13

Adaptive Channel Estimation in Space-Time Coded MIMO Systems

Murilo B. Loiola[1] and Renato R. Lopes and João M. T. Romano[2]
[1]*Universidade Federal do ABC (UFABC)*
[2]*University of Campinas (UNICAMP)*
Brazil

1. Introduction

Space-time codes are an effective and practical way to exploit the benefits of spatial diversity in multiple-input, multiple-output (MIMO) systems. As the name suggests, the coding of space-time codes is performed in both the spatial and the temporal domains, introducing correlation between signals transmitted by different antennas at different time instants. At the receiver, this space-time correlation is exploited to mitigate the detrimental effects of fading and to improve the quality of received signal. The use of space-time codes allows the system to profit from transmit diversity without any power increment and with no channel knowledge at the transmitter.

Among the existing space-time coding schemes, orthogonal space-time block codes (OSTBCs) (Alamouti, 1998; Duman & Ghrayeb, 2007; Larsson & Stoica, 2003; Tarokh et al., 1999) are of particular interest because they achieve full diversity at low receiver complexity. More specifically, the maximum likelihood (ML) receiver for OSTBCs consists of a linear receiver followed by a symbol-by-symbol decoder.

In order to correctly decode the received signals, the ML receiver for OSTBCs must have perfect channel knowledge. Unfortunately, this channel information is not normally available to the receivers; therefore channel estimation techniques are essential for the system to work properly. When the channel is static, methods such as those presented in (Vucetic & Yuan, 2003) and (Larsson et al., 2003) can be successfully used. However, the channel may be time-varying due to the mobility of the transmitter and/or receiver, to changes in the environment, or to carrier frequency mismatch between transmitter and receiver. In these cases, the estimation algorithm must be able to track the channel variations. One of the most widely known approaches to channel tracking is Kalman filtering (Enescu et al., 2007; Kailath et al., 2000; Komninakis et al., 2002; Piechocki et al., 2003; Simon, 2006). An important characteristic of the Kalman filter (KF) is its inherent ability to deal with nonstationary environments. In (Komninakis et al., 2002) for instance, a Kalman filter that uses the outputs of a minimum mean squared error (MMSE) decision-feedback equalizer (DFE) is developed to track Rice MIMO frequency-selective channels. Channel estimation using Kalman filters for MIMO-OFDM systems is studied in (Enescu et al., 2007; Piechocki et al., 2003). The use of Kalman filters to estimate channels in space-time block coded MIMO systems is also developed in the literature. In (Liu et al., 2002), a KF is used to estimate fast flat fading MIMO channels in Alamouti-based schemes. Therefore, it is limited to the case of two transmit

antennas. An extension of (Liu et al., 2002) for any type of OSTBCs is presented in (Balakumar et al., 2007). It is also shown in (Balakumar et al., 2007) that the KF can be significantly simplified due to the orthogonality of OSTBC codewords.

The algorithms developed in (Liu et al., 2002) and (Balakumar et al., 2007) assume that the channel coefficients are uncorrelated and independent from each other. However, spatial correlation between transmit and/or receive antennas usually exists in practical scenarios. This can occur, for instance, if the separation of adjacent antennas is not sufficient to allow the signals undergo different channel fades. Besides affecting the performance of OSTBCs, spatial correlations reduce the capacity gains of MIMO channels. Furthermore, the complexity reduction in (Balakumar et al., 2007) cannot be achieved if the channels were correlated. A generalization of (Balakumar et al., 2007) for correlated channels is presented in (Loiola et al., 2009) with a complexity similar to that in (Balakumar et al., 2007).

As with most Kalman channel estimators (KCEs), the KCE in (Loiola et al., 2009), is a time-varying filter whose coefficients need to be computed anew for each time instant. This computation increases the complexity of the filter, especially because it involves a matrix inversion. However, in many scenarios with constant modulus signal constellations the filter coefficients converge quickly. Taking advantage of this fact, the authors in (Loiola et al., 2009) derive a steady-state Kalman channel estimator (SS-KCE) (Simon, 2006), where the asymptotic value of the filter coefficients are computed at the initialization and used in a time-invariant filter. In spite of the significant complexity reduction, it is shown in (Loiola et al., 2009) that the SS-KCE suffers negligible performance degradation compared to the regular KCE, especially when channel variations are fast. However, the SS-KCE of (Loiola et al., 2009) depends on the solution of a Riccati equation. This chapter extends (Loiola et al., 2009) by providing an explicit expression for the SS-KCE and by proving that, under mild conditions, the SS-KCE is stable. It also proves that, at worst, the SS-KCE is marginally stable, but it is never unstable.

In this chapter, we also present a KCE with improved robustness to channel model mismatch. In fact, all the KCEs mentioned so far rely on an autoregressive model of the channel dynamics (Enescu et al., 2007; Komninakis et al., 2002; Li & Wong, 2007). This model is an approximation of Bello's wide-sense stationary uncorrelated scattering (WSSUS) model (Jakes, 1974), which in turn is an approximation of what actually governs the channel variations. In other words, the KCEs rely on a channel dynamics that may be far from true. To mitigate this problem, in this chapter we develop a fading-memory Kalman channel estimator (FM-KCE) (Simon, 2006) for estimating MIMO channels with OSTBC. This filter artificially increases the process noise in the state equations, 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.

An outline of the chapter is as follows. We begin with a basic channel model and some definitions. We then proceed to developed the KCEs mentioned so far, including the SS-KCE and the FM-KCE. We also prove that the SS-KCE is never unstable. Simulations results are presented in the sequel. Finally, the last section presents a summary of the chapter.

2. System model

We consider a MIMO system with N_T transmit antennas sending data blocks of length T to N_R receive antennas. The channel is assumed to be flat and constant during the transmission of each data block and can change between consecutive blocks. The relationship between transmitted and received signals for a data block k can be expressed as (Duman & Ghrayeb,

2007; Larsson & Stoica, 2003)

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

where $\mathbf{Y}_k$ is an $N_R \times T$ matrix with the received signals, $\mathbf{X}_k$ is an $N_T \times T$ matrix containing the transmitted signals at block k, the $N_R \times T$ matrix $\mathbf{N}_k$ contains samples of independent, zero mean, circularly symmetric, white Gaussian noise with covariance σ_n^2, and the frequency-flat channel is represented by the $N_R \times N_T$ matrix $\mathbf{H}_k$.

In space-time block coding, the matrix $\mathbf{X}_k$ represents a mapping that transforms a block of M complex symbols, $\mathbf{x}_k = [x_{k,1}\ x_{k,2}\ \cdots\ x_{k,M}]^{\mathsf{T}}$, to an $N_T \times T$ complex matrix. The space-time codeword $\mathbf{X}_k$ is then used to transmit these M symbols in T time slots, achieving a rate of M/T. The matrix $\mathbf{X}_k$ is an OSTBC if (Duman & Ghrayeb, 2007; Larsson & Stoica, 2003; Tarokh et al., 1999; Vucetic & Yuan, 2003): 1) all elements of $\mathbf{X}_k$ are linear functions of symbols of $\mathbf{x}_k$ and their complex conjugates and 2) for an arbitrary $\mathbf{x}_k$, the matrix $\mathbf{X}_k$ satisfies $\mathbf{X}_k \mathbf{X}_k^{\mathsf{H}} = \|\mathbf{x}_k\|^2 \mathbf{I}_{N_T}$, 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. From the definition of OSTBC the space-time codeword $\mathbf{X}_k$ can be written as (Duman & Ghrayeb, 2007; Larsson & Stoica, 2003)

$$\mathbf{X}_k = \sum_{m=1}^{M} \left(x_{k,m} \mathbf{A}_m + x_{k,m}^* \mathbf{B}_m \right), \tag{2}$$

where $x_{k,m}, m = 1, \ldots, M$ is the m^{th} symbol of the data vector $\mathbf{x}_k$ and the $N_T \times T$ matrices $\mathbf{A}_m$ and $\mathbf{B}_m$ describe the code.

One of the main characteristics of OSTBCs is the simplicity of the decoder. More specifically, the ML space-time decoder for OSTBCs consists of a linear receiver followed by a symbol-by-symbol decoder (Duman & Ghrayeb, 2007; Larsson & Stoica, 2003; Tarokh et al., 1999; Vucetic & Yuan, 2003). The linear receiver generates sufficient statistics through a linear combination of the received signals. Then, these sufficient statistics are passed to a symbol-by-symbol decision device. It is important to highlight that the computation of the sufficient statistics depends on the channel knowledge at the receiver.

As mentioned before, the channel is assumed to be fixed during the transmission of a space-time codeword $\mathbf{X}_k$, but can vary between consecutive codewords. According to the widely used wide-sense stationary uncorrelated scattering (WSSUS) model (Jakes, 1974), the channel coefficients are modeled as independent, zero-mean, complex Gaussian random variables with time autocorrelation function

$$\mathrm{E}\left[h_{k,i,j} h_{t,i,j}^* \right] \approx \mathcal{J}_0(2\pi f_D T_s |k - t|), \tag{3}$$

where $h_{k,i,j}, i = 1, \ldots, N_R, j = 1, \ldots, N_T$ is the (i, j) element of 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 symbol period.

