Modified Kalman Filters for Channel Estimation in Orthogonal Space-Time Coded Systems

Murilo B. Loiola*, *Member, IEEE,* Renato R. Lopes,
and João M. T. Romano, *Senior Member, IEEE*

Abstract

We derive and analyze two modified Kalman channel estimators (KCE) for time-varying, flat, spatially correlated MIMO channels in systems employing orthogonal space-time block codes: the steady-state KCE, which is less complex than the KCE, and the fading memory KCE, which is more robust to model mismatch.

Index Terms

Channel tracking, Kalman filtering, space-time coding, MIMO systems, algebraic Riccati equation, steady-state Kalman filter, fading memory Kalman filter

I. Introduction

Orthogonal space-time block codes (OSTBC) [1]–[3] achieve full spatial diversity in multiple-input, multiple-output (MIMO) wireless systems with low complexity, since their maximum likelihood (ML) receiver consists of a linear processing followed by a symbol-by-symbol decoder. However, the receiver

This work was supported in part by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP), Brazil, under grant $05/55310-8$ and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), under grant 304155/2009-8.

*M. B. Loiola was with the School of Electrical and Computer Engineering, University of Campinas. He is now with Universidade Federal do ABC, Avenida dos Estados, 5001, Bloco A, torre 1, sala 723 – CEP 09210-170, Santo André, Brazil. Fax: +55 11 44960089; (e-mail:murilo.loiola@ufabc.edu.br).

R. R. Lopes and J. M. T. Romano are with the School of Electrical and Computer Engineering, University of Campinas, Avenida Albert Einstein, 400 CEP 13083-852, Campinas, Brazil. Phone/Fax: +55 19 35213857 (e-mails: rlopes@decom.fee.unicamp.br; romano@dmo.fee.unicamp.br).

Manuscript received January 21, 2011.

for OSTBCs requires channel knowledge, so that channel estimation techniques are essential. When the channel is static, estimators such as those presented in [4] and [5] can be successfully used. However, in wireless communications the channel is often time-varying. In this case, the estimation algorithm must be able to track the channel variations.

Kalman filters [6], [7] are widely used for channel tracking, especially due to their ability to deal with nonstationary environments. In [8] for instance, the authors derive a Kalman channel estimator (KCE) that uses the outputs of a minimum mean square error (MMSE) decision-feedback equalizer (DFE) to track Ricean MIMO frequency-selective channels. Channel estimation using KCEs for MIMO-OFDM systems is studied in [9]. Kalman filters can also be used to track the channel in MIMO systems employing OSTBCs. In [10], a KCE is used to estimate fast flat fading MIMO channels in systems with two transmit antennas employing Alamouti code. This KCE is generalized in [11] for any type of OSTBCs and spatially uncorrelated channels. A KCE for correlated channels was derived in [12].

As with most KCEs, the KCE in [12] is a time-varying filter whose coefficient matrices need to be computed anew for each time instant. This computation, which involve a matrix inversion, increases the complexity of the filter. To reduce the complexity, a steady-state Kalman channel estimator (SS-KCE) [6] is also derived in [12]. The SS-KCE is a time-invariant filter, whose coefficients are given by the asymptotic value of the filter matrices. In spite of the significant complexity reduction, it is shown in [12] that the SS-KCE suffers negligible performance degradation compared to the regular KCE, especially when channel variations are fast.

However, the SS-KCE of [12] depends on the solution of a Riccati equation. The first contribution of the present paper is to derive an explicit expression for the SS-KCE, and to prove that, under mild conditions, the SS-KCE is stable. We also prove that, at worst, the SS-KCE is marginally stable, but it is never unstable.

Besides complexity, another drawback of the KCEs mentioned so far is that they rely on an autoregressive model of the channel dynamics [8], [11]. Unfortunately, this model is only an approximation of the actual dynamics behind the channel variations. To improve the KCE robustness to channel model mismatch, in this paper we derive a fading-memory Kalman channel estimator (FM-KCE) [6] for estimating MIMO channels with OSTBC. This filter attributes a larger value to the variance of the process noise in the state equation, so that the filter must rely more on the measurement than on the prediction step of the KCE. As a result, the filter is more robust to model mismatch. We also derive a steady-state version of the FM-KCE, and present simulation results that attest the improved performance of the fading memory estimator when compared to the KCE.

The remainder of the paper is organized as follows: in section II we present the system model, briefly describing orthogonal space-time block codes and an autoregressive model for spatially correlated and time-varying MIMO channels. In section III we derive the steady-state Kalman filter and analyze the conditions for its existence. In section IV, we derive the fading memory Kalman channel estimator. Simulation results are shown in section V. Finally, section VI concludes the paper.

