences total energies respectively. We can estimate $\sigma^2_{e_{d,k}}$ and $\sigma^2_{e_{c,k}}$ from $\hat{h}_{d,k}$ and $\hat{h}_{c,k}$ at delays k where we expect the channel not to carry any energy. That can be achieved by, e.g., overestimating the channel delay spread, and using the tails of the channel estimates to obtain unbiased estimates $\sigma^2_{e_{d,k}}$ and $\sigma^2_{e_{c,k}}$.

4. OPTIMAL RECURSIVE APPROACH: JOINT KALMAN FILTERING AND SMOOTHING

Channel Dynamics (State Vector)

$\boldsymbol{h}(n) = \begin{bmatrix} h_d(n) \\ h_c(n) \end{bmatrix}$: Present State Vector

$\boldsymbol{h}(n+1) = \rho\boldsymbol{h}(n) + \boldsymbol{Bu}(n)$: State Transition Process

$\boldsymbol{B} = \sqrt{1-|\rho^2|}\begin{bmatrix} 1 & 0 \\ \alpha & \sqrt{1-\frac{\sigma^2_{h_d}}{\sigma^2_{h_c}}|\alpha^2|} \end{bmatrix}$: Input Gain

$\boldsymbol{u}(n) = \begin{bmatrix} \Delta h_d(n) \\ \Delta h_c(n) \end{bmatrix}$: Input Vector, α: CPICH-DPCH corr. coef.

$\boldsymbol{R_{uu}} = \begin{bmatrix} \sigma^2_{\Delta h_d} & 0 \\ 0 & \sigma^2_{\Delta h_c} \end{bmatrix}$: Input Covariance

$\boldsymbol{Bu}(n)$: Process Noise

$\boldsymbol{Q} = \boldsymbol{BR_{uu}B}^H$: Process Noise Covariance

First Step LS Estimation (State Measurement)

$\hat{\boldsymbol{h}}(n) = \begin{bmatrix} \hat{h}_d(n) \\ \hat{h}_c(n) \end{bmatrix} = \boldsymbol{h}(n) + \boldsymbol{w}(n)$: Measurement (*LS estimates*)

$\boldsymbol{w}(n) = \begin{bmatrix} e_d(n) \\ e_c(n) \end{bmatrix}$: Measurement Noise

$\boldsymbol{R_{ww}} = \begin{bmatrix} \sigma^2_{e_d} & 0 \\ 0 & \sigma^2_{e_c} \end{bmatrix}$: Measurement Noise Covariance

Algorithm Initialization

$\gamma(0) = 0$: Moving Averaging Weight
$\lambda = 0.95$: Forgetting Factor
$\hat{\rho}(0) = 0.999$: Temporal Correlation Coefficient Estimate
$\hat{\hat{\boldsymbol{h}}}(0 \mid 0) = \hat{\boldsymbol{h}}(0 \mid 0)$: Initial State Estimate
$\boldsymbol{S}(0 \mid 0) = \boldsymbol{R_{ww}}$: Initial State Error Covariance
$\boldsymbol{S}(1 \mid 0) = \mid \hat{\rho}(0) \mid^2 \boldsymbol{R_{ww}}$: Initial Prediction Error Covariance
$\hat{\boldsymbol{Q}}(0) = \boldsymbol{0}_{2\times 2}$: Process Noise Covariance Estimate
$\boldsymbol{M}_1 = \boldsymbol{M}_2 = \boldsymbol{M}_{12} = \boldsymbol{0}_{2\times 2}$: Supporting Adaptation Parameters

Kalman Filtering and Smoothing (E-Step)