Although exact modeling of channel dynamics by finite length autoregressive (AR) processes is impossible because the time autocorrelation function (3) is nonrational and its spectrum is bandlimited, we can approximate the time evolution of channel coefficients by low-order AR processes. This is possible because the first few correlation terms of (3), for small lags, capture most of the channel dynamics (Komninakis et al., 2002; Li & Wong, 2007). Therefore, following (Komninakis et al., 2002; Li & Wong, 2007), we herein approximate the MIMO channel variations by a first order AR process. Thus, defining vec$(\cdot)$ as the operator that

stacks the columns of a matrix on top of each other, the time evolution of the channel is given by

$$\mathbf{h}_k^{\text{ind}} = \beta \mathbf{h}_{k-1}^{\text{ind}} + \mathbf{w}_k, \tag{4}$$

where

$$\mathbf{h}_k^{\text{ind}} = \text{vec}(\mathbf{H}_k), \tag{5}$$

$\beta = \mathcal{J}_0(2\pi f_D T_s)$, $\mathbf{w}_k$ is a vector of length $N_R N_T$ containing independent samples of circularly symmetric, zero-mean, 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 superscript ind indicates that the channel coefficients are independent and spatially uncorrelated. In general, however, the channel coefficients present a spatial correlation that depends on the propagation environment as well as the polarization of the antenna elements and the spacing between them. Among the spatially correlated channel models, one of the most used splits the fading correlation into two independent components called receive correlation and transmit correlation, respectively. This amounts to modeling the channel as (Larsson & Stoica, 2003)

$$\mathbf{H}_k = \mathbf{R}_R^{1/2} \mathbf{H}_k^{\text{ind}} \left(\mathbf{R}_T^{1/2} \right)^{\mathsf{T}}, \tag{6}$$

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 (Gantmacher, 1959; Golub & Van Loan, 1996) and $\mathbf{H}_k^{\text{ind}}$ is a MIMO channel with independent, uncorrelated and unit variance Gaussian elements.
As the model (6) is calculated from the channel with uncorrelated coefficients, it would be natural to try to formulate the time evolution of spatially correlated channel in a way analog to (4). Therefore, in the next subsection we derive an AR model for time-varying, spatially correlated MIMO channels.

2.1 Time-varying spatially correlated state-space channel model

If we apply the $\text{vec}(\cdot)$ operator to (6), we obtain

$$\text{vec}\left(\mathbf{H}_k\right) = \mathbf{h}_k = \left(\mathbf{R}_T^{1/2} \otimes \mathbf{R}_R^{1/2} \right) \text{vec}\left(\mathbf{H}_k^{\text{ind}} \right) = \mathbf{G} \mathbf{h}_k^{\text{ind}}, \tag{7}$$

where $\mathbf{G} = \left(\mathbf{R}_T^{1/2} \otimes \mathbf{R}_R^{1/2} \right)$, $\otimes$ represents the Kronecker product and we use the fact that, for any matrices $\mathbf{D} = \mathbf{ABC}$ of compatible sizes, $\text{vec}\left(\mathbf{D}\right) = \left(\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A}\right) \text{vec}\left(\mathbf{B}\right)$ (Golub & Van Loan, 1996).
Using the properties of Kronecker products (Golub & Van Loan, 1996; Horn & Johnson, 1991), we can define the matrix $\mathbf{R}_{\mathbf{h}}$ as

$$\begin{aligned} \mathbf{R}_{\mathbf{h}} = \mathbf{G}\mathbf{G}^{\mathsf{H}} &= \left(\mathbf{R}_T^{1/2} \otimes \mathbf{R}_R^{1/2} \right) \left(\mathbf{R}_T^{1/2} \otimes \mathbf{R}_R^{1/2} \right)^{\mathsf{H}} = \left(\mathbf{R}_T^{1/2} \otimes \mathbf{R}_R^{1/2} \right) \left(\mathbf{R}_T^{\mathsf{H}/2} \otimes \mathbf{R}_R^{\mathsf{H}/2} \right) \\ &= \left(\mathbf{R}_T^{1/2} \mathbf{R}_T^{\mathsf{H}/2} \otimes \mathbf{R}_R^{1/2} \mathbf{R}_R^{\mathsf{H}/2} \right) = \left(\mathbf{R}_T \otimes \mathbf{R}_R \right). \end{aligned} \tag{8}$$

Pre-multiplying (4) by $\mathbf{G}$ results in

$$\mathbf{G}\mathbf{h}_k^{\text{ind}} = \beta \mathbf{G}\mathbf{h}_{k-1}^{\text{ind}} + \mathbf{G}\mathbf{w}_k. \tag{9}$$

Comparing (9) to (7), we can describe the dynamics of spatially correlated MIMO channels as

$$\mathbf{h}_k = \beta \mathbf{h}_{k-1} + \mathbf{G}\mathbf{w}_k. \tag{10}$$

This model is similar to (4), but in (10) the excitation noise is correlated.

3. Adaptive channel estimation using Kalman filters

In order to formulate the problem of channel estimation as one of state estimation, we need two equations named process and measurement equations, respectively (Haykin, 2002; Kailath et al., 2000; Simon, 2006). The process equation describes the dynamic behavior of the state variables to be estimated, while the measurement equation presents the relationship between the state variables and the observed system output. As we focus on channel tracking, we can use (10) as the process equation and $\mathbf{h}_k$ as the state vector. The system output, in our case, is the channel output $\mathbf{Y}_k$ in (1). Thus, the measurement equation can be formed by stacking the columns of $\mathbf{Y}_k$, $\mathbf{H}_k$ and $\mathbf{N}_k$ in (1) on top of each other, resulting in

$$\mathbf{y}_k = \boldsymbol{\mathcal{X}}_k \mathbf{h}_k + \mathbf{n}_k, \tag{11}$$

where $\boldsymbol{\mathcal{X}}_k = \mathbf{X}_k^{\mathsf{T}} \otimes \mathbf{I}_{N_R}$ and $\mathbf{R}_n = \sigma_n^2 \mathbf{I}$ is the covariance matrix of the measurement noise $\mathbf{n}_k$. As mentioned in section 2, one of the main properties of any OSTBC is the orthogonality of the codewords. This characteristic is still valid for the transformed codeword $\boldsymbol{\mathcal{X}}_k$, as shown in the following lemma:

Lemma 1 (Orthogonality of $\boldsymbol{\mathcal{X}}_k$)**.** *The matrix $\boldsymbol{\mathcal{X}}_k$ satisfies*

$$\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \boldsymbol{\mathcal{X}}_k = \|\mathbf{x}_k\|^2 \mathbf{I}_{N_R N_T}. \tag{12}$$

Proof. Given that $\mathbf{X}_k$ is an OSTBC and using the properties of Kronecker products (Golub & Van Loan, 1996; Horn & Johnson, 1991), we can write

$$\begin{aligned} \boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \boldsymbol{\mathcal{X}}_k &= \left(\mathbf{X}_k^{\mathsf{T}} \otimes \mathbf{I}_{N_R}\right)^{\mathsf{H}} \left(\mathbf{X}_k^{\mathsf{T}} \otimes \mathbf{I}_{N_R}\right) = \left(\mathbf{X}_k^{*} \otimes \mathbf{I}_{N_R}\right) \left(\mathbf{X}_k^{\mathsf{T}} \otimes \mathbf{I}_{N_R}\right) = \mathbf{X}_k^{*} \mathbf{X}_k^{\mathsf{T}} \otimes \mathbf{I}_{N_R} \\ &= \left(\mathbf{X}_k \mathbf{X}_k^{\mathsf{H}}\right)^{\mathsf{T}} \otimes \mathbf{I}_{N_R} = \|\mathbf{x}_k\|^2 \mathbf{I}_{N_T} \otimes \mathbf{I}_{N_R} = \|\mathbf{x}_k\|^2 \mathbf{I}_{N_R N_T} \end{aligned} \tag{13}$$

where $(\cdot)^*$ is the conjugate of a matrix. □

The state-space formulation of the problem of estimation of flat, time-varying and spatially correlated MIMO channels is then given by (10) and (11). As both (10) and (11) are linear functions of the state vector $\mathbf{h}_k$ and the noises $\mathbf{w}_k$ and $\mathbf{n}_k$ are independent, white and Gaussian, the Kalman filter provides the optimal recursive estimates, in the MMSE sense, for the channel coefficients (Kailath et al., 2000; Simon, 2006). Hence, the classical KF for estimating the MIMO channel is given by

`Prediction`

$$\hat{\mathbf{h}}_{k|k-1} = \beta \hat{\mathbf{h}}_{k-1|k-1} \tag{14a}$$

$$\mathbf{P}_{k|k-1} = \beta^2 \mathbf{P}_{k-1|k-1} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathsf{H}} \tag{14b}$$

`Filtering`

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \left(\boldsymbol{\mathcal{X}}_k \mathbf{P}_{k|k-1} \boldsymbol{\mathcal{X}}_k^{\mathsf{H}} + \mathbf{R}_n\right)^{-1} \tag{15a}$$

$$\boldsymbol{\alpha}_k = \mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{h}}_{k|k-1} \tag{15b}$$

$$\hat{\mathbf{h}}_{k|k} = \hat{\mathbf{h}}_{k|k-1} + \mathbf{K}_k \boldsymbol{\alpha}_k \tag{15c}$$

$$\mathbf{P}_{k|k} = \left[\mathbf{I} - \mathbf{K}_k \boldsymbol{\mathcal{X}}_k\right] \mathbf{P}_{k|k-1} \tag{15d}$$