II. System Model

We consider a MIMO system with N_T transmit antennas sending data blocks of length T to N_R receive antennas, through a frequency flat channel. The received signal for data block k, $\mathbf{Y}_k$, can be expressed as [3]

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{N}_k, \tag{1}$$

where $\mathbf{X}_k$ is an $N_T \times T$ space-time codeword, the elements of the $N_R \times T$ matrix $\mathbf{N}_k$ are independent, zero mean, circularly symmetric, white Gaussian noise with variance σ_n^2, and the channel is represented by the $N_R \times N_T$ matrix $\mathbf{H}_k$. We assume the use of an OSTBC [2]–[4], so the elements of $\mathbf{X}_k$ are linear combinations of the transmitted information symbols in the data block $\mathbf{x}_k$ and their complex conjugates. Also, the codewords satisfy $\mathbf{X}_k \mathbf{X}_k^{\mathsf{H}} = \|\mathbf{x}_k\|^2 \mathbf{I}_{N_T}$, where $\mathbf{I}_{N_T}$ is the identity matrix of order N_T, $\|\cdot\|$ represents the Euclidean norm and $(\cdot)^{\mathsf{H}}$ denotes the conjugate transpose of a matrix.

The channel matrix $\mathbf{H}_k$ is assumed to be fixed during the transmission of a data block, and is assumed to change between blocks. We use the wide-sense stationary uncorrelated scattering (WSSUS) model [13], where the channel coefficients are modeled as zero-mean, complex Gaussian random variables with time autocorrelation function

$$\mathrm{E}\left[h_{k,i,j} h_{t,i,j}^*\right] \approx \mathcal{J}_0(2\pi f_D T_s |k-t|), \tag{2}$$

where $h_{k,i,j}$, $i = 1, \ldots, N_R$, $j = 1, \ldots, N_T$ is the (i,j) element of the channel matrix $\mathbf{H}_k$, $\mathcal{J}_0$ is the zero-order Bessel function of the first kind, $f_D T_s$ is the normalized Doppler rate (assumed the same for all transmit-receive antenna pairs) and T_s is the time necessary to transmit the space-time codeword $\mathbf{X}_k$. Note that we assume that the channel coefficients have unit variance. Spatial correlation is captured by the matrix $\mathbf{R}_{\mathbf{h}} = \mathrm{E}\left[\mathbf{h}_k \mathbf{h}_k^{\mathsf{H}}\right]$, where $\mathbf{h}_k = \mathrm{vec}(\mathbf{H}_k)$ is the vector obtained by stacking the columns of $\mathbf{H}_k$ on top of each other. As in [9], [12], we approximate the autocorrelation function (2) by the first order autoregressive (AR) process given by

$$\mathbf{h}_k = \beta \mathbf{h}_{k-1} + \mathbf{G}\mathbf{w}_k, \tag{3}$$

where $\beta = \mathcal{J}_0(2\pi f_D T_s)$, $\mathbf{w}_k$ is a vector of length $N_R N_T$ containing samples of circularly symmetric, zero-mean, white Gaussian excitation noise with covariance matrix $\mathbf{Q} = \sigma_w^2 \mathbf{I}_{N_R N_T}$, and $\sigma_w^2 = (1 - \beta^2)$.

The matrix $\mathbf{G}$ is such that $\mathbf{R_h} = \mathbf{G}\mathbf{G}^{\mathsf{H}}$. Note that $\beta = 1$ when $f_D = 0$, i.e., when there is no mobility. In all other cases, we have $|\beta| < 1$. Although higher-order AR models provide better approximation to (2) and the extension of (3) to that higher-order models is straightforward [8], we use the AR(1) model due to its simplicity. In Section IV, we derive a channel estimator to cope with this modeling error.

III. Steady-state Kalman channel estimator

Equation (3) can be seen as a process equation in a state-space description of the channel dynamics [6], [7]. To complete the state-space model required by the Kalman filter, we need to derive the measurement equation. The system output, in our case, is the channel output matrix $\mathbf{Y}_k$ in (1). Thus, as in [12], the measurement equation is formed by $\mathbf{y}_k = \mathrm{vec}(\mathbf{Y}_k)$, resulting in

$$\mathbf{y}_k = \boldsymbol{\mathcal{X}}_k \mathbf{h}_k + \mathbf{n}_k, \tag{4}$$

where $\boldsymbol{\mathcal{X}}_k = \mathbf{X}_k^{\mathsf{T}} \otimes \mathbf{I}_{N_R}$ and $\mathbf{R}_n = \sigma_n^2 \mathbf{I}_{N_R N_T}$ is the covariance matrix of the measurement noise $\mathbf{n}_k$, and $\otimes$ represents the Kronecker product [14].