$\hat{\hat{\boldsymbol{h}}}(n+1 \mid n) = \hat{\rho}\hat{\hat{\boldsymbol{h}}}(n \mid n)$: Time Update
$\boldsymbol{G}(n+1) = \boldsymbol{S}(n+1 \mid n)\,[\boldsymbol{S}(n+1 \mid n) + \boldsymbol{R_{ww}}]^{-1}$: Filter Gain
$\hat{\hat{\boldsymbol{h}}}(n+1 \mid n+1) = [\boldsymbol{I}_{2\times 2} - \boldsymbol{G}(n+1)]\,\hat{\hat{\boldsymbol{h}}}(n+1 \mid n) + \boldsymbol{G}(n+1)\hat{\boldsymbol{h}}(n)$: Measurement Update
$\boldsymbol{S}(n+1 \mid n+1) = [\boldsymbol{I}_{2\times 2} - \boldsymbol{G}(n+1)]\,\boldsymbol{S}(n+1 \mid n)$: Filtered State Covariance
$\boldsymbol{A}(n) = \hat{\rho}^H \boldsymbol{S}(n \mid n)\boldsymbol{S}(n+1 \mid n)^{-1}$: Smoothing Gain
$\hat{\hat{\boldsymbol{h}}}(n \mid n+1) = \hat{\hat{\boldsymbol{h}}}(n \mid n) + \boldsymbol{A}(n)\left(\hat{\hat{\boldsymbol{h}}}(n+1 \mid n+1)\right) - \hat{\hat{\boldsymbol{h}}}(n+1 \mid n)$: Smoothing Update
$\boldsymbol{S}(n \mid n+1) = \boldsymbol{S}(n \mid n) + \boldsymbol{A}(n)\,(\boldsymbol{S}(n+1 \mid n+1) - \boldsymbol{S}(n+1 \mid n))\,\boldsymbol{A}(n)^H$: Smoothed State Error Covariance

Adaptive Estimation of Model Parameters (M-Step)

$\boldsymbol{M}_1 = \lambda \boldsymbol{M}_1 + \hat{\hat{\boldsymbol{h}}}(n+1 \mid n+1)\hat{\hat{\boldsymbol{h}}}(n+1 \mid n+1)^H + \boldsymbol{S}(n+1 \mid n+1)$
$\boldsymbol{M}_2 = \lambda \boldsymbol{M}_2 + \hat{\hat{\boldsymbol{h}}}(n \mid n+1)\hat{\hat{\boldsymbol{h}}}(n \mid n+1)^H + \boldsymbol{S}(n \mid n+1)$
$\boldsymbol{M}_{12} = \lambda \boldsymbol{M}_{12} + \hat{\hat{\boldsymbol{h}}}(n+1 \mid n+1)\hat{\hat{\boldsymbol{h}}}(n \mid n+1)^H + \boldsymbol{S}(n+1 \mid n+1)^H \boldsymbol{A}(n)^H$
$\gamma(n+1) = \lambda\gamma(n) + 1$
$\hat{\rho}(n+1) = Trace\{\boldsymbol{M}_{12}\boldsymbol{M}_2^{-1}\}/2$
$\hat{\boldsymbol{Q}}(n+1) = \frac{1}{\gamma(n+1)}\left(\boldsymbol{M}_1 - \boldsymbol{M}_{12}\boldsymbol{M}_2^{-1}\boldsymbol{M}_{12}^H\right)$
$\boldsymbol{S}(n+2 \mid n+1) = \hat{\rho}(n+1)\boldsymbol{S}(n+1 \mid n+1)\hat{\rho}(n+1)^H + \hat{\boldsymbol{Q}}(n+1)$: Prediction Error Covariance

Steady State Performance

$\boldsymbol{S}(\infty \mid \infty) = \boldsymbol{R}_{w,w}\left[\mid \hat{\rho}(\infty) \mid^2 \boldsymbol{S}(\infty \mid \infty) + \hat{\boldsymbol{Q}}(\boldsymbol{\infty}) + \boldsymbol{R}_{w,w}\right]^{-1} \times \left[\mid \hat{\rho}(\infty) \mid^2 \boldsymbol{S}(\infty \mid \infty) + \hat{\boldsymbol{Q}}(\boldsymbol{\infty})\right]$: Steady State Error Variance
$\boldsymbol{S}(\infty+1 \mid \infty) = \mid \hat{\rho}(\infty) \mid^2 \boldsymbol{S}(\infty \mid \infty) + \hat{\boldsymbol{Q}}(\boldsymbol{\infty})$: Steady State Prediction Error Variance
$\boldsymbol{S}(\infty \mid \infty+1) = \boldsymbol{S}(\infty \mid \infty) + \mid \hat{\rho}(\infty) \mid^2 \boldsymbol{S}(\infty \mid \infty)\boldsymbol{S}(\infty+1 \mid \infty)^{-1}\left[\boldsymbol{S}(\infty \mid \infty) - \boldsymbol{S}(\infty+1 \mid \infty)\right]\boldsymbol{S}(\infty+1 \mid \infty)^{-H}\boldsymbol{S}(\infty \mid \infty)^H$: Steady State Smoothed Error Variance