In these equations, $\mathbf{P}_{k|k-1}$ is the prediction error covariance matrix and $\hat{\mathbf{h}}_{k|k-1}$ is the vector with estimated channel coefficients, at a data block k, computed from the observation of previous blocks.
As done in (Balakumar et al., 2007), the complexity of (14a)–(15d) can be reduced. To that end, as shown in (Loiola et al., 2009), we exploit the orthogonality of OSTBCs as well as the model characteristics. We begin by substituting (8) into (14b)

$$\mathbf{P}_{k|k-1} = \beta^2 \mathbf{P}_{k-1|k-1} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathsf{H}} = \beta^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{G}\mathbf{G}^{\mathsf{H}} = \beta^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{R}_{\mathbf{h}}, \tag{16}$$

where we use the fact that $\mathbf{Q} = \sigma_w^2 \mathbf{I}_{N_R N_T}$.
Remembering that $\mathbf{R}_n = \sigma_n^2 \mathbf{I}_{N_R T}$ and using the matrix inversion lemma and lemma 1, it is possible to rewrite (15a) as

$$\begin{aligned}
\mathbf{K}_k &= \mathbf{P}_{k|k-1}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\left[\mathbf{R}_n^{-1} - \mathbf{R}_n^{-1}\boldsymbol{\mathcal{X}}_k\left(\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{R}_n^{-1}\boldsymbol{\mathcal{X}}_k + \mathbf{P}_{k|k-1}^{-1}\right)^{-1}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{R}_n^{-1}\right] \\
&= \mathbf{P}_{k|k-1}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\left[\frac{1}{\sigma_n^2}\mathbf{I}_{N_R T} - \frac{1}{\sigma_n^4}\boldsymbol{\mathcal{X}}_k\left(\frac{1}{\sigma_n^2}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\boldsymbol{\mathcal{X}}_k + \mathbf{P}_{k|k-1}^{-1}\right)^{-1}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\right] \\
&= \mathbf{P}_{k|k-1}\left[\frac{1}{\sigma_n^2}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} - \frac{1}{\sigma_n^4}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\boldsymbol{\mathcal{X}}_k\left(\frac{\|\mathbf{x}_k\|^2}{\sigma_n^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}^{-1}\right)^{-1}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\right] \\
&= \frac{1}{\sigma_n^2}\mathbf{P}_{k|k-1}\left[\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} - \frac{\|\mathbf{x}_k\|^2}{\sigma_n^2}\left(\frac{\|\mathbf{x}_k\|^2}{\sigma_n^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}^{-1}\right)^{-1}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\right] \\
&= \frac{1}{\sigma_n^2}\mathbf{P}_{k|k-1}\left[\mathbf{I}_{N_R N_T} - \frac{\|\mathbf{x}_k\|^2}{\sigma_n^2}\left(\frac{\|\mathbf{x}_k\|^2}{\sigma_n^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}^{-1}\right)^{-1}\right]\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}.
\end{aligned} \tag{17}$$

Employing the matrix inversion lemma once more, it is possible to write the inverse matrix of the last expression of (17) as

$$\begin{aligned}
&\left(\frac{\|\mathbf{x}_k\|^2}{\sigma_n^2}\mathbf{I}_{N_R N_T} + \mathbf{I}_{N_R N_T}\mathbf{P}_{k|k-1}^{-1}\mathbf{I}_{N_R N_T}\right)^{-1} = \\
&\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T} - \frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}\right)^{-1}\left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T}\right) = \\
&\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T} - \left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\right)^2\left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}\right)^{-1}.
\end{aligned} \tag{18}$$

Substituting now (18) into (17), the Kalman gain adaptation equation is expressed as

$$\begin{aligned}
\mathbf{K}_k &= \frac{1}{\sigma_n^2}\mathbf{P}_{k|k-1}\left[\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}\right)^{-1}\right]\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \\
&= \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{P}_{k|k-1}\left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}\right)^{-1}\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} = \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}},
\end{aligned} \tag{19}$$

where we define

$$\mathbf{A}_k = \mathbf{P}_{k|k-1}\left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1}\right)^{-1}. \tag{20}$$

Using lemma 1, (14a), (15b), and (19) into (15c), we have

$$\begin{aligned}\hat{\mathbf{h}}_{k|k} &= \hat{\mathbf{h}}_{k|k-1} + \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{h}}_{k|k-1}\right)\\ &= \hat{\mathbf{h}}_{k|k-1} - \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\boldsymbol{\mathcal{X}}_k\hat{\mathbf{h}}_{k|k-1} + \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{y}_k\\ &= \hat{\mathbf{h}}_{k|k-1} - \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\left(\|\mathbf{x}_k\|^2\mathbf{I}_{N_RN_T}\right)\hat{\mathbf{h}}_{k|k-1} + \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{y}_k\\ &= \left(\mathbf{I}_{N_RN_T} - \mathbf{A}_k\right)\hat{\mathbf{h}}_{k|k-1} + \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{y}_k = \left(\mathbf{I}_{N_RN_T} - \mathbf{A}_k\right)\left(\beta\hat{\mathbf{h}}_{k-1|k-1}\right) + \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{y}_k\\ &= \beta\mathbf{B}_k\hat{\mathbf{h}}_{k-1|k-1} + \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{y}_k,\end{aligned} \tag{21}$$

where $\mathbf{B}_k$ is defined as

$$\mathbf{B}_k = \mathbf{I}_{N_RN_T} - \mathbf{A}_k. \tag{22}$$

Finally, using (19) into (15d) we obtain

$$\mathbf{P}_{k|k} = \left(\mathbf{I}_{N_RN_T} - \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\boldsymbol{\mathcal{X}}_k\right)\mathbf{P}_{k|k-1} = \left(\mathbf{I}_{N_RN_T} - \mathbf{A}_k\right)\mathbf{P}_{k|k-1} = \mathbf{B}_k\mathbf{P}_{k|k-1}. \tag{23}$$

Putting together (16), (20), (21), (22) and (23), the reduced complexity Kalman channel estimator (KCE) for correlated MIMO-OSTBC systems is given by (Loiola et al., 2009)

$$\mathbf{P}_{k|k-1} = \beta^2\mathbf{P}_{k-1|k-1} + \sigma_w^2\mathbf{R}_{\mathbf{h}} \tag{24a}$$

$$\mathbf{A}_k = \mathbf{P}_{k|k-1}\left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2}\mathbf{I}_{N_RN_T} + \mathbf{P}_{k|k-1}\right)^{-1} \tag{24b}$$

$$\mathbf{B}_k = \mathbf{I}_{N_RN_T} - \mathbf{A}_k \tag{24c}$$

$$\hat{\mathbf{h}}_{k|k} = \beta\mathbf{B}_k\hat{\mathbf{h}}_{k-1|k-1} + \frac{1}{\|\mathbf{x}_k\|^2}\mathbf{A}_k\boldsymbol{\mathcal{X}}_k^{\mathsf{H}}\mathbf{y}_k \tag{24d}$$

$$\mathbf{P}_{k|k} = \mathbf{B}_k\mathbf{P}_{k|k-1} \tag{24e}$$

It is important to note that one of the key assumptions to the complexity reduction in (Balakumar et al., 2007) is the uncorrelated nature of the channel coefficients. In this case, and supposing that the initial value $\mathbf{P}_{0|0}$ is also a diagonal matrix, it is shown in (Balakumar et al., 2007) that $\mathbf{P}_{k|k-1}$ is always diagonal, which simplifies all subsequent calculations. However, for a general spatial correlation matrix $\mathbf{R}_{\mathbf{h}}$, it is not possible to simplify the computation of the matrix inversion in (24b). For this reason, the approach taken in (Loiola et al., 2009) to reduce the complexity of KCE (24a)–(24e) is the development of a steady-state Kalman channel estimator, which is presented in section 4. It will be shown in section 4 that the steady-state Kalman channel estimator has a complexity order less than or equal to that of the algorithm in (Balakumar et al., 2007) and works also for non-diagonal spatial correlation matrices.

It is also worth observing that the channel estimates produced by the Kalman filter (24a)–(24e) correspond to weighted sums of instantaneous ML channel estimates. To see this, first

consider the instantaneous ML channel estimates, i.e., the estimates computed by using only the k^{th} data block, which is given by (Kaiser et al., 2005)

$$\hat{\mathbf{h}}_k^{(\text{ML})} = \left(\boldsymbol{\mathcal{X}}_k^{\text{H}} \boldsymbol{\mathcal{X}}_k\right)^{-1} \boldsymbol{\mathcal{X}}_k^{\text{H}} \mathbf{y}_k. \tag{25}$$

For OSTBCs, thanks to lemma 1, (25) reduces to

$$\hat{\mathbf{h}}_k^{(\text{ML})} = \left(\|\mathbf{x}_k\|^2 \mathbf{I}_{N_R N_T}\right)^{-1} \boldsymbol{\mathcal{X}}_k^{\text{H}} \mathbf{y}_k = \frac{1}{\|\mathbf{x}_k\|^2} \boldsymbol{\mathcal{X}}_k^{\text{H}} \mathbf{y}_k. \tag{26}$$

Thus, using (26) the channel estimate (24d) can be rewritten as

$$\hat{\mathbf{h}}_{k|k} = \beta \mathbf{B}_k \hat{\mathbf{h}}_{k-1|k-1} + \mathbf{A}_k \hat{\mathbf{h}}_k^{(\text{ML})}. \tag{27}$$

Consequently, the KF proposed in (Loiola et al., 2009) updates the channel estimates through weighted sums of instantaneous maximum likelihood channel estimates. It is important to note that the weights are time-varying and optimally calculated, in the MMSE sense, for each data block.