The state equation (3) and the observation equation (4) are linear functions of the state vector $\mathbf{h}_k$, and the noises $\mathbf{w}_k$ and $\mathbf{n}_k$ are white, Gaussian and mutually statistically independent. Thus, the Kalman filter provides the optimal recursive estimates, in the MMSE sense, for the channel coefficients [6], [7]. Using the orthogonality of the transformed space-time codeword $\boldsymbol{\mathcal{X}}_k$ [12], the Kalman channel estimator (KCE) for correlated MIMO-OSTBC systems can be written as [12]

$$\mathbf{P}_{k|k-1} = \beta^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{R_h} \tag{5a}$$

$$\mathbf{A}_k = \mathbf{P}_{k|k-1} \left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2} \mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1} \right)^{-1} \tag{5b}$$

$$\mathbf{P}_{k|k} = \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_k \right) \mathbf{P}_{k|k-1} \tag{5c}$$

$$\hat{\mathbf{h}}_{k|k} = \beta \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_k \right) \hat{\mathbf{h}}_{k-1|k-1} + \mathbf{A}_k \frac{\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \mathbf{y}_k}{\|\mathbf{x}_k\|^2} \tag{5d}$$

In these equations, $\mathbf{P}_{i|j}$ an $\hat{\mathbf{h}}_{i|j}$ are, respectively, the prediction error covariance matrix and estimated channel vector at a data block i given the observation up to block j, and $\mathbf{A}_k$ is simply an auxiliary matrix.

The matrices in (5a)–(5c), and in particular the matrix inversion in (5b), have to be computed at every iteration. However, for constant modulus constellations, $n_s \triangleq \|\mathbf{x}_k\|^2$ is constant, so (5a)–(5c) represent a time-invariant system that quickly converges to a constant value. To reduce complexity, the steady-state Kalman filter [6] computes the asymptotic values of these equations, and uses these values to update the estimated channel in (5d).

Now, as shown in [12], the steady-state value of $\mathbf{P}_{k|k-1}$, denoted by $\mathbf{P}_\infty$, is a solution of the following discrete algebraic Riccati equation (DARE) [6], [7]:

$$\mathbf{P}_\infty = \beta^2 \mathbf{P}_\infty - \beta^2 \mathbf{P}_\infty \left(\mathbf{P}_\infty + \frac{\sigma_n^2}{n_s} \mathbf{I}_{N_R N_T} \right)^{-1} \mathbf{P}_\infty + \sigma_w^2 \mathbf{R}_\mathbf{h}. \tag{6}$$

If (6) can be solved, we can use $\mathbf{P}_\infty$ in (5b) to calculate the steady-state value of $\mathbf{A}$, denoted by $\mathbf{A}_\infty$. In this case, the proposed steady-state Kalman channel estimator (SS-KCE) is given simply by

$$\hat{\mathbf{h}}_{k|k} = \beta \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_\infty \right) \hat{\mathbf{h}}_{k-1|k-1} + \mathbf{A}_\infty \frac{\boldsymbol{\mathcal{X}}_k^\mathsf{H} \mathbf{y}_k}{\|\mathbf{x}_k\|^2}. \tag{7}$$

In the next section, we present the solution to (6), and discuss the stability of the resulting filter (7).

We note that the KCE and the SS-KCE can operate in both training and decision-directed (DD) modes. First, when pilot symbols are available, the matrix $\boldsymbol{\mathcal{X}}_k$ is known. Once the transmission of pilot symbols is finished, the algorithms enter in DD mode, where the matrix $\boldsymbol{\mathcal{X}}_k$ is formed by the decisions provided by the ML space-time decoder. Possible decoding errors in these decisions could be modeled as an extra term in the measurement noise, which could then be non-Gaussin. These decoding errors could be minimized by the use of a turbo-like receiver, where the decisions used by the channel estimator would be iteratively improved by a soft-decision decoder of an outer code. However, this solution demands more computational power. Hence, in this paper we use the usual simplifying assumption, as in [8], that the decisions provided by the ML space-time decoder are correct. Despite this simplification, the channel estimators proposed in this paper present very good performances, as can be seen in Section V.

A. Derivation and Analysis of the SS-KCE

We now begin extending [12], showing the only possible solution of the DARE in (6) that results in a valid covariance matrix $\mathbf{P}_\infty$. Then, we show that this solution leads to a stable filter in most cases of interest. To begin the derivation, let $\mathbf{R}_\mathbf{h} = \mathbf{Q}^\mathsf{H} \boldsymbol{\Lambda} \mathbf{Q}$ be the eigendecomposition of $\mathbf{R}_\mathbf{h}$. Since $\mathbf{Q}$ is unitary, it is easy to verify that $\mathbf{P}_\infty = \mathbf{Q}^\mathsf{H} \boldsymbol{\Sigma} \mathbf{Q}$ is a solution of the DARE (6), as long as the diagonal matriz $\boldsymbol{\Sigma}$ satisfies

$$\boldsymbol{\Sigma} = \beta^2 \boldsymbol{\Sigma} - \beta^2 \boldsymbol{\Sigma} \left(\boldsymbol{\Sigma} + \frac{\sigma_n^2}{n_s} \mathbf{I}_{N_R N_T} \right)^{-1} \boldsymbol{\Sigma} + \sigma_w^2 \boldsymbol{\Lambda}. \tag{8}$$

