EURASIP Journal on Applied Signal Processing 2005:5, 611–625

A Kalman-Filter Approach to Equalization of CDMA Downlink Channels

Hoang Nguyen
Nokia Research Center, San Diego, CA 92131, USA
Email: hoang.nguyen@nokia.com

Jianzhong Zhang
Nokia Research Center, Irving, TX 75039, USA
Email: charlie.zhang@nokia.com

Balaji Raghothaman
Nokia Research Center, San Diego, CA 92131, USA
Email: balaji.raghothaman@nokia.com

Received 31 July 2003; Revised 26 February 2004

An efficient method for equalization of downlink CDMA channels is presented. By describing the observed signal in terms of a state-space model, the method employs the Kalman filter (KF) to achieve an unbiased signal estimate satisfying the linear minimum mean-squared error (LMMSE) criterion. The state-space model is realized at the symbol and chip levels. With the symbol-level model, the KF is used to estimate the transmitted chips that correspond to each symbol interval; whereas at the chip level, the transmitted chips are estimated individually. The symbol-level KF has a built-in tracking capability that takes advantage of the a priori known scrambling sequence, which renders the transmitted signal nonstationary. The chip-level KF reduces the complexity of the symbol-level KF significantly by ignoring the nonstationarity introduced by scrambling. A simple method for further reducing the KF complexity is also presented. The computational complexity of the proposed technique is analyzed and compared with that of several linear approaches based on finite-impulse response (FIR) filtering. Simulations under realistic channel conditions are carried out which indicate that the KF-based approach is superior to FIR equalizers by 1–2 dBs in error-rate performance.

Keywords and phrases: Kalman filter, CDMA, single-user detection, oversampling, color noise, state-space models.

1. INTRODUCTION

The existence of dispersive effects of the channel, such as multipath, destroys the orthogonality of the spreading codes in CDMA systems. As a result, the conventional RAKE receiver in many cases reaches a noise floor at a fairly high frame error rate [1, 2]. There are two major categories of methods by which channel dispersion can be dealt with, namely multiuser detection and single-user detection. In multiuser detection, which generally requires knowledge of all user codes, the channel is equalized and every user is detected so that this approach is suitable for uplink CDMA. In downlink CDMA, however, all users share the same physical channel and multiuser detection is inefficient since only one user needs to be demodulated. To efficiently demodulate just the desired user, the channel is equalized in order to restore the code orthogonality inherent in the signal structure. Restoration of code orthogonality is thus motivated by the fact that to demodulate the desired user, (i) the receiver does not need to know the spreading codes of the interfering users, and (ii) the computational complexity of the detection mechanism does not increase with the number of users.

Well-known methods in the literature for orthogonality restoration rely on the LMMSE criterion attained via FIR filtering, which can be realized at the chip or symbol level in adaptive or batch form. For each segment of data over which the observation sequence is considered stationary, the batch implementation requires the solution to a matrix equation which can be obtained by means of LU decomposition or matrix inversion. Frequent matrix solution imposes a heavy computational burden. In an effort to reduce this computational cost, an approximate solution to the matrix inversion problem is developed in [3] based on FFT and IFFT operations by taking advantage of the near-circulant Toeplitz structure of the observed-data autocorrelation matrix. We note parenthetically that symmetric Toeplitz equations can also be solved by means of the Levinson recursion [4, 5, 6] whose complexity is on the order of the square

of the system dimension. Even for stationary channels, the symbol-level LMMSE equalizer of [7] requires one matrix solution per symbol period. This equalizer in essence is a chip-level FIR (cFIR) LMMSE equalizer that encompasses the descrambling and despreading operations, normally done outside the equalizer. Its objective, however, is to minimize the error variance of the symbol rather than chip estimate. On the other hand, matrix inversion can be avoided by means of chip-adaptive FIR equalizers. Adaptive implementations typically employ well-known algorithms such as stochastic gradient, LMS, and RLS. Adaptive FIR equalizers unfortunately tend to be unstable, sensitive to initialization, and slow to converge. Examples of stochastic gradient descent implementations of the LMMSE criterion can be found in [8, 9].