Considering communication systems where pilot sequences are periodically inserted between information symbols, the algorithm in (24a)–(24e) can operate in both training and decision-directed (DD) modes. First, when pilot symbols are available, the matrix $\boldsymbol{\mathcal{X}}_k$ in (24d) is constructed from them. Once the transmission of pilot symbols is finished, the algorithm enters in decision-directed mode and the matrix $\boldsymbol{\mathcal{X}}_k$ is then formed by the decisions provided by the ML space-time decoder. Note that these decisions are based on the channel estimates generated by the algorithm in the previous iteration.

4. Steady-state Kalman channel estimator

The measurement equation (11) represents a time-varying system, since the matrix $\boldsymbol{\mathcal{X}}_k$ changes at each transmitted data block. However, in the Kalman channel estimator (24a)–(24e), only (24d) has an explicit dependence on $\boldsymbol{\mathcal{X}}_k$. Because of the orthogonality of OSTBC codewords, all other expressions in this recursive estimator depend only on the energy of the uncoded data block, i.e. $\|\mathbf{x}_k\|^2$. Now, for constant modulus signal constellations such as M-PSK, $\|\mathbf{x}_k\|^2$ is a constant. In this case, (24a)–(24c) and (24e) are just functions of the initial estimate of $\mathbf{P}_{k|k}$, the normalized Doppler rate, the spatial correlation matrix, a constant equal to the energy of the constellation symbols and the variance of the measurement noise.

These parameters can be estimated ahead of time using, for example, the methods proposed in (Jamoos et al., 2007) and in the references therein. Thus, we assume that the parameters in (24a)–(24c) and (24e) are known. Furthermore, we can analyze the state-space model (10) and (11) to check if the matrices $\mathbf{P}_{k|k}$, $\mathbf{A}_k$ and $\mathbf{B}_k$ converge to steady-state values. If this is the case, and if these values can be found, the time-varying matrices could be replaced by constant matrices, originating a low complexity sub-optimal estimator known as the steady-state Kalman channel estimator (SS-KCE) (Loiola et al., 2009). As pointed out in (Simon, 2006), the steady-state filter often performs nearly as well as the optimal time-varying filter.

To determine the SS-KCE, we begin by substituting (24e) into (24a), which yields

$$\mathbf{P}_{k|k-1} = \beta^2 \mathbf{B}_{k-1} \mathbf{P}_{k-1|k-2} + \sigma_w^2 \mathbf{R}_{\mathbf{h}}. \tag{28}$$

Now substitute (24c) into (28) to obtain

$$\mathbf{P}_{k|k-1} = \beta^2 \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_{k-1}\right) \mathbf{P}_{k-1|k-2} + \sigma_w^2 \mathbf{R}_{\mathbf{h}}. \tag{29}$$

Taking into account (24b), we can rewrite (29) as

$$\mathbf{P}_{k|k-1} = \beta^2 \mathbf{P}_{k-1|k-2} - \beta^2 \mathbf{P}_{k-1|k-2} \left(\frac{\sigma_n^2}{n_s} \mathbf{I}_{N_R N_T} + \mathbf{P}_{k-1|k-2}\right)^{-1} \mathbf{P}_{k-1|k-2} + \sigma_w^2 \mathbf{R}_{\mathbf{h}}, \tag{30}$$

where $n_s = \|\mathbf{x}\|^2$ corresponds to the energy of each uncoded data block $\mathbf{x}_k$, assumed to be a constant.
If $\mathbf{P}_{k|k-1}$ converges to a steady-state value, then $\mathbf{P}_{k|k-1} = \mathbf{P}_{k-1|k-2}$ for large k. Denoting this steady-state value as $\mathbf{P}_\infty$, we rewrite (30) as

$$\mathbf{P}_\infty = \beta^2 \mathbf{P}_\infty - \beta^2 \mathbf{P}_\infty \left(\mathbf{P}_\infty + \frac{\sigma_n^2}{n_s} \mathbf{I}_{N_R N_T}\right)^{-1} \mathbf{P}_\infty + \sigma_w^2 \mathbf{R}_{\mathbf{h}}. \tag{31}$$

Equation (31) is a discrete algebraic Riccati equation (DARE) (Kailath et al., 2000; Simon, 2006). If it can be solved, we can use $\mathbf{P}_\infty$ in (24b) and (24c) to calculate the steady-state values of matrices $\mathbf{A}$ and $\mathbf{B}$, denoted $\mathbf{A}_\infty$ and $\mathbf{B}_\infty$, respectively. Hence, the steady-state Kalman channel estimator proposed in (Loiola et al., 2009) is given simply by

$$\hat{\mathbf{h}}_{k|k} = \beta \mathbf{B}_\infty \hat{\mathbf{h}}_{k-1|k-1} + \frac{1}{n_s} \mathbf{A}_\infty \boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \mathbf{y}_k. \tag{32}$$

As in (27), the steady-state KF generates channel estimates by averaging instantaneous ML channel estimates. However, as opposed to (27), the weights in (32) are not time-varying.
The problem now is to determine the solution of (31). As the DARE is highly nonlinear, its solutions $\mathbf{P}_\infty$ may or may not exist, they may or may not be unique or indeed they may or may not generate a stable steady-state filter. In the next subsection, we present the solution to (31), and discuss the stability of the resulting filter (32).

4.1 Existence of DARE solutions

To show one possible solution of the DARE in (31), let $\mathbf{R}_{\mathbf{h}} = \mathbf{Q}_U^{\mathsf{H}} \mathbf{\Lambda} \mathbf{Q}_U$ be the eigendecomposition of $\mathbf{R}_{\mathbf{h}}$. Since $\mathbf{Q}_U$ is unitary, it is easy to verify that $\mathbf{P}_\infty = \mathbf{Q}_U^{\mathsf{H}} \mathbf{\Sigma} \mathbf{Q}_U$ is a solution of the DARE, as long as the diagonal matriz $\mathbf{\Sigma}$ satisfies

$$\mathbf{\Sigma} = \beta^2 \mathbf{\Sigma} - \beta^2 \mathbf{\Sigma} \left(\mathbf{\Sigma} + \frac{\sigma_n^2}{n_s} \mathbf{I}_{N_R N_T}\right)^{-1} \mathbf{\Sigma} + \sigma_w^2 \mathbf{\Lambda}. \tag{33}$$

Now let σ_i and λ_i be the i-th diagonal element of $\mathbf{\Sigma}$ and $\mathbf{\Lambda}$, respectively. Then, since all the matrices in (33) are diagonal, σ_i must satisfy

$$\sigma_i^2 + b\sigma_i + c = 0, \tag{34}$$

where $b = \sigma_n^2(1-\beta^2)/n_s - \sigma_w^2 \lambda_i$ and $c = -\sigma_n^2 \sigma_w^2 \lambda_i / n_s$.
Equation (34) has two possible solutions. We now show that only one of these solutions is valid, in the sense that the resulting $\mathbf{P}_\infty$ is a valid autocorrelation matrix. To that end, we need to show that the eigenvalues of $\mathbf{P}_\infty$ are real and non-negative. We begin by noting that $\mathbf{R}_{\mathbf{h}}$ is a correlation matrix, so $\lambda_i \geq 0$. As the remaining terms of c also are positive, we conclude

that $c \leq 0$. Thus, the discriminant of (34), given by $b^2 - 4c$, is non-negative. We identify two possibilities. First, the discriminant is zero if and only if $b = c = 0$. This happens if and only if there is no mobility, in which case $\beta = 1$ and $\sigma_w^2 = 0$. In this case, $\sigma_i = 0$, so $\mathbf{P}_\infty$ does not have full rank. On the other hand, if there is mobility, the discriminant of (34) is strictly positive. In this case, the quadratic equation in (34) has two distinct real solutions. Furthermore, since $c \leq 0$, we have that $b^2 - 4c \geq b^2$, so the solution given by $(-b + \sqrt{b^2 - 4c})/2$ is non-negative, which concludes the proof.

We also need to prove that the SS-KCE in (32) is stable. To that end, note that stability holds as long as the eigenvalues of $\mathbf{I} - \mathbf{A}_\infty$ have magnitude less than one. Now, using the fact that $\mathbf{P}_\infty = \mathbf{Q}_U^{\mathsf{H}} \mathbf{\Sigma} \mathbf{Q}_U$, it is easy to verify that the eigenvalues of $\mathbf{I} - \mathbf{A}_\infty$, ρ_i, are given by

$$\rho_i = \frac{\sigma_n^2/n_s}{\sigma_n^2/n_s + \sigma_i}. \tag{35}$$

Note that $\sigma_i \geq 0$, so that $0 < \rho_i \leq 1$. Also, note that $\rho_i = 1$ if and only if $\sigma_i = 0$, which happens if and only if $\lambda_i = 0$, i.e., when the spatial correlation matrix $\mathbf{R_h}$ does not have full rank. In this case, the SS-KCE is marginally stable. In all other cases, the filter is stable.

Finally, we note that the SS-KCE does not work very well in low mobility. In fact, we will show that, as $\beta \to 1$, the SS-KCE in (32) tends to $\hat{\mathbf{h}}_{k|k} = \hat{\mathbf{h}}_{k-1|k-1}$. In other words, as $\beta \to 1$, the SS-KCE does not update the channel estimate, simply keeping the initial guess for all iterations while ignoring the channel output. This makes intuitive sense. Indeed, as $\beta \to 1$, the state equation (10) tends to $\mathbf{h}_k = \mathbf{h}_{k-1}$, i.e., the channel becomes static. In this case, as we have more and more observations, the variance of the estimation error in the Kalman filter tends to zero. Thus, in steady-state, the filter stops updating the channel estimates. To prove this result in our case, we note that, as $\beta \to 1$, $\sigma_w^2 \to 0$, so the solution of (34) tends to $\sigma_i = 0$. Using again the fact that $\mathbf{P}_\infty = \mathbf{Q}_U^{\mathsf{H}} \mathbf{\Sigma} \mathbf{Q}_U$, we see that the eigenvalues of $\mathbf{A}_\infty$ are given by $\sigma_i/(\sigma_n^2/n_s + \sigma_i)$. Thus, as $\beta \to 1$, these eigenvalues tend to zero, so that $\mathbf{A}_\infty \to \mathbf{0}$, and the result follows.