Now let σ_i and λ_i be the i-th diagonal element of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$, respectively. Then, since all the matrices in (8) are diagonal, σ_i must satisfy

$$\sigma_i^2 + b\sigma_i + c = 0, \tag{9}$$

where $b = \sigma_n^2(1-\beta^2)/n_s - \sigma_w^2 \lambda_i$ and $c = -\sigma_n^2 \sigma_w^2 \lambda_i / n_s$. In other words, the eigenvalues of $\mathbf{P}_\infty$ are given by

$$\sigma_i = \frac{-b \pm \sqrt{b^2 - 4c}}{2}. \tag{10}$$

We now show that only one of the solutions in (10) yield a valid autocorrelation matrix $\mathbf{P}_\infty$, with eigenvalues σ_i that are real and non-negative. We begin by noting that $\mathbf{R_h}$ is a correlation matrix, so $\lambda_i \geq 0$. We then identify three possibilities, all leading to a valid $\mathbf{P}_\infty$:

- When there is no mobility, $\beta = 1$ and $\sigma_w^2 = 0$, so that $b = c = 0$. In this case, $\sigma_i = 0$.
- When $\lambda_i = 0$ and there is mobility, *i.e.*, $\beta < 1$, we have that $b > 0$ and $c = 0$. In this case, σ_i may be equal to $-b$ or 0. The second solution is the only valid one.
- When $\lambda_i > 0$ and there is mobility, *i.e.*, $\beta < 1$, we have that $c < 0$. Furthermore, since $c < 0$, we have that $b^2 < b^2 - 4c$, so the solution given by $\sigma_i = (-b + \sqrt{b^2 - 4c})/2$ is non-negative.

We also need to prove that the SS-KCE in (7) is stable. To that end, note from (7) that stability holds as long as the eigenvalues of $\beta(\mathbf{I} - \mathbf{A}_\infty)$ have magnitude less than one. Now, using the fact that $\mathbf{P}_\infty = \mathbf{Q}^\mathsf{H}\mathbf{\Sigma}\mathbf{Q}$, it is easy to verify that the eigenvalues of $\beta(\mathbf{I} - \mathbf{A}_\infty)$, ρ_i, are given by

$$\rho_i = \beta \frac{\sigma_n^2/n_s}{\sigma_n^2/n_s + \sigma_i}. \tag{11}$$

Note that $\sigma_i \geq 0$, so that $0 < \rho_i \leq \beta$. Also, note that $\rho_i = 1$ if and only if $\sigma_i = 0$ and $\beta = 1$, which happens if and only if there is no mobility. In these cases, the SS-KCE is marginally stable. In all other cases, the filter is stable.

Finally, we note that the SS-KCE is more sensitive to the initial conditions than the KCE. In fact, in the KCE, the initial value $\mathbf{P}_{0|0}$ controls the influence of the initial value of the channel estimate. Large values of $\mathbf{P}_{0|0}$ indicate low reliability of $\hat{\mathbf{h}}_{0|0}$, which in turn decreases the impact of $\hat{\mathbf{h}}_{0|0}$ in the computation of $\hat{\mathbf{h}}_{1|1}$. In the limit, when $\mathbf{P}_{0|0}$ grows very large, $\mathbf{A}_1 \to \mathbf{I}$, so that $\hat{\mathbf{h}}_{1|1}$ tends to $\boldsymbol{\mathcal{X}}_1^\mathsf{H}\mathbf{y}_1/\|\mathbf{x}_1\|^2$ [1]. However, $\mathbf{P}_{k|k}$ is fixed in the SS-KCE, so the impact of the initial value of the channel estimate cannot be controlled. Thus, the initialization should be chosen carefully. To that end, we propose using the one-shot maximum-likelihood estimate $\hat{\mathbf{h}}_{1|1} = \boldsymbol{\mathcal{X}}_1^\mathsf{H}\mathbf{y}_1/\|\mathbf{x}_1\|^2$ in the first iteration. Equation (7) is used to update $\hat{\mathbf{h}}_{k|k}$ only when $k \geq 2$.

IV. Fading Memory Kalman Channel Estimator

As mentioned in Section II, the first order AR model used in (3) is only an approximate description of the time evolution of channel coefficients. This modeling error can degrade the performance of the proposed channel estimators. One possible solution to mitigate this performance degradation in the KCE is to give more emphasis to the most recent received data, thus increasing the importance of the observations and decreasing the importance of the process equation [6], [15]. This can be accomplished with an

[1]Note that this is the maximum likelihood estimate of the channel given only the observation $\mathbf{y}_1$.

exponential data weighting, controlled by a scalar $\alpha \geq 1$, which increases the importance of the most recent observations by giving them a higher weight than past data [6], [15].