Above is given the optimal EM-Kalman filtering and smoothing algorithm corresponding to the scheme shown in Figure 1.b1 which we apply *independently* for each channel tap, hence the tap indices dropped for simplicity. It has been used in several other different contexts in order to estimate the states and the unknown model parameters of dynamic systems [9, 10, 11, 12, 13]. On its core lies the Expectation Maximization (EM) algorithm which is the most referred method when the problem in hand is suffering from incomplete data and the solution requires both completing this data and estimating some parameters [14]. The algorithm iterates between the E-phase which is the expected log-likelihood computation of the missing (imputed) data by using both the observed data and the present parameter estimates and the M-phase which computes the maximum likelihood (ML) value of the parameters by conditioning on the imputed data as if it were the correct data. For this section $\boldsymbol{h}(n) = [h_d(n)\ h_c(n)]^T$ channel parameters are the missing data and $\{\hat{\rho}, \hat{\boldsymbol{Q}}\}$ are the only needed parameter estimates. Fitting the EM mechanism to Kalman filtering context requires also smoothing in the E-phase. In all the mentioned papers fixed-interval smoothing mechanism is used which is very complex and large buffer sizes are required, except for [13] where single delay fixed-lag smoothing is considered. In this paper we follow the latter strategy due to its suitability for implementation. We slightly modify the M-phase by taking the average of the two $\hat{\rho}$ estimates (diagonal components of $\boldsymbol{M}_{12}\boldsymbol{M}_2^{-1}$) via the *Trace* operation, considering the temporal correlation coefficients of dedicated and common taps as equal.

5. SUBOPTIMAL SCHEME 1: EM-KALMAN PROCEDURE AFTER ULMMSE COMBINING

This scheme corresponds to Figure 1.b2, it runs independently for each channel tap and it has two phases as explained in the sequel.

5.1 Unbiased LMMSE Combining of LS Estimates

Let $\hat{\boldsymbol{h}}_k(n) = [\hat{h}_{d,k}(n)\ \hat{h}_{c,k}(n)]^T$ denote the vector of the LS estimates of the k-th elements of the dedicated and common pilot channel FIR responses at slot n, i.e.,

$$\hat{\boldsymbol{h}}_k(n) = \begin{bmatrix} \hat{h}_{d,k}(n) \\ \hat{h}_{c,k}(n) \end{bmatrix} = \begin{bmatrix} h_{d,k}(n) \\ h_{c,k}(n) \end{bmatrix} + \begin{bmatrix} e_{d,k}(n) \\ e_{c,k}(n) \end{bmatrix}. \quad (5)$$

In order for our derivation to be fully general, we introduce the following dedicated and common channel correlation model

$$h_{c,k}(n) = \alpha_k h_{d,k}(n) + x_{c,k}(n) \quad (6)$$

where $\alpha_k h_{d,k}(n)$ represents the short-term ULMMSE estimate of $h_{c,k}(n)$ on the basis of $h_{d,k}(n)$, and $x_{c,k}(n)$ represents the associated estimation error. Then, a refined estimate can be obtained as $\overline{h}_{d,k}(n) = \boldsymbol{f}_k\hat{\boldsymbol{h}}_k(n)$ by optimal combining of common and dedicated LS channel estimates. In order not to introduce bias for the processing in the next estimation step, we shall determine $\boldsymbol{f}$ as the ULMMSE filter, i.e. by solving for all k's the optimization problem

$$\min_{\boldsymbol{f}_k} \mathrm{E}|h_{d,k}(n) - \boldsymbol{f}_k\hat{\boldsymbol{h}}_k(n)|^2 \quad \text{s.t. } \boldsymbol{f}_k[1\ \alpha_k]^T = 1$$