5. Fading-memory Kalman channel estimator

As mentioned in Section 2, the first order AR model used in (10) is only an approximate description of the time evolution of channel coefficients. This modeling error can degrade the performance of Kalman-based 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 (Anderson & Moore, 1979; Simon, 2006). To understand how this can be done, we consider the state-space model (10) and (11). For this model, it is possible to show (Anderson & Moore, 1979; Simon, 2006) that the sequence of estimates produced by the KCE minimizes $\mathrm{E}[J_N]$, where the cost function J_N is given by

$$J_N = \sum_{k=1}^{N} \left[\left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{h}}_{k|k-1} \right)^{\mathsf{H}} \mathbf{R}_n^{-1} \left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{h}}_{k|k-1} \right) + \mathbf{w}_k^{\mathsf{H}} \left(\sigma_w^2 \mathbf{R_h} \right)^{-1} \mathbf{w}_k \right]. \tag{36}$$

The importance of the most recent observations can be increased if they receive a higher weight than past data. This can be accomplished with an exponential weight, controlled by a scalar $\alpha \geq 1$. In this case, the cost function can be rewritten as (Anderson & Moore, 1979;

Simon, 2006)

$$\tilde{J}_N = \sum_{k=1}^{N} \left[\left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{h}}_{k|k-1} \right)^{\mathsf{H}} \alpha^{2k} \mathbf{R}_n^{-1} \left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{h}}_{k|k-1} \right) + \mathbf{w}_k^{\mathsf{H}} \alpha^{2k+2} \left(\sigma_w^2 \mathbf{R}_{\mathbf{h}} \right)^{-1} \mathbf{w}_k \right]. \quad (37)$$

Following (Anderson & Moore, 1979; Simon, 2006), it is possible to show that the minimization of $\mathrm{E}[\tilde{J}_N]$ for OSTBC systems leads to the fading-memory Kalman channel estimator (FM-KCE), given by

$$\mathbf{P}_{k|k-1} = (\alpha\beta)^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{R}_{\mathbf{h}} \quad (38a)$$

$$\mathbf{A}_k = \mathbf{P}_{k|k-1} \left(\frac{\sigma_n^2}{\|\mathbf{x}_k\|^2} \mathbf{I}_{N_R N_T} + \mathbf{P}_{k|k-1} \right)^{-1} \quad (38b)$$

$$\hat{\mathbf{h}}_{k|k} = \beta \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_k \right) \hat{\mathbf{h}}_{k-1|k-1} + \mathbf{A}_k \frac{\boldsymbol{\mathcal{X}}_k^{\mathsf{H}} \mathbf{y}_k}{\|\mathbf{s}_k\|^2} \quad (38c)$$

$$\mathbf{P}_{k|k} = \left(\mathbf{I}_{N_R N_T} - \mathbf{A}_k \right) \mathbf{P}_{k|k-1} \quad (38d)$$

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 (38a). 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. 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.

As an aside, we note that the FM-KCE can be interpreted as a result of adding a fictitious process noise (Anderson & Moore, 1979; Simon, 2006), which in consequence reduces the confidence of the KCE in the system model and increases the importance of observed data. To see that this fictitious process noise addition is mathematically equivalent to the FM-KCE, we rewrite (38a) as

$$\begin{aligned} \mathbf{P}_{k|k-1} &= (\alpha\beta)^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{R}_{\mathbf{h}} = \left(\alpha^2 - 1 + 1 \right) \beta^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{R}_{\mathbf{h}} \\ &= \beta^2 \mathbf{P}_{k-1|k-1} + \sigma_w^2 \mathbf{R}_{\mathbf{h}} + \left(\alpha^2 - 1 \right) \beta^2 \mathbf{P}_{k-1|k-1} = \beta^2 \mathbf{P}_{k-1|k-1} + \tilde{\mathbf{Q}}, \end{aligned} \quad (39)$$

where

$$\tilde{\mathbf{Q}} = \sigma_w^2 \mathbf{R}_{\mathbf{h}} + \left(\alpha^2 - 1 \right) \beta^2 \mathbf{P}_{k-1|k-1} \quad (40)$$

and $\left(\alpha^2 - 1 \right) \beta^2 \mathbf{P}_{k-1|k-1}$ corresponds to the covariance matrix of the fictitious process noise.

Due to the similarity between the KCE (24a)–(24d) and the FM-KCE (38a)–(38d), one could think that the FM-KCE should also have a steady-state version. Following the same steps described in Section 4 to the derivation of (31), it is not hard to show that the Riccati equation for the FM-KCE is given by

$$\mathbf{P}_\infty = (\alpha\beta)^2 \mathbf{P}_\infty - (\alpha\beta)^2 \mathbf{P}_\infty \left(\mathbf{P}_\infty + \frac{\sigma_n^2}{n_s} \mathbf{I}_{N_R N_T} \right)^{-1} \mathbf{P}_\infty + \sigma_w^2 \mathbf{R}_{\mathbf{h}}. \quad (41)$$

Its solution is also of the form $\mathbf{P}_\infty = \mathbf{Q}_U^{\mathsf{H}} \boldsymbol{\Sigma} \mathbf{Q}_U$. The elements of the diagonal matrix $\boldsymbol{\Sigma}$ are given by $\sigma_i = -b + \sqrt{b^2 - 4c}$, where $b = \sigma_n^2 (1 - \alpha^2\beta^2)/n_s - \sigma_w^2 \lambda_i$ and $c = -\sigma_n^2 \sigma_w^2 \lambda_i / n_s$. Since $c \geq 0$,

we conclude that $\sigma_i \geq 0$, so the solution leads to a valid autocorrelation matrix, as before. Also, as before, we see that the steady-state filter is stable as long as $\sigma_i > 0$. Now, $\sigma_i = 0$ if and only if $c = 0$, which happens if $\lambda_i = 0$, i.e., $\mathbf{R_h}$ does not have full rank, or if $\sigma_w^2 = 0$, i.e., if there is no mobility. In either of these cases, the steady-state filter is marginally stable. Otherwise, the filter is stable.

Finally, we note that the DARE (41) could also be derived from the process equation

$$\mathbf{h}_k = \alpha\beta\mathbf{h}_{k-1} + \mathbf{G}\mathbf{w}_k. \tag{42}$$

Comparing (10) to (42), we see that the state transition matrix in (42) is modified by the scalar $\alpha \geq 1$, while the variance of the process noise remains the same. As shown in (Simon, 2006), this could be interpreted as an artificial increase in the process noise variance and hence equivalent to that done in (40).

6. Simulation results

In this section, we present some simulation results to illustrate the performance of the presented channel estimation algorithms. In all simulations the correlated channels are generated by (7), where the elements of $\mathbf{h}_k^{\text{ind}}$ are Rayleigh distributed with time autocorrelation function given by (3). It is worth emphasizing that the estimators presented in this chapter approximate the channel dynamics by the first order AR model (10). The receiver operates in decision-directed mode, i.e. after a certain number of space-time training codewords, the channel estimators employ the decisions provided by the ML space-time decoder. Unless stated otherwise, we insert 25 OSTBC training codewords between every 225 OSTBC data codewords.

Supposing that the spatial correlation coefficient between any two adjacent receive (transmit) antennas is given by p_r (p_t), it is possible to express each (i, j) element of the spatial correlation matrices R_R and R_T as $p_r^{|i-j|}$, $i, j = 1, \ldots, N_R$ and $p_t^{|i-j|}$, $i, j = 1, \ldots, N_T$, respectively. 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 $f_D T_s$. The simulation results presented in the sequel correspond to averages of 10 channel realization, in each of which we simulate the transmission of 1×10^6 orthogonal space-time codewords. For comparison purposes, we also simulate a channel estimator implemented by the well known RLS adaptive filter (Haykin, 2002), with a forgetting factor of 0.98. This value was determined by trial and error to yield the best performance of the RLS.

To verify if there is any performance degradation of the SS-KCE (32) compared to the KCE (24a)–(24e), we simulate the transmission of 8-PSK symbols from $N_T = 2$ transmit antennas to $N_R = 2$ receive antennas using the Alamouti space-time block code (Alamouti, 1998). We also assume $p_t = 0.4$, $p_r = 0$ and different normalized Doppler rates. Fig. 1 shows the estimation mean squared error (MSE) for KCE and SS-KCE as a function of $f_D T_s$. We observe that the smaller the value of $f_D T_s$ (i.e. the smaller the relative velocity between transmitter and receiver), the greater the gap between KCE and SS- KCE. In the limit when $f_D T_s = 0$, the channel is time-invariant, the solution of (31) is null and the SS-KCE does not update the channel estimates. On the other hand, for channels varying at typical rates, both algorithms have equivalent performances. This can be seen in Fig. 2, which presents the symbol error rates at the output of ML space-time decoders fed with channel state information (CSI) provided by KCE and SS-KCE, as well as at the output of an ML decoder with perfect channel knowledge. Clearly, SS-KCE has the same performance of the KCE for the two values of $f_D T_s$ considered while demanding just a fraction of the complexity.