Hence, following [6], [15], it is possible to show that this exponential data weighting for OSTBC systems leads to the fading-memory Kalman channel estimator (FM-KCE), given by

$$\mathbf{P}_{k|k-1} = (\alpha\beta)^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{R}_{\mathbf{h}} \tag{12a}$$

$$\mathbf{A}_k = \mathbf{P}_{k|k-1} \left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2} \mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1} \right)^{-1} \tag{12b}$$

$$\hat{\mathbf{h}}_{k|k} = \beta \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_k \right) \hat{\mathbf{h}}_{k-1|k-1} + \mathbf{A}_k \frac{\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \mathbf{y}_k}{\|\mathbf{s}_k\|^2} \tag{12c}$$

$$\mathbf{P}_{k|k} = \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_k \right) \mathbf{P}_{k|k-1} \tag{12d}$$

The only difference between the KCE and the FM-KCE is the existence of the scalar α^2 in the update equation of prediction error covariance matrix of the FM-KCE in (12a). This increases the variance of the prediction error, to which the filter responds by giving less importance to the system equation. The same could also be accomplished by using a system equation with a noise term of increased variance, which could be interpreted as a result of adding a fictitious process noise [6], [15]. It is worth noting that when $\alpha = 1$, the FM-KCE reduces to the KCE. On the other hand, when $\alpha \to \infty$, the channel estimates provided by the FM-KCE are solely based on the received signals, and the system model is not taken into account.

The FM-KCE also has a steady-state version. In fact, noting the similarities between the FM-KCE (12a)–(12d) and the KCE in (5a)–(5d), and following the same steps described in Section III for the derivation of (6), it can be shown that the Riccati equation for the FM-KCE is given by

$$\mathbf{P}_\infty = (\alpha\beta)^2 \mathbf{P}_\infty - (\alpha\beta)^2 \mathbf{P}_\infty \left(\mathbf{P}_\infty + \frac{\sigma_n^2}{n_s} \mathbf{I}_{N_R N_T} \right)^{-1} \mathbf{P}_\infty + \sigma_w^2 \mathbf{R}_{\mathbf{h}}. \tag{13}$$

The solution of the DARE in (13) is also of the form $\mathbf{P}_\infty = \mathbf{Q}^{\mathsf{H}} \boldsymbol{\Sigma} \mathbf{Q}$. As was done in section III-A, it can be shown that the elements of the diagonal matrix $\boldsymbol{\Sigma}$ are given by $\sigma_i = (-b + \sqrt{b^2 - 4c})/2$, where $b = \sigma_n^2(1 - \alpha^2\beta^2)/n_s - \sigma_w^2 \lambda_i$ and $c = -\sigma_n^2 \sigma_w^2 \lambda_i / n_s$. Furthermore, when $c \neq 0$, *i.e.*, when there is mobility and $\mathbf{R}_{\mathbf{h}}$ is full rank, we can follow the steps in section III-A to show that the filter has a unique solution that leads to a stable filter. The steps in section III-A can also be followed when $c = 0$ and $\alpha\beta \leq 1$ to show that the filter has a unique solution that leads to a marginally stable filter.

However, when $c = 0$ and $\alpha\beta \leq 1$, b becomes a negative number, so that $-b$ leads to a valid autocorrelation matrix that satisfies (13). In this case, it can be shown that the eigenvalues of $\mathbf{P}_{k|k-1}$ remain at zero when initialized at this value, and they converge to $-b$ when initialized at a non-zero value. In other words, the value of P_∞ depends on the initialization. Since this initialization is generally

of full rank, normally the steady-state value will be given by $-b$. Also, note that, as $-b > 0$, the resulting filter is stable, even if, in this case, there is no mobility or $\mathbf{R_h}$ is not full rank.

V. Simulation results

In this section, we present simulation results that illustrate the performance of the proposed channel estimation algorithms. The simulation results presented in the sequel correspond to averages of 10 channel realizations, in each of which we simulate the transmission of 1×10^6 orthogonal space-time codewords. Unless stated otherwise, we insert 25 OSTBC training codewords between every 225 OSTBC data codewords. During the detection of the data codewords, the channel estimators operate in the decision-directed mode, using the detected data as training information. We assume that the receiver has perfect knowledge of the variances of process and measurement noises, the spatial correlation matrix and the normalized Doppler rate $f_D T_s$[2].