The optimal ULMMSE filter $\boldsymbol{f}_k$ is obtained as

$$\begin{aligned} \boldsymbol{f}_{k,\mathrm{ULMMSE}} &= ([1\ \alpha_k^*]\boldsymbol{R}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k}^{-1}[1\ \alpha_k]^T)^{-1}[1\ \alpha_k^*]\boldsymbol{R}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k}^{-1} \\ &= ([1\ \alpha_k^*]\boldsymbol{R}^{-1}[1\ \alpha_k]^T)^{-1}[1\ \alpha_k^*]\boldsymbol{R}^{-1} \end{aligned}$$

where $\boldsymbol{R}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k} = \mathrm{E}\hat{\boldsymbol{h}}_k(n)\hat{\boldsymbol{h}}_k^H(n)$, $\boldsymbol{R} = \mathrm{diag}\,(\sigma^2_{e_{d,k}},\ (\sigma^2_{e_{c,k}} + \sigma^2_{x_{c,k}}))$, with $\sigma^2_{x_{c,k}} = \mathrm{E}|\hat{x}_{c,k}(n)|^2$. Notice that the covariance matrix $\boldsymbol{R}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k}$ is equal to

$$\boldsymbol{R}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \sigma^2_{h_{d,k}}\begin{bmatrix} 1 \\ \alpha_k \end{bmatrix}\begin{bmatrix} 1 \\ \alpha_k \end{bmatrix}^H + \begin{bmatrix} \sigma^2_{e_{d,k}} & 0 \\ 0 & \sigma^2_{e_{c,k}} + \sigma^2_{x_{c,k}} \end{bmatrix}$$

Having an estimate of the matrix $\boldsymbol{R}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k}$, e.g. by temporal averaging, we can apply the *covariance matching* criterion so that $\sigma^2_{h_{d,k}} = r_{11} - \sigma^2_{e_{d,k}}$, $\alpha_k = r_{21}/(r_{11} - \sigma^2_{e_{d,k}})$, (i.e. α_k has the same phase as r_{21}), where the following bound $|\alpha_k| \leq \sigma_{h_{c,k}}/\sigma_{h_{d,k}} = \sqrt{(r_{22} - \sigma^2_{e_{c,k}})/(r_{11} - \sigma^2_{e_{d,k}})}$ can be used in actual estimation. Furthermore, since $\sigma^2_{x_{c,k}} = r_{22} - \sigma^2_{e_{c,k}} - |r_{21}|^2/(r_{11} - \sigma^2_{e_{d,k}})$.

Finally, the variance of the estimation error $\hat{\hat{e}}_{d,k}(n)$ after ULMMSE combining, is obtained as

$$\overline{\sigma}^2_{e_{d,k}} = \frac{\sigma^2_{e_{d,k}}(\sigma^2_{e_{c,k}} + \sigma^2_{x_{c,k}})}{\sigma^2_{e_{d,k}}|\alpha_k|^2 + \sigma^2_{e_{c,k}} + \sigma^2_{x_{c,k}}} \quad (7)$$

The dedicated channel estimate after ULMMSE combining, $\overline{h}_{d,k}(n) = h_{d,k}(n) + \overline{\sigma}^2_{e_{d,k}}$, is such that the post-combining estimation error $\overline{\sigma}^2_{e_{d,k}}$ is mutually uncorrelated with $h_{d,k}(n)$, $\overline{\sigma}^2_{e_{d,k}}$ and $\overline{\sigma}^2_{e_{d,j}}$ are mutually uncorrelated for any $k \neq j$, and the variance $\overline{\sigma}^2_{e_{d,k}}$ is independent of k while it depends on the Doppler spread, on the channel power and on the SINR.