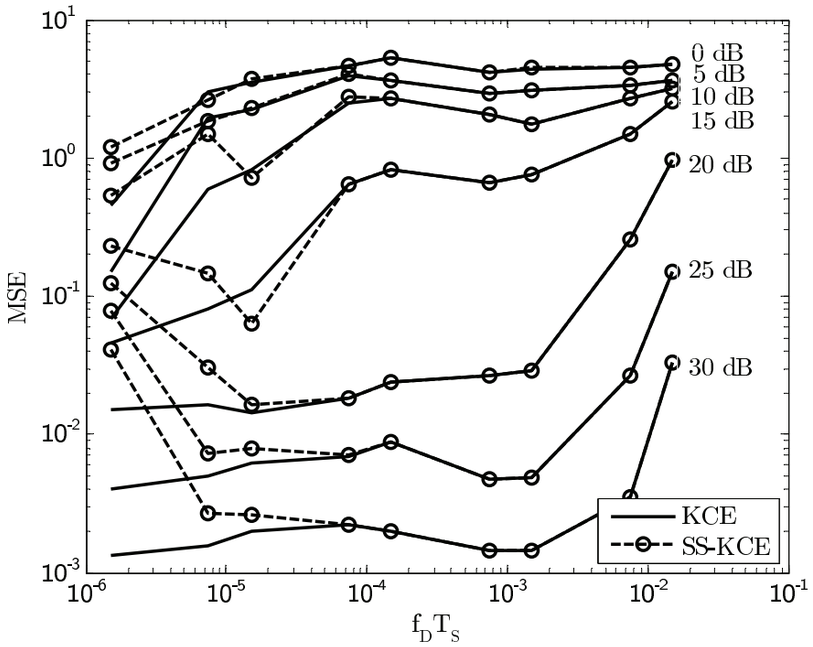

Fig. 1. Estimation mean squared error for KCE and SS-KCE.

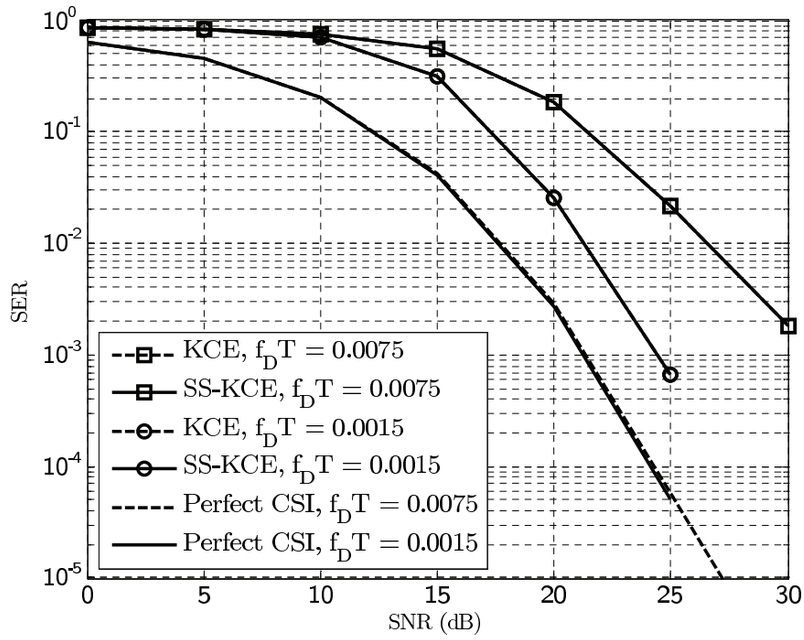

Fig. 2. Symbol error rates of ML decoders fed with channel estimates provided by KCE and SS-KCE.

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

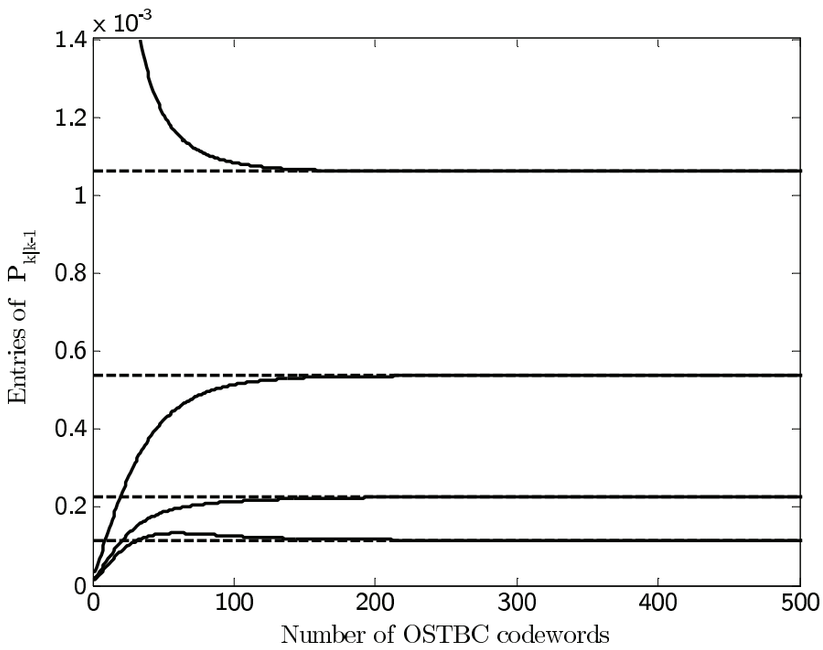

Fig. 3. Evolution of the entries of $\mathbf{P}_{k|k-1}$.

values of $\mathbf{A}_k$ and $\mathbf{B}_k$ after just a few blocks. Consequently, after these few blocks, the estimates provided by the SS-KCE are the same as those generated by the optimal 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 $\mathbf{P}_{0|0} = \mathbf{I}_{N_R N_T}$. 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 even before the second training period. Due to the similar performances of KCE and SS-KCE, we hereinafter present just SS-KCE results.

It is important to observe that the gap in the symbol error rate curves of Fig. 2, between the decoders with perfect CSI and with estimated CSI, is due in great part to the use of the first order AR approximation to the channel dynamics. To show this, in Fig. 4 we present the symbol error rates at the output of decoders with perfect CSI and with SS-KCE estimates for the same scenario used in Fig 2, except that in Fig. 4 the channel is also generated by a first order AR process. As we can see, for $f_D T_s = 0.0015$, the receiver composed by SS-KCE and the space-time decoder has the same performance as the ML decoder with perfect CSI. For $f_D T_s = 0.0075$ and an SER of 10^{-3}, the receiver using SS-KCE is about 5 dB from the decoder with perfect CSI. This value is half of that shown in Fig. 2.

To analyze the impact of spatial channel correlation in the performance of the channel estimation algorithms, the next scenario simulates the transmission of QPSK symbols to 2 receive antennas using Alamouti's code for a normalized Doppler rate of 0.0045. The receiver correlation coefficient p_r is set to zero while the transmitter correlation coefficient p_t assumes values of 0.2 and 0.8. Fig. 5 presents the channel estimation MSE for SS-KCE and RLS algorithms for both p_t considered. From this figure, we note that the performances of the estimation algorithms are hardly affected by transmitter spatial correlation and that the

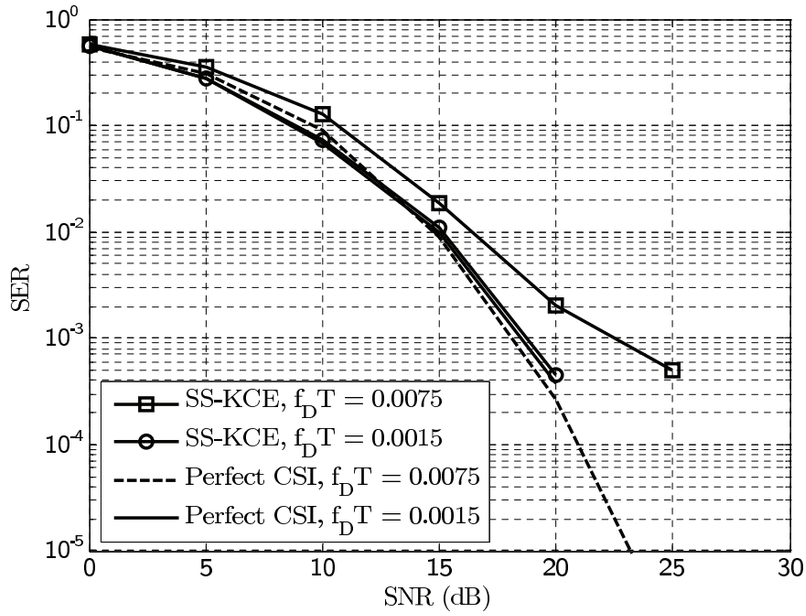

Fig. 4. Symbol error rates of ML space-time decoders for a first order AR channel.

curves for RLS are indistinguishable. It is also clear that the SS-KCE performs much better than the classical RLS algorithm. The symbol error rates at the output of ML decoders using the channel estimates provided by SS-KCE and RLS filters are shown in Fig. 6. Since the simulated RLS adaptive filter is not able to track the channel variations, the decoder can not correctly decode the space-time codewords, leading to a poor receiver performance. On the other hand, the receiver fed with SS-KCE estimates is 3 dB from the decoder with perfect CSI for both values of p_t at an SER of 10^{-4}.