The spatial correlation is modeled as [3] $\mathbf{H}_k = \mathbf{R}_R^{1/2} \mathbf{H}_k^{\text{ind}} \mathbf{R}_T^{\mathsf{T}/2}$, where $\mathbf{R}_R$ models the correlation between receive antennas, $\mathbf{R}_T$ models the correlation between transmit antennas, $(\cdot)^{1/2}$ stands for the Hermitian square root of a matrix [14] and $\mathbf{H}_k^{\text{ind}}$ is a MIMO channel with spatially independent and unit variance Gaussian elements with time autocorrelation function given by WSSUS model in (2). As shown in [12], this results in a channel that also satisfies (3), with spatial autocorrelation matrix given by $\mathbf{R_h} = \mathbf{G}\mathbf{G}^{\mathsf{H}}$, where $\mathbf{G} = \mathbf{R}_T^{1/2} \otimes \mathbf{R}_R^{1/2}$. We further assume that the spatial correlation coefficient between any two adjacent receive (transmit) antennas is given by p_r (p_t), so that each (i, j) element of the spatial correlation matrix R_R (R_T) can be expressed as $p_r^{|i-j|}$, $i, j = 1, \ldots, N_R$ ($p_t^{|i-j|}$, $i, j = 1, \ldots, N_T$).

To illustrate the performance of the Kalman channel estimator and its steady-state version, we simulate a system sending QPSK symbols from $N_T = 4$ transmit to $N_R = 4$ receive antennas. For comparison purposes, we also simulate a channel estimator implemented by the well known RLS adaptive filter [17], with a forgetting factor of 0.98. This value was determined by trial and error to yield the best performance of the RLS. We employ the $1/2$-rate OSTBC of [2] and assume $p_t = 0.8$ and $p_r = 0.4$. The mean squared error (MSE) for the RLS and the SS-KCE is shown in Fig. 1. The curves for the KCE are undistinguishable from those of the SS-KCE, so they are not shown in the figure. We observe that the estimates produced by the RLS algorithm are affected by the rate of channel variation. On the other hand, for this scenario, the proposed SS-KCE has the same performance for both values of $f_D T_s$ considered and the MSE decreases with the SNR. The similar performances of SS-KCE for $f_D T_s = 0.0015$ and $f_D T_s = 0.0045$ are also reflected in the symbol error rates (SER) at the output of the ML decoders, as shown in Fig. 2. In terms

[2]These parameters can be estimated using, for example, the methods proposed in [16].

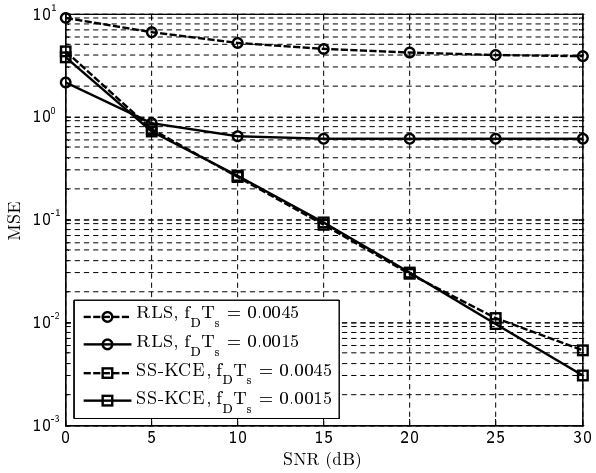

Fig. 1. Estimation MSE for different values of $f_D T_s$. The performances of the KCE and the SS-KCE cannot be distinguished.

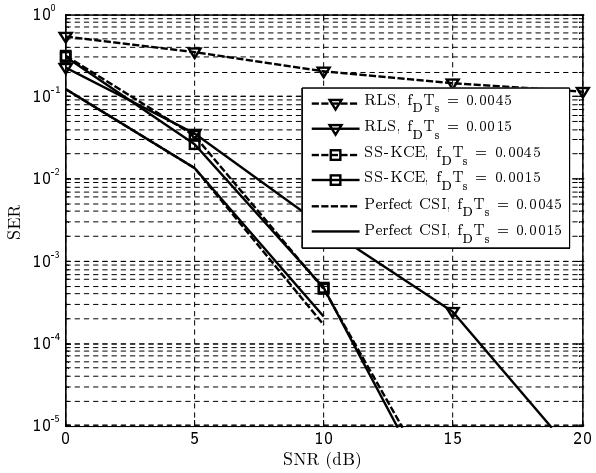

Fig. 2. Symbol error rate for different values of $f_D T_s$. The performances of the KCE and the SS-KCE cannot be distinguished.

of SER, the KCE also has the exact same performance as the SS-KCE. For an SER of 10^{-3}, the decoders using the channels estimates provided by the SS-KCE are about 1 dB from the curves of the ML decoders with perfect channel state information (CSI). For an SER of 10^{-3} and $f_D T_s = 0.0015$ the decoder fed with RLS channels estimates is approximately 4 dB from the optimal decoder, while for $f_D T_s = 0.0045$ the RLS-based decoder presents an SER no smaller than 10^{-1} in the simulated SNR range.

We can explain the performance equivalence of KCE and SS-KCE by the fast convergence of the matrix $\mathbf{P}_{k|k-1}$ to its steady-state value. This means that the SS-KCE uses the optimal value of $\mathbf{A}_k$ after just a

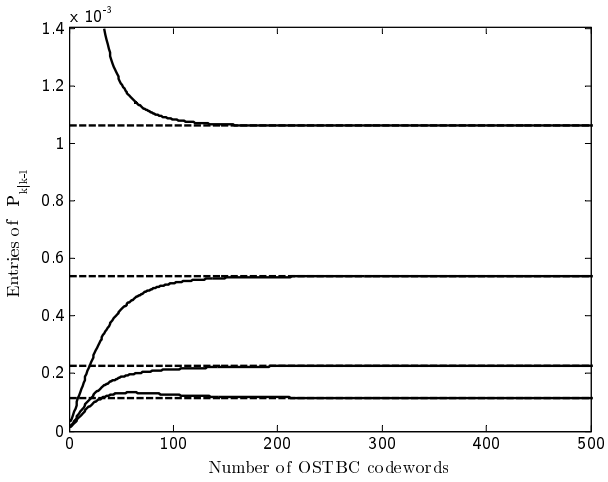

Fig. 3. Evolution of the entries of $\mathbf{P}_{k|k-1}$.

few blocks. Consequently, after these few blocks, the estimates provided by the SS-KCE are the same as those generated by the KCE. To exemplify the fast convergence of $\mathbf{P}_{k|k-1}$, Fig. 3 shows the evolution of the values of the elements of $\mathbf{P}_{k|k-1}$ for an 8-PSK, Alamouti coded system with $N_R = N_T = 2$, $f_D T_s = 0.0015$, $p_r = 0.4$, $p_t = 0.8$, SNR $= 15$ dB and with the initial condition $\mathbf{P}_{0|0} = \mathbf{I}_{N_R N_T}$. The dashed lines show the solution of the Riccati equation. It is clear from this figure that the elements of the matrix $\mathbf{P}_{k|k-1}$ reach their steady-state values before the transmission of 200 blocks. As the simulated system inserts 25 training blocks between 225 data blocks, we see that $\mathbf{P}_{k|k-1}$ converges to the solution of the Riccati equation even before the second training period.

In Section IV, we claimed that much of the estimation errors in the KCE is due to the modeling error introduced by the use of the first-order AR channel model. To cope with this error, we proposed the fading-memory estimator and its steady-state version. To illustrate the performance improvement of FM-KCE in comparison to the SS-KCE, we simulate a MIMO system with 2 transmit antennas sending Alamouti-coded QPSK symbols to 2 receive antennas. The normalized Doppler rate is set to 0.0015, the receiver correlation coefficient p_r is set to zero while the transmitter correlation coefficient assumes the value $p_t = 0.4$. To take into account a possible mismatch in the spatial channel model, we add a white noise with variance 0.1 to all elements of matrix $\mathbf{R}_T$, while maintaining it Hermitian. We vary the number of training codewords from 4 to 32 while maintaining the total number of blocks (training + data) fixed to 160 codewords. Also, we assume the weight of the FM-KCE $\alpha = 1.1$.

In Fig. 4 we present the estimation MSE for SS-KCE and for the steady-state version of FM-KCE, computed from the solution of the Riccati equation (13), with $4, 8, 12, 16, 20, 24, 28$ and 32 training

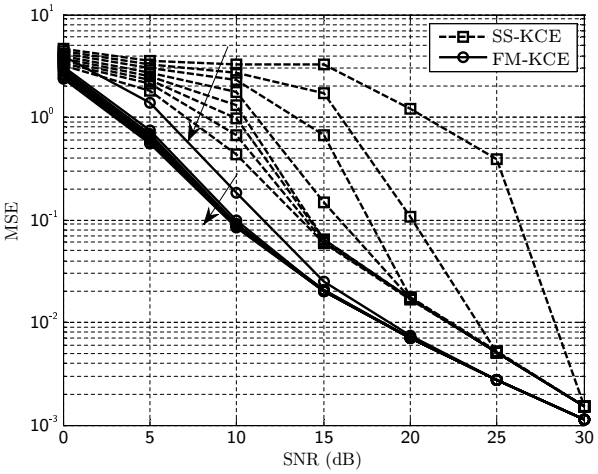

Fig. 4. Estimation MSE for SS-KCE and the steady-state version of FM-KCE.

codewords. The arrows in this figure indicate the number of training codewords in ascending order. From Fig. 4, the superiority of FM-KCE over SS-KCE is evident. Differently from SS-KCE, whose performance improves with the increase in the number of training codewords, the FM-KCE presents similar performances for the whole range of training codewords considered. For instance, for an MSE of 10^{-2} the FM-KCE performs 10 dB better than the SS-KCE with 4 training codewords and about 3.5 dB better than the SS-KCE with 32 training codewords.

The superior performance of the FM-KCE can also be observed in Fig. 5, which shows the SER at the output of ML decoders fed with CSI provided by SS-KCE and FM-KCE, as well as with perfect channel knowledge, for different training sequence lengths. For an SER of 10^{-3}, the receiver with the FM-KCE is about 1.3 dB from the decoder with perfect CSI, while the receiver using channel estimates provided by the SS-KCE presents performance losses of 2.2 and 8 dB from the decoder with perfect CSI for 32 and 4 training codewords, respectively. For an SER of 10^{-4}, the receiver with the FM-KCE performs 0.25 and 5 dB better than the receiver with SS-KCE for 32 and 4 training codewords, respectively, and presents a loss of 0.8 dB from the ML space-time decoder with perfect CSI. Thus, from Figs. 4 and 5, we see that the FM-KCE allows the use of a small number of training codewords without compromising the performance of the receiver.

VI. Conclusion

In this paper, we proposed channel estimation algorithms for systems employing orthogonal space-time block codes. The proposed algorithms are based on generalizations of the traditional Kalman filter.

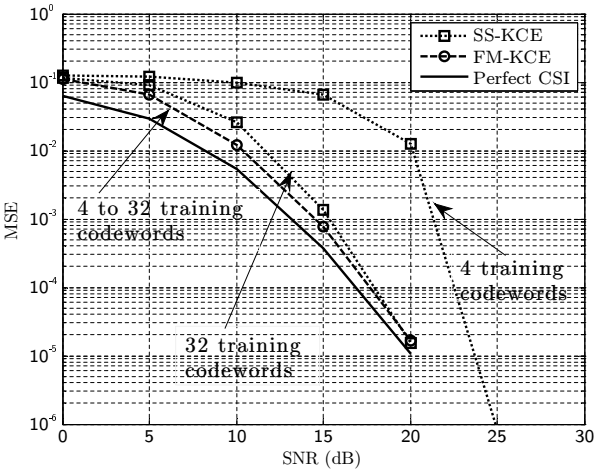

Fig. 5. Symbol error rate for SS-KCE and the steady-state version of FM-KCE.

The first is the steady-state filter, which is applicable to systems employing constant modulus signal constellations. In this case, several matrices in the Kalman filter quickly converge to steady-state values. These can be used instead of their time-varying counterparts, thus reducing the complexity of the filter. We derived explicit expressions for these steady-state values, showing that the resulting filter is stable in most scenarios of interest. We also proposed an efficient initialization of the filter. In our simulations, the performance of the Kalman estimator and its steady-state version cannot be distinguished.

The second generalization is the fading memory filter, which is more robust to errors in the channel model. This is particularly important in the estimation of wireless channels, since the first order autoregressive model for the channel dynamics, used in the derivation of the Kalman filters, is only a rough approximation to the channel dynamics. To achieve this robustness, the fading memory filter decreases the importance of the process equation in the estimation process. Simulation results show that the fading memory estimator outperforms the traditional Kalman filter by as much as 5 dB for a SER of 10^{-3}.

REFERENCES

[1] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 1451–1458, Oct. 1998.

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-Time Block Codes from Orthogonal Designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[3] E. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge University Press, 2003.

[4] B. Vucetic and J. Yuan, *Space-Time Coding*. John Wiley and Sons, 2003.

[5] E. Larsson, P. Stoica, and J. Li, "Orthogonal Space-Time Block Codes: Maximum Likelihood Detection for Unknown Channels and Unstructured Interferences," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 362–372, Feb. 2003.

[6] D. Simon, *Optimal State Estimation - Kalman, H_∞, and Nonlinear Approaches*. John Wiley and Sons, 2006.

[7] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[8] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-Input Multi-Output Fading Channel Tracking and Equalization Using Kalman Estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[9] M. Enescu, T. Roman, and V. Koivunen, "State-Space Approach to Spatially Correlated MIMO OFDM Channel Estimation," *Signal Processing*, vol. 87, no. 1, pp. 2272–2279, 2007.

[10] Z. Liu, X. Ma, and G. B. Giannakis, "Space-Time Coding and Kalman Filtering for Time-Selective Fading Channels," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 183–186, Feb. 2002.

[11] B. Balakumar, S. Shahbazpanahi, and T. Kirubarajan, "Joint MIMO Channel Tracking and Symbol Decoding Using Kalman Filtering," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5873–5879, Dec. 2007.

[12] M. B. Loiola, R. R. Lopes, and J. M. T. Romano, "Kalman Filter-Based Channel Tracking in MIMO-OSTBC Systems," in *Global Telecommun. Conf., GLOBECOM 2009*, Honolulu, 2009.

[13] W. C. Jakes, *Microwave Mobile Communications*. New York: John Wiley and Sons, 1974.

[14] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. John Hopkins University Press, 1996.

[15] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall, 1979.

[16] A. Jamoos, E. Grivel, W. Bobillet, and R. Guidorzi, "Errors-In-Variables-Based Approach for the Identification of AR Time-Varying Fading Channels," *IEEE Signal Process. Lett.*, vol. 14, no. 11, pp. 793–796, Nov. 2007.

[17] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice-Hall, 2002.