5.2 Kalman Filtering of ULMMSE Combined Estimates

Once the ULMMSE combined dedicated channel estimates are obtained, we apply the causal Kalman filtering and smoothing to obtain the final estimates $\overline{\overline{h}}_{d,k}(n \mid n)$ and $\overline{\overline{h}}_{d,k}(n \mid n+1)$. This algorithm is similar to the one in Section 4 with the single difference that the state vector has now only one element, which is the *complexity advantage* w.r.t. the optimal scheme that has a state vector of two elements. This scheme is identical to the optimal one when the normalized correlation factor $\mid \zeta_k \mid = |\alpha_k| \sigma_{h_{d,k}} / \sigma_{h_{c,k}} \leq 1$ is unity, i.e. $\zeta_k = 1, \forall k$.

6. SUBOPTIMAL SCHEME 2: ULMMSE COMBINING AFTER TWO SEPARATE EM-KALMAN PROCEDURES

In this case we change the order of EM-Kalman procedure and ULMMSE combining as shown in Figure 1.b3. The EM-Kalman filtering outputs dedicated and common channel parameter estimates $\tilde{h}_{d,k}(n \mid n)$ and $\tilde{h}_{c,k}(n \mid n)$ and their associated error variances $\tilde{\sigma}^2_{e_{d,k}}(n \mid n)$ and $\tilde{\sigma}^2_{e_{c,k}}(n \mid n)$ are fed to the following ULMMSE block to obtain the final estimate $\tilde{\tilde{h}}_{d,k}(n \mid n)$. Similar procedure is applied to obtain the refined smoothed estimate $\tilde{\tilde{h}}_{d,k}(n \mid n+1)$ from the EM-Kalman smoothing output parameters $\tilde{h}_{d,k}(n \mid n+1)$, $\tilde{h}_{c,k}(n \mid n+1)$, $\tilde{\sigma}^2_{e_{d,k}}(n \mid n+1)$ and $\tilde{\sigma}^2_{e_{c,k}}(n \mid n+1)$. Other necessary parameters are estimated similar to what is done in Section 5.1. This two stage procedure is suboptimal (due to the coloring of noises at EM-Kalman outputs) unless $\rho = 1$ and it is not as attractive for implementation as the first suboptimal scheme since the Kalman state vector has two elements as in the optimal case.

7. SIMULATIONS AND CONCLUSIONS

The performances of the presented channel estimation methods in the presence of dedicated transmit beamforming are presented in Figure 2 to Figure 7 in terms of the channel estimate NMSE. We assume the DPCCH to occupy 20% of the UMTS slot, and the DPCH spreading factor to be equal to 128. We define the normalized correlation factor $\mid \zeta_k \mid = |\alpha_k| \sigma_{h_{d,k}} / \sigma_{h_{c,k}} \leq 1$. Being interested in the impact of dedicated and common channel correlation we set, for the sake of simplicity, $|\alpha_k| = |\alpha_0|$ constant $\forall k$. We initially assume the DPCCH and the CPICH to be respectively assigned 5% and 10 % of the base station transmitted power. We also assume an additional DPCH beamforming gain of 6 dB, yielding to a power offset between DPCCH and CPICH equal to $\sigma^2_{h_{c,k}} / \sigma^2_{h_{d,k}} = 0.5$ for all k's, so that $\zeta = \zeta_k = \sqrt{2}|\alpha_0|$. Channels are randomly generated from the power delay profile of the *UMTS Vehicular A* channel [1]. Temporal correlation coeffcients $\rho = 0.99$ and $\rho = 0.9$ correspond in Jakes model to vehicle speeds 29km/h and 92km/h for a transmission at 1.8GHz. The legends {Dedicated LS, Kalman Filtering of Dedicated LS, Kalman Smoothing of Dedicated LS, ULMMSE Combining of Dedicated LS, Kalman Filtering After ULMMSE, Kalman Smoothing After ULMMSE, ULMMSE After Kalman Filtering, ULMMSE After Kalman Smoothing, Optimal Kalman Filtering, Optimal Kalman Smoothing} on the figures corresponds in the same order to the NMSE performances of the channel estimates $\{\hat{\boldsymbol{h}}_d(n), \tilde{\boldsymbol{h}}_d(n \mid n), \tilde{\boldsymbol{h}}_d(n \mid n+1), \bar{\boldsymbol{h}}_d(n), \bar{\bar{\boldsymbol{h}}}_d(n \mid n), \bar{\bar{\boldsymbol{h}}}_d(n \mid n+1), \tilde{\tilde{\boldsymbol{h}}}_d(n \mid n), \tilde{\tilde{\boldsymbol{h}}}_d(n \mid n+1), \hat{\hat{\boldsymbol{h}}}_d(n \mid n), \hat{\hat{\boldsymbol{h}}}_d(n \mid n+1)\}$ at the steady states of the EM-Kalman procedures.