In the previous simulations, the channel estimators tracked simultaneously the 4 possible channels between 2 transmit and 2 receive antennas. If the number of antennas increases, the number of channels to be tracked simultaneously also increases. To illustrate the capacity of the KF-based algorithms to track a larger number of channels, we simulate a system sending QPSK symbols from $N_T = 4$ transmit to $N_R = 4$ receive antennas. We employ the $1/2$-rate OSTBC of (Tarokh et al., 1999) and assume $p_t = 0.8$ and $p_r = 0.4$. The MSE for the RLS and the SS-KCE is shown in Fig. 7. We observe that the estimates produced by the RLS algorithm are affected by the rate of channel variation. Moreover, the RLS MSE flattens out for SNR's greater than 10 dB. On the other hand, for this scenario, the SS-KCE has the same performance for both values of $f_D T_s$ considered and the MSE presents a linear decrease 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 at the output of the ML decoders, as shown in Fig. 8. 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 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.

To cope with the modeling error introduced by the use of the first-order AR channel model, we show the FM-KCE in Section 5. Hence, to illustrate the performance improvement of FM-KCE

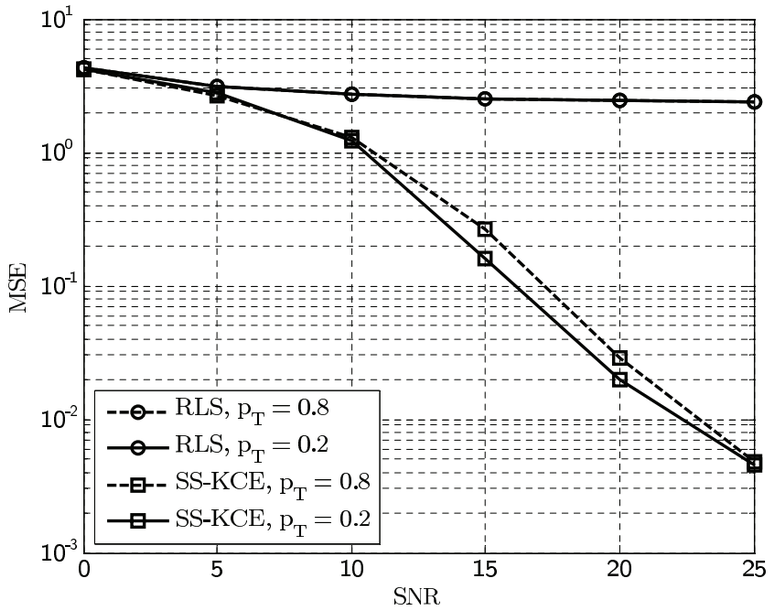

Fig. 5. Estimation mean square error for different transmitter correlation coefficient.

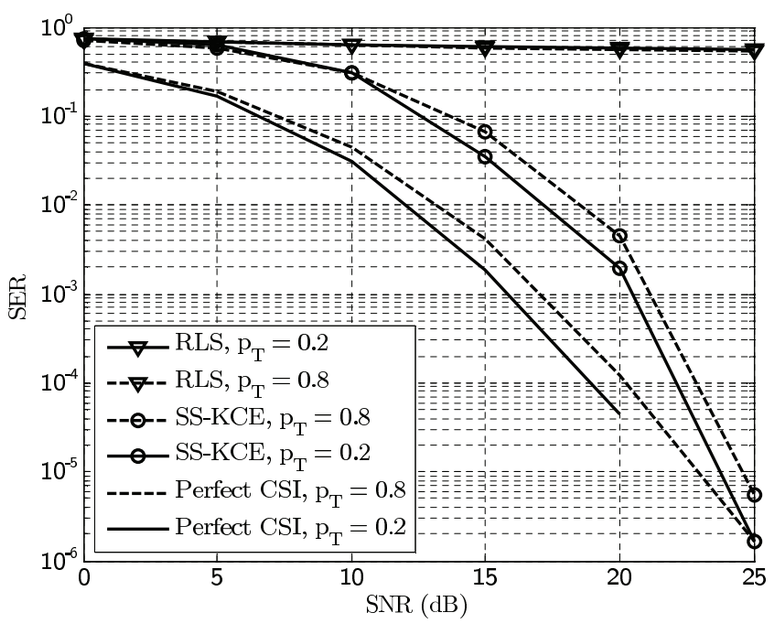

Fig. 6. Symbol error rate for different transmitter correlation coefficient.

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 $p_t = 0.4$. We vary the number of training codewords from 4

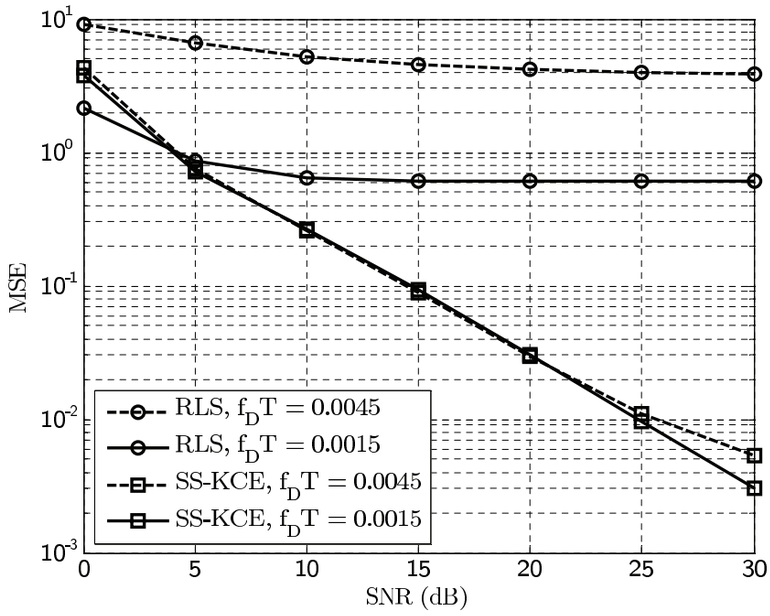

Fig. 7. Estimation mean square error for different values of $f_D T$.

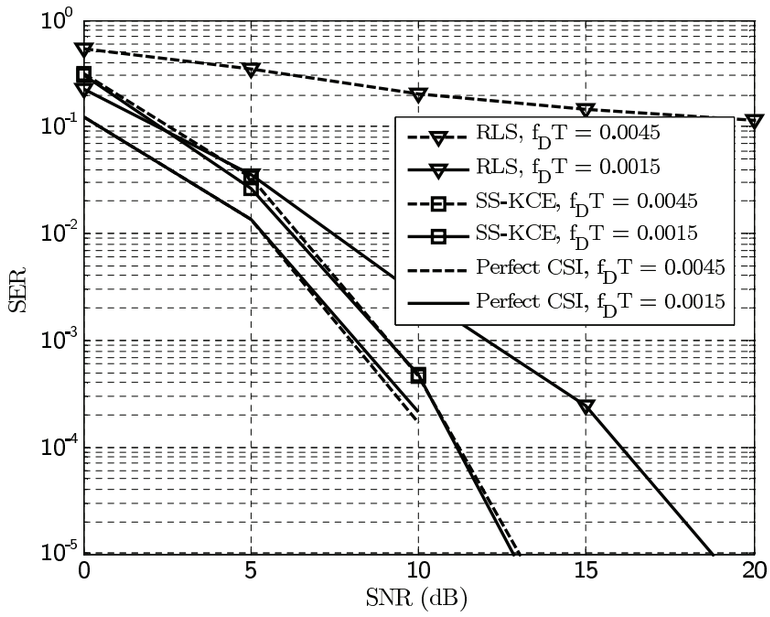

Fig. 8. Symbol error rate for different values of $f_D T$.

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. 9 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 (41), with $4, 8, 12, 16, 20, 24, 28$ and 32 training codewords. The arrows in this figure indicate the number of training

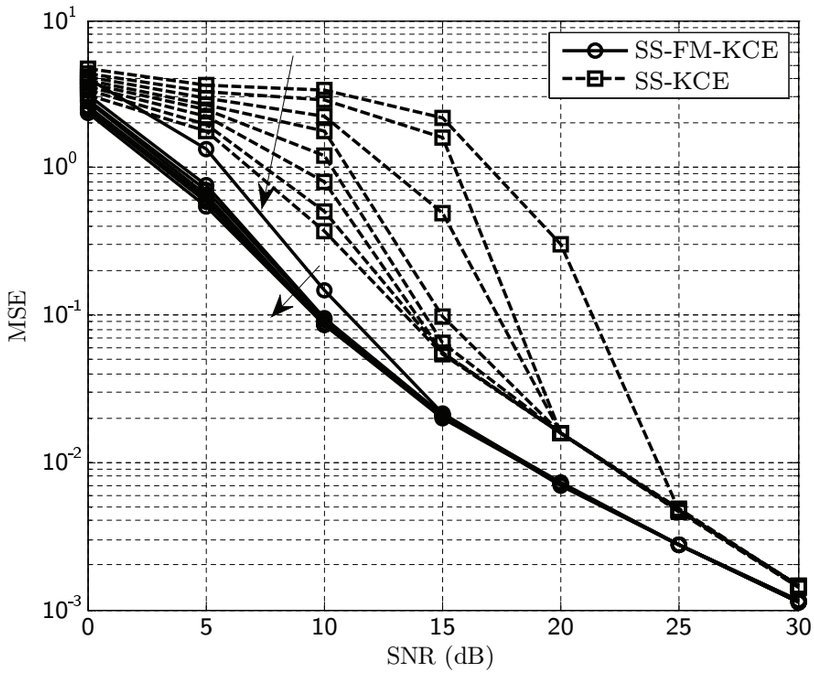

Fig. 9. Estimation mean square error for SS-KCE and FM-KCE.

codewords in ascending order. From Fig. 9, it is evident the superiority of FM-KCE over SS-KCE. 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 5 dB better than the SS-KCE with 4 training codewords and about 3.5 dB better than the SS-KCE with 32 traininig codeowrds.