An attractive alternative to FIR filtering is to use the Kalman filter (KF) which we consider in this paper. The KF is preferred to FIR approaches for several important reasons. First, the KF has the capability to track nonstationarities which result typically from the time-variant characteristics of the channel, noise processes, and the underlying signal to be estimated. Also, the KF is well known as an optimal linear estimation method in the mean-squared error (MSE) sense.

To clearly differentiate our proposed method from other applications of the KF, it is worth discussing CDMA state-space formulations proposed by several researchers. Most notable perhaps is the work of Iltis et al. [10, 11] where a state-space model is developed for estimation of the path gains and delays of multipath channels. The observation equation of this model depends nonlinearly on the path delays and is therefore linearized with respect to these variables so that the extended Kalman filter (EKF) [12] is suitable for parameter tracking. Further, in [11], multiuser detection at each time epoch is performed based on the maximum a posteriori probability (MAP) criterion evaluated from the estimated model up to the previous time epoch; in this respect, one can view the multiuser detector as one of decision-feedback type. Similar use of the EKF with linearized state-space models for estimation and tracking of multipath gains and delays can also be found in [13, 14, 15, 16]. For the case of flat Rayleigh fading channels, relative delay estimation is no longer necessary since there is only one path to consider. If the fade process admits a Gauss-Markov model, then it is possible to employ the KF in the decision-feedback mode to track the fade [17, 18, 19].

Apart from the above applications of the KF, we discuss several state-space formulations which are more closely related to our proposed approach. In particular, the state-space models of [20, 21, 22, 23] have a resemblance to one another as seen from the fact that the measurement matrix consists of all spreading codes and channel coefficients, while the state vector consists of multiuser data symbols. Similarly, the models of [24, 25] multiplicatively lump the channel coefficients and user symbols into the state vector while the spreading codes are incorporated in the measurement matrix. It should be noted that the state-space model of [21] for multiuser detection (not equalization) is a special case of that developed in [20] without channel dispersion.

In consideration of complexity, it is worth paying special attention to the state-space model of [24]. In this model, each element of the state vector is the product of a channel tap value and a user-transmitted symbol. What makes this formulation interesting, in particular, is the fact that it yields the symbol estimate, rather than chip estimate, while the KF for this model operates at the chip level rather than symbol level since the model supplies chip-wise observations. This strategy avoids matrix inversion in computing the Kalman gain, while taking advantage of the ability of the KF to handle nonstationary state dynamics which occurs at the symbol boundaries.

In conclusion, we note that all of the above-mentioned state-space formulations require knowledge of all spreading codes, normally unavailable in practical downlink CDMA receivers. Such models are appropriate for multiuser detection and equalization, perhaps in the uplink, but inefficient for channel equalization and detection of just one desired user in a downlink multiple-access channel. We observe also that none of the aforementioned state-space models encompass a scrambling code, except for [13, 14] which include the possibility of handling long codes but no scrambling is incorporated explicitly. Scrambling codes are used in the original and evolving CDMA standards such as IS-95 and 1X EV-DV to mitigate intercell interferences; their use, however, renders the transmitted signal nonstationary. This nonstationarity is taken into account by our symbol-level state-space formulation.

An anonymous referee points our attention to [26] published during the review phase of this paper. In [26], a nonlinear state-space model is developed for joint channel estimation and equalization via the EKF. Though the model is valid, there appears to be an implementation error in [26, Section V.B.] in applying the EKF to the linearized model: in addition to replacing the measurement matrix $\mathbf{H}_k$ in [26, (8), (9), and (10)] with the Jacobian $\mathbf{H}'_k$ as indicated, the observation $z[k]$ in [26, (8)] must also be replaced with $z[k] - \mathbf{H}(\mathbf{p}_k, \hat{\mathbf{x}}_{k|k-1}) + \mathbf{H}'_k \hat{\mathbf{x}}_{k|k-1}$, since $z[k]$ is linearized about $\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}$. No mention about this later replacement is made in [26]; if this was indeed overlooked, it could be the main reason causing the performance failure of the EKF reported in [26, Figure 1]. Nonetheless, the model presented in [26, Section IV] represents a special case of our chip-level state-space model applied to single-input single-output systems.

In this paper, we elaborate on the results presented in [27, 28] where we apply the Kalman filter to single-user detection in downlink CDMA. The objective is to combat channel distortions by restoring the signal orthogonality so that demodulation can be done for just the desired user. To achieve this objective, we develop two state-space models for downlink CDMA channels, namely the symbol- and chip-level models. The symbol-level model incorporates the nonstationarity of the transmitted signal and, hence, has an enhanced tracking capability. On the other hand, the chip-level model ignores this nonstationarity in order to achieve a lower computational complexity. In addition, the computational complexity of the proposed method is analyzed for each state-space model and compared with several FIR approaches.

To take advantage of slowly fading channels which exist in many environments, a complexity reduction technique is described which yields a simple mechanism for making tradeoffs between tracking and complexity. We demonstrate also that the proposed state-space models have a signal delay structure so that the KF naturally yields fixed-lag signal estimates. In this paper, we assume that the channel impulse response is available since it can be estimated relatively accurately from the CDMA pilot tone having sufficiently high power.

The rest of the paper is organized as follows. In Section 2, we develop the symbol- and chip-level state-space models for the downlink CDMA channels. In Section 3, a complexity reduction technique is described. Fixed-lag smoothing via several ways of model augmentation is discussed in Section 4. Section 5 contains a complexity analysis and a brief description of several FIR filtering methods to which the performance of the KF techniques is compared. In Section 6, results of numerical simulations obtained in accordance with industrial standards under realistic multipath fade conditions are given. Finally, concluding remarks are given in Section 7.

2. STATE-SPACE DESCRIPTIONS OF DOWNLINK CDMA CHANNELS

For a multiple-input multiple-output (MIMO) system with M transmit and N receive antennas, we assume that the transmit antennas emit uncorrelated data streams. Nonetheless, the same set of spreading codes is used for every transmit antenna. In the sequel, the superscript or subscript t, such as in $x_u^t(i)$, indexes the tth transmit antenna. Since the Kalman filter is well known, no discussion thereof is given here. A comprehensive treatment of the Kalman filter and its properties can be found in [12, 29]. The one-step-ahead prediction and the filtered estimates of the state vector $\mathbf{x}(k)$ are denoted by $\mathbf{x}(k|k-1)$ and $\mathbf{x}(k|k)$, respectively, where k represents a generic time index which may have different interpretations for different state-space models.

2.1. MIMO CDMA signal model

Consider a U-user system where the transmitted chip sequence of the uth user is

$$x_u^t(i) = A_u a_u^t\left(\left\lfloor \frac{i}{F} \right\rfloor\right) s_u(|i|_F). \tag{1}$$

Here, i denotes the chip index; A_u is the signal amplitude; $a_u^t(n)$ is a possibly coded sequence of i.i.d. data symbols; $\underline{\mathbf{s}}_u = [s_u(0), \ldots, s_u(F-1)]^T$ is the spreading sequence; and F represents the spreading factor. The notations $\lfloor \cdot \rfloor$ and $|\cdot|_F$ denote the floor function (rounding toward $-\infty$) and the modulo-F reduction, respectively. Let $c(i)$ denote the base-station dependent scrambling sequence which has a constant modulus. The total signal transmitted via the tth transmit antenna takes the form

$$x^t(i) = c(i) \sum_{u=1}^{U} x_u^t(i), \tag{2}$$

which is the sum of all user signals scrambled by the scrambling code. The signal observed at the rth receive antenna can be expressed as

$$\begin{aligned} y_r(i) &= \sum_{t=1}^{M} \sum_{l=0}^{D} h_{r,l}^t(i) x^t(i-l) + v_r(i) = \sum_{t=1}^{M} \mathbf{h}_r^{tT}(i) \underline{x}^t(i) + v_r(i) \\ &= \mathbf{h}_r^T(i) \underline{x}(i) + v_r(i), \end{aligned} \tag{3}$$

where for $t = 1, \ldots, M$ and $r = 1, \ldots, N$,

$$\mathbf{h}_r^t(i) \triangleq \left[h_{r,0}^t, \ldots, h_{r,D}^t\right]^T, \tag{4}$$

$$\underline{x}^t(i) \triangleq \left[x^t(i), \ldots, x^t(i-D)\right]^T, \tag{5}$$

$$\mathbf{h}_r(i) \triangleq \left[\mathbf{h}_r^{1T}(i), \ldots, \mathbf{h}_r^{MT}(i)\right]^T, \tag{6}$$

$$\underline{x}(i) \triangleq \left[\underline{x}^{1T}(i), \ldots, \underline{x}^{MT}(i)\right]^T. \tag{7}$$

Here, $h_{r,l}^t$ represents the lth tap of the composite channel impulse response between the tth transmit and the rth receive antennas sampled at the chip rate, and D denotes the channel time span in chip durations. The measurement noises $v_r(i)$ are assumed to be uncorrelated white Gaussian processes with variance $\sigma_v^2 \equiv N_o$. The superscript T denotes matrix transposition; for example, $\underline{x}^{tT}(i)$ is the transpose of $\underline{x}^t(i)$. Note that for each transmit antenna, the pilot tone is contained among the U user signals defined by (1). Typically, the pilot tones account for about 10% of the total transmit power.

It is worth pointing out that the above signal model holds for the case of fractional sampling with an integer oversampling factor, denoted by P. In that case, N represents the number of "virtual" receive antennas which is equal to P times the number of actual receive antennas. Fractional sampling, however, colorizes the noise processes $v_r(i)$ spatially (over r) and temporally (over i) if the nominal bandwidth of the receive filter is kept at the chip rate $1/T_c$, with T_c denoting the chip duration. Assuming additive white Gaussian noise (WGN) at the receiver front end, the statistics of $v_r(i)$ are completely determined by the receive filter and therefore it is possible in principle to whiten the measurement noise by employing spectral factorization [4, page 104], which requires passing the sampled signal through a whitening filter. Instead of noise whitening, which is difficult for a general process, we describe in the appendix how the measurement noise $v_r(i)$ can be treated as the output of a discrete-time filter driven by WGN. We also show that the filtered noise model can then be incorporated into the state-space model of the downlink signal, yielding a white-noise model to which the ordinary KF is applicable. Due to space limit, however, we restrict our treatment of fractional sampling to the appendix only.

2.2. Symbol-level state-space model

The symbol-level state-space model is developed in this section. We call the estimator operating based on this model the symbol Kalman filter (SKF). During the nth symbol interval,

we seek the best linear unbiased estimate (BLUE) $\mathbf{x}^t(n|n)$ of the vector $\mathbf{x}^t(n) = [x^t(nF+F-1),\ldots,x^t(nF-D)]$, for $t = 1,\ldots,M$, based on all past and present observations $\{y_r(i) \mid i \le nF+F-1; r = 1,\ldots,N\}$. The characterizer *best* here means minimum error variance. Note that, once an estimate of the chip-level sequence $\mathbf{x}^t(n|n)$ has been obtained, an estimate of the nth symbol $a_u^t(n)$ of the desired user u transmitted from antenna t can be produced from a simple despreading and descrambling operation on the first F elements of $\mathbf{x}^t(n|n)$. We define the observation at time n as

$$\mathbf{y}(n) \triangleq \left[\mathbf{y}_1^T(n),\ldots,\mathbf{y}_N^T(n)\right]^T, \tag{8}$$

where $\triangleq$ denotes definition and for $r = 1,\ldots,N$,

$$\mathbf{y}_r(n) \triangleq \left[y_r(nF+F-1),\ldots,y_r(nF)\right]^T. \tag{9}$$

From (3) and (9) we can write

$$\mathbf{y}_r(n) = \mathbf{H}_r(n)\mathbf{x}(n) + \mathbf{v}_r(n), \tag{10}$$

where

$$\mathbf{H}_r(n) = \left[\mathbf{H}_r^1(n),\ldots,\mathbf{H}_r^M(n)\right], \tag{11}$$

$$\mathbf{x}(n) = \left[\mathbf{x}^{1T}(n),\ldots,\mathbf{x}^{MT}(n)\right]^T \tag{12}$$

with

$$\mathbf{H}_r^t(n) = \begin{bmatrix} h_{r,0}^t(n) & \cdots & h_{r,D}^t(n) & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & h_{r,0}^t(n) & \cdots & h_{r,D}^t(n) \end{bmatrix}, \tag{13}$$

$$\mathbf{x}^t(n) = \left[x^t(nF+F-1),\ldots,x^t(nF-D)\right]^T, \tag{14}$$

$$\mathbf{v}_r(n) = \left[v_r(nF+F-1),\ldots,v_r(nF)\right]^T \sim \text{i.i.d. } \mathcal{CN}(\mathbf{0}, N_o\mathbf{I}_F). \tag{15}$$

In (13) we have assumed for notational simplicity that the channel is constant over each symbol period. Observe that the channel matrix $\mathbf{H}_r^t(n)$ has dimensions $F \times (F+D)$. The measurement equation can be written as

$$\mathbf{y}(n) = \mathbf{H}(n)\mathbf{x}(n) + \mathbf{v}(n), \tag{16}$$

where

$$\begin{aligned} \mathbf{H}(n) &= \left[\mathbf{H}_1^T(n),\ldots,\mathbf{H}_N^T(n)\right]^T, \\ \mathbf{v}(n) &= \left[\mathbf{v}_1^T(n),\ldots,\mathbf{v}_N^T(n)\right]^T \sim \text{i.i.d. } \mathcal{CN}\left(\mathbf{0}, N_o\mathbf{I}_{NF}\right). \end{aligned} \tag{17}$$

To describe the state dynamics, we define

$$\begin{aligned} \mathbf{w}_t(n) &\triangleq \left[x^t(nF+F-1),\ldots,x^t(nF),\mathbf{0}_{1\times D}\right]^T, \\ \tilde{\mathbf{A}} &\triangleq \begin{bmatrix} \mathbf{0}_{F\times D} & \mathbf{0}_{F\times F} \\ \mathbf{I}_D & \mathbf{0}_{D\times F} \end{bmatrix}, \\ \mathbf{w}(n) &\triangleq \left[\mathbf{w}_1^T(n),\ldots,\mathbf{w}_M^T(n)\right]^T, \\ \mathbf{A} &\triangleq \mathbf{I}_M \otimes \tilde{\mathbf{A}}, \end{aligned} \tag{18}$$

where $\mathbf{I}_a$ denotes the identity matrix of size a; $\mathbf{0}_{a\times b}$ represents an $a \times b$ zero matrix; and $\otimes$ denotes the Kronecker product operator. By considering $\mathbf{x}(n)$ as the state vector, the state dynamics admits the description

$$\mathbf{x}(n) = \mathbf{A}\mathbf{x}(n-1) + \mathbf{w}(n). \tag{19}$$

The state noise process $\mathbf{w}(n)$ is temporally white but nonstationary. In particular, its covariance matrix is time varying and has the form

$$\mathbf{Q}_w(n) \triangleq E\left[\mathbf{w}(n)\mathbf{w}^H(n)\right] = \mathbf{I}_M \otimes \mathbf{Q}(n), \tag{20}$$

where

$$\mathbf{Q}(n) \triangleq E\left[\mathbf{w}_t(n)\mathbf{w}_t^H(n)\right] = \begin{bmatrix} \tilde{\mathbf{Q}}(n) & \mathbf{0}_{F\times D} \\ \mathbf{0}_{D\times F} & \mathbf{0}_{D\times D} \end{bmatrix}, \tag{21}$$

$$\tilde{\mathbf{Q}}(n) = \mathbf{C}(n)\mathbf{R}_{ss}\mathbf{C}^H(n), \tag{22}$$

$$\mathbf{R}_{ss} = E\left[\sum_{u=1}^{U} |A_u|^2 \tilde{\mathbf{s}}_u\tilde{\mathbf{s}}_u^H\right], \tag{23}$$

$$\tilde{\mathbf{s}}_u = \left[s_u(F-1),\ldots,s_u(0)\right]^T, \tag{24}$$

$$\mathbf{C}(n) = \text{diag}\left\{c(nF+F-1),\ldots,c(nF)\right\}. \tag{25}$$

Observe that the transmitted signal is nonstationary due to the a priori known scrambling sequence $c(i)$. The measurement noise $\mathbf{v}(n)$ is assumed to be a white Gaussian process with covariance matrix

$$\mathbf{Q}_v = N_o\mathbf{I}_{NF}. \tag{26}$$

It is reasonable to assume that the measurement noise variance N_o is known from the front-end noise figure. The expectation (23) can be thought of as a long-term average taken over all unknown quantities such as the spreading codes of interfering users. The expectation operator can be dropped if the argument is given. However, if $\mathbf{R}_{ss}$ is unknown to the receiver, it can be approximated by

$$\hat{\mathbf{R}}_{ss} = N_x\mathbf{I}_F, \tag{27}$$

where

$$N_x = E\left[|x(i)|^2\right] = \sum_{u=1}^{U} |A_u|^2 \tag{28}$$

which represents the total transmitted power. Simulations indicate that $\mathbf{R}_{ss}$ and $\hat{\mathbf{R}}_{ss}$ yield similar performances. The approximation (27) is motivated by the fact that when the system is fully loaded (i.e., $U = F$) and the users have equal powers, there holds

$$\mathbf{R}_{ss}|_{U=F,A_u=A_o} = A_o^2F\mathbf{I}_F \tag{29}$$

for any orthogonal set of spreading codes. The rows or columns of an $F \times F$ Walsh-Hadamard matrix composed of ± 1's form such set of codes (see, e.g., [30, page 48], [31, page 422]).

2.3. Chip-level state-space model

The chip-level state-space model developed in this section is used to obtain the BLUE of each transmitted chip $x^t(i)$ based on all past observations $\{y_r(k) \mid k \le i,\ r = 1,\dots,N\}$ measured from an N-antenna array. We call the estimator operating based on this model the chip Kalman filter (CKF). To construct this model, we define by

$$\mathbf{y}(i) \triangleq \left[y_1(i),\dots,y_N(i)\right]^T, \tag{30}$$

$$\mathbf{H}(i) \triangleq \left[\mathbf{h}_1(i),\dots,\mathbf{h}_N(i)\right]^T, \tag{31}$$

$$\mathbf{x}(i) \triangleq \left[\underline{x}^{1T}(i),\dots,\underline{x}^{MT}(i)\right]^T, \tag{32}$$

the array observation, channel matrix, and state vector at chip time i, respectively. Recall that, for $t \in \{1,\dots,M\}$, $\underline{x}^t(i)$ is defined by (5). In (31),

$$\mathbf{h}_r(i) \triangleq \left[\mathbf{h}_r^{1T}(i),\dots,\mathbf{h}_r^{MT}(i)\right]^T, \tag{33}$$

where $\mathbf{h}_r^t(i)$ is given by (4). By defining the state vector at chip epoch i as $\mathbf{x}(i)$, the chip-level state-space model can be written as

$$\mathbf{x}(i) = \mathbf{A}\mathbf{x}(i-1) + \mathbf{w}(i), \tag{34}$$

$$\mathbf{y}(i) = \mathbf{H}(i)\mathbf{x}(i) + \mathbf{v}(i), \tag{35}$$

with

$$\begin{aligned} \mathbf{A} &\triangleq \mathbf{I}_M \otimes \tilde{\mathbf{A}}, \\ \mathbf{w}(i) &= \left[\mathbf{w}_1^T(i),\dots,\mathbf{w}_M^T(i)\right]^T, \\ \mathbf{v}(i) &= \left[v_1(i),\dots,