Some interpretations of figures are as follows:

- $\bar{\boldsymbol{h}}_d(n)$ brings moderate improvement w.r.t $\hat{\boldsymbol{h}}_d(n)$ at reasonably high cross correlations.
- $\tilde{\boldsymbol{h}}_d(n \mid n)$ performs much better than $\bar{\boldsymbol{h}}_d(n)$.
- $\hat{\hat{\boldsymbol{h}}}_d(n \mid n)$ is the best (optimal causal filter in MMSE sense), performs also as close as 0.2dB w.r.t. knowing the ρ and $\boldsymbol{Q}$ parameters (latter case not shown on the plots), but it is at the same time the most complex.
- $\bar{\bar{\boldsymbol{h}}}_d(n \mid n)$ is equivalent to $\hat{\hat{\boldsymbol{h}}}_d(n \mid n)$ when the channels associated with DPCH and CPICH are fully correlated. Their performance difference is non-negligible only when the Doppler spread and DPCH-CPICH correlations are both low. It is attractive also for the non-beamforming case, especially in order to increase the coverage since in that case DPCH power can become comparable to or even exceed the CPICH power at cell edges due to power control.
- $\tilde{\tilde{\boldsymbol{h}}}_d(n \mid n)$ performs better than the $\bar{\bar{\boldsymbol{h}}}_d(n \mid n)$ when DPCH and CPICH are not very much correlated.
- smoothing (backward pass) improves the performance w.r.t filtering (only forward pass) in all the cases.

All the three methods are feasible for implementation since complexity is proportional to the number of channel taps and Kalman state vectors for each tap have at most two elements.

Possible extensions to the covered schemes are

- straightforward extension for 3 or more pilot sequences in case of one or more S-CPICH assignments
- handling channel variations within the slot
- taking into account also the correlations among FIR channel taps
- sparsification and hybrid treatment of different taps

REFERENCES

[1] "3GPP Technical Specifications," available at: http://www.3gpp.org/specs/specs.html

[2] G. Montalbano and D.T.M. Slock, "Joint Common-Dedicated Pilots Based Estimation of Time-Varying Channels for W-CDMA Receivers,"*IEEE VTC'03 Fall*, Orlando, FL, Oct. 2003.

[3] J. Bsltersee, G. Fock, P. Schultz-Rittich, and H. Meyr, "Performance analysis of phasor estimation algorithms for FDD-UMTS RAKE receiver," *IEEE 6th Symp. on Spread Spectrum Technologies and Applications*, NJIT, NJ, September 2000.

[4] M. Lenardi and D.T.M. Slock, "Estimation of time-varying wireless channels and application to the UMTS WCDMA FDD downlink," *ECWT'02*, Firenze, Italy, Feb. 2002.

[5] M. Benthin and K. Kammayer, "Influence of channel estimation on the performance of a coherent DS-CDMA system," *IEEE Trans. on VT,* Vol. 46, No. 2, pp. 262–268, May 1997

[6] M. Usada, Y. Ishikawa and S. Onoe, "Optimizing the number of dedicated pilot symbols for forward link in WCDMA systems," *IEEE VTC Spring'00*

[7] K. A. Qaraqe and S. Roe, "Channel estimation algorithms for third generation WCDMA communication systems," *IEEE VTC Fall'01*

[8] J. G. Proakis *Digital Communications*, NY: McGraw-Hill, 3rd ed., 1995

[9] B.R. Musicus, "Iterative Algorithms for Optimal Signal Reconstruction and Parameter Identification Given Noisy and Incomplete Data," Ph.D. Thesis, Mass. Inst. Tech., Sept. 1982

[10] R.H. Shumway and D.S.Stoffer, "An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm," *J. Time Series Anal.,* Vol.3, no.4, pp.253-264, 1982

[11] E. Weinstein, A.V. Oppenheim, M. Feder and J.R. Buck, "Iterative and Sequential Algorithms for Multisensor Signal Enhancement," *IEEE Trans. Signal Proc.,* Vol.42, no.4, pp.846-859,April 1994

[12] V. Digalakis, J.R. Rohlicek and M. Ostendorf,"ML Estimation of a Stochastic Linear System with the EM Algorithm and Its Application to Speech Recognition," *IEEE Trans. Speech and Audio Proc.,*Vol.1, no.4, October 1993

[13] W. Gao, S.Tsai and J.S. Lehnert, " Diversity Cobining for DS/SS Systems With Time-Varying Correlated Fading Branches," *IEEE Trans. Communications,* Vol.51, no.2, Feb. 2003

[14] A.P. Dempster, N.M. Laird and D.B. Rubin, "Maximum Likelihood From Incomplete Data via the EM Algorithm," *J. Roy. Statist. B,* vol.39, no.1, pp.1-38,1077

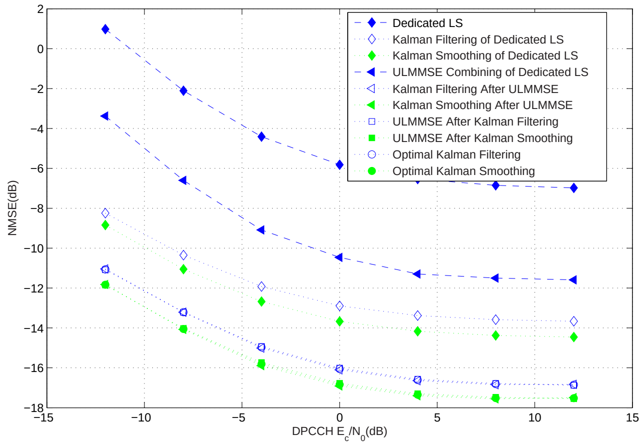

Figure 2: NMSE vs DPCCH E_c/N_0, $r = 1$, $\rho = 0.99$

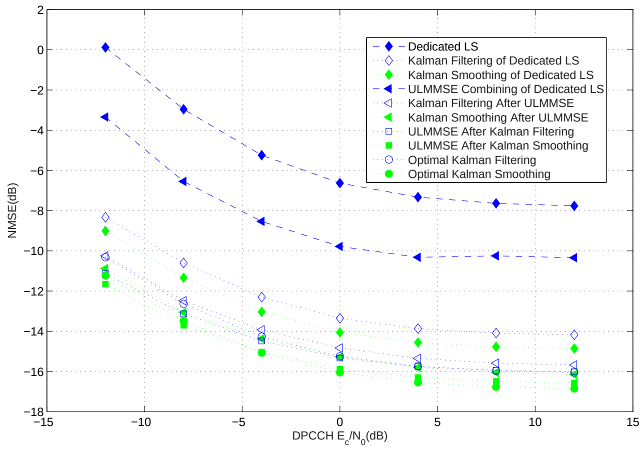

Figure 3: NMSE vs DPCCH E_c/N_0, $r = 0.95$, $\rho = 0.9$

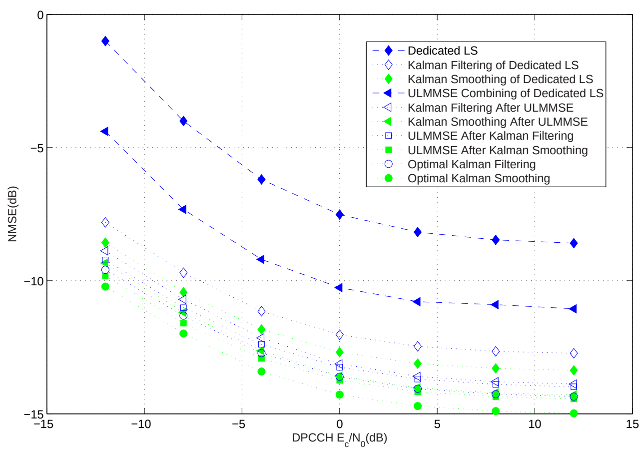

Figure 4: NMSE vs DPCCH E_c/N_0, $r = 0.9$, $\rho = 0.9$

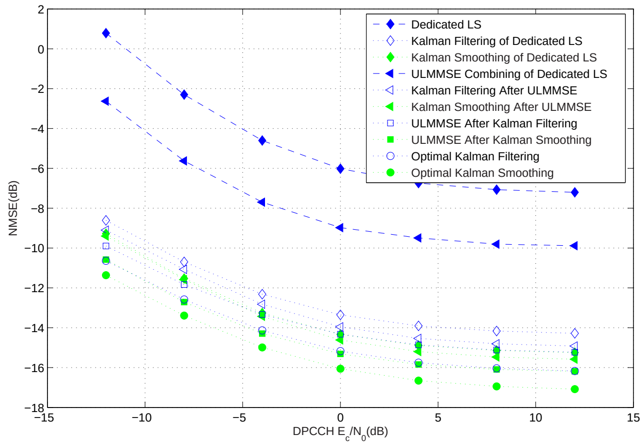

Figure 5: NMSE vs DPCCH E_c/N_0, $r = 0.9$, $\rho = 0.99$

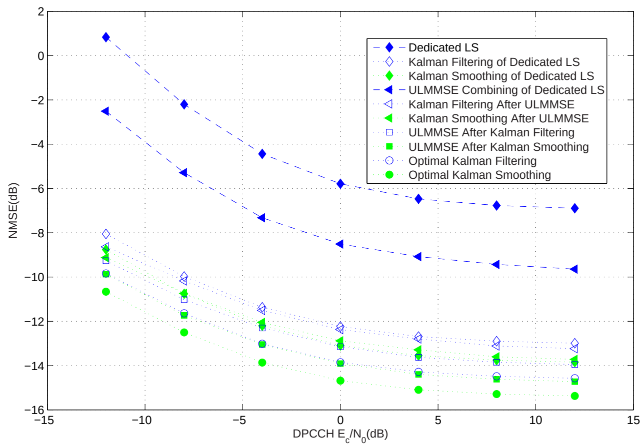

Figure 6: NMSE vs DPCCH E_c/N_0, $r = 0.8$, $\rho = 0.99$

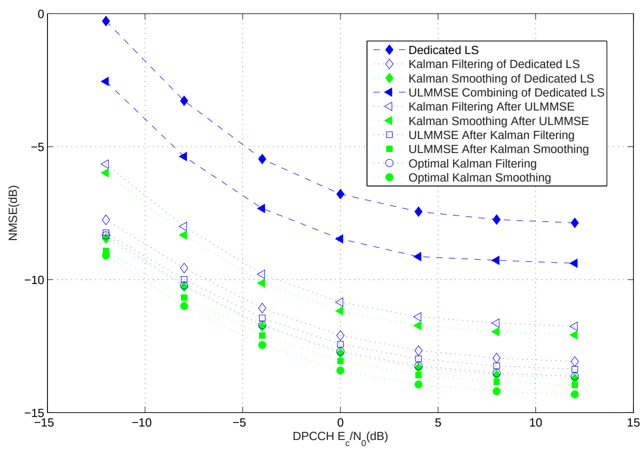

Figure 7: NMSE vs DPCCH E_c/N_0, $r = 0.6$, $\rho = 0.9$