The superior performance of the FM-KCE can also be observed in Fig. 10, 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 0.8 dB from the decoder with perfect CSI, while the receiver using channel estimates provided by the SS-KCE presents performance losses of 3 and 5.5 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 2 and 3.5 dB better than the receiver with SS-KCE for 32 and 4 training codewords, respectively, and presents a loss of 0.5 dB from the ML space-time decoder with perfect CSI. Thus, from Figs. 9 and 10, we see that the FM-KCE allows the use of a small number of training codewords without compromising the performance of the receiver.

7. Summary

In this chapter, we presented channel estimation algorithms intended for systems employing orthogonal space-time block codes. Before developing the channel estimators, we construct a state-space model to describe the dynamic behavior of spatially correlated MIMO channels. Using this channel model, we formulate the problem of channel estimation as one of state

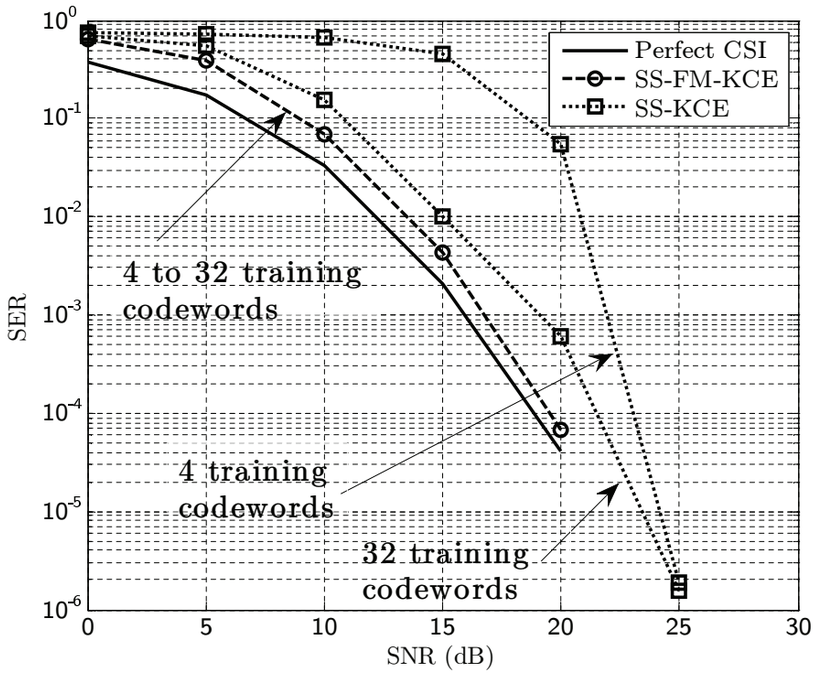

Fig. 10. Symbol error rate for SS-KCE and FM-KCE.

estimation. Thus, by applying the well-known Kalman filter to that state-space model, and using the orthogonality of OSTBCs, we arrive at a low-complexity optimal Kalman channel estimator. We also show that the channel estimates provided by the KCE in fact correspond to weighted sums of instantaneous maximum likelihood channel estimates.

For constant modulus signal constellations, a reduced complexity estimator is give by the steady-state Kalman filter. This filter also generates channel estimates by averaging instantaneous ML channel estimates. The existence and stability of the steady-state Kalman channel estimator is intimately related to the existence of solutions to the discrete algebraic Riccati equation derived from the KCE.

Simulation results indicate that the SS-KCE performs nearly as well as the optimal KCE, while demanding just a fraction of the calculations. They also show that the fading memory estimator outperforms the traditional Kalman filter by as much as 5 dB for a symbol error rate of 10^{-3}.

8. Acknowledgments

We acknowledge the financial support received from CAPES.

9. References

Alamouti, S. M. (1998). A Simple Transmit Diversity Technique for Wireless Communications, *IEEE Journal on Selected Areas in Communications* 16(10): 1451–1458.

Anderson, B. D. O. & Moore, J. B. (1979). *Optimal Filtering*, Prentice-Hall.

Balakumar, B., Shahbazpanahi, S. & Kirubarajan, T. (2007). Joint MIMO Channel Tracking and Symbol Decoding Using Kalman Filtering, *IEEE Transactions on Signal Processing* 55(12): 5873–5879.

Duman, T. M. & Ghrayeb, A. (2007). *Coding for MIMO Communication Systems*, John Wiley and Sons.

Enescu, M., Roman, T. & Koivunen, V. (2007). State-Space Approach to Spatially Correlated MIMO OFDM Channel Estimation, *Signal Processing* 87(9): 2272–2279.

Gantmacher, F. R. (1959). *The Theory of Matrices*, Vol. 1, AMS Chelsea Publishing.

Golub, G. H. & Van Loan, C. F. (1996). *Matrix Computations*, 3 edn, John Hopkins University Press.

Haykin, S. (2002). *Adaptive Filter Theory*, 4 edn, Prentice-Hall.

Horn, R. A. & Johnson, C. R. (1991). *Topics in Matrix Analysis*, Cambridge University Press.

Jakes, W. C. (1974). *Microwave Mobile Communications*, John Wiley and Sons.

Jamoos, A., Grivel, E., Bobillet, W. & Guidorzi, R. (2007). Errors-In-Variables-Based Approach for the Identification of AR Time-Varying Fading Channels, *IEEE Signal Processing Letters* 14(11): 793–796.

Kailath, T., Sayed, A. H. & Hassibi, B. (2000). *Linear Estimation*, Prentice Hall.

Kaiser, T., Bourdoux, A., Boche, H., Fonollosa, J. R., Andersen, J. B. & Utschick, W. (eds) (2005). *Smart Antennas – State of the Art*, Hindawi Publishing Corporation.

Komninakis, C., Fragouli, C., Sayed, A. H. & Wesel, R. D. (2002). Multi-Input Multi-Output Fading Channel Tracking and Equalization Using Kalman Estimation, *IEEE Transactions on Signal Processing* 50(5): 1065–1076.

Larsson, E. & Stoica, P. (2003). *Space-Time Block Coding for Wireless Communications*, Cambridge University Press.

Larsson, E., Stoica, P. & Li, J. (2003). Orthogonal Space-Time Block Codes: Maximum Likelihood Detection for Unknown Channels and Unstructured Interferences, *IEEE Transactions on Signal Processing* 51(2): 362–372.

Li, X. & Wong, T. F. (2007). Turbo Equalization with Nonlinear Kalman Filtering for Time-Varying Frequency-Selective Fading Channels, *IEEE Transactions on Wireless Communications* 6(2): 691–700.

Liu, Z., Ma, X. & Giannakis, G. B. (2002). Space-Time Coding and Kalman Filtering for Time-Selective Fading Channels, *IEEE Transactions on Communications* 50(2): 183–186.

Loiola, M. B., Lopes, R. R. & Romano, J. M. T. (2009). Kalman Filter-Based Channel Tracking in MIMO-OSTBC Systems, *Proceedings of IEEE Global Communications Conference, 2009 – GLOBECOM 2009.*, IEEE, Honolulu, HI.

Piechocki, R. J., Nix, A. R., McGeehan, J. P. & Armour, S. M. D. (2003). Joint Blind and Semi-Blind Detection and Channel Estimation for Space-Time Trellis Coded Modulation Over Fast Faded Channels, *IEE Proceedings on Communications* 150(6): 419–426.

Simon, D. (2006). *Optimal State Estimation - Kalman, H_∞, and Nonlinear Approaches*, John Wiley and Sons.

Tarokh, V., Jafarkhani, H. & Calderbank, A. R. (1999). Space-Time Block Codes from Orthogonal Designs, *IEEE Transactions on Information Theory* 45(5): 1456–1467.

Vucetic, B. & Yuan, J. (2003). *Space-Time Coding*, John Wiley and Sons.

Adaptive Filtering Applications

Edited by Dr Lino Garcia

ISBN 978-953-307-306-4
Hard cover, 400 pages
Publisher InTech
Published online 24, June, 2011
Published in print edition June, 2011

Adaptive filtering is useful in any application where the signals or the modeled system vary over time. The configuration of the system and, in particular, the position where the adaptive processor is placed generate different areas or application fields such as: prediction, system identification and modeling, equalization, cancellation of interference, etc. which are very important in many disciplines such as control systems, communications, signal processing, acoustics, voice, sound and image, etc. The book consists of noise and echo cancellation, medical applications, communications systems and others hardly joined by their heterogeneity. Each application is a case study with rigor that shows weakness/strength of the method used, assesses its suitability and suggests new forms and areas of use. The problems are becoming increasingly complex and applications must be adapted to solve them. The adaptive filters have proven to be useful in these environments of multiple input/output, variant-time behaviors, and long and complex transfer functions effectively, but fundamentally they still have to evolve. This book is a demonstration of this and a small illustration of everything that is to come.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Murilo B. Loiola, Renato R. Lopes and João M. T. Romano (2011). Adaptive Channel Estimation in Space-Time Coded MIMO Systems, Adaptive Filtering Applications, Dr Lino Garcia (Ed.), ISBN: 978-953-307-306-4, InTech, Available from: http://www.intechopen.com/books/adaptive-filtering-applications/adaptive-channel-estimation-in-space-time-coded-mimo-systems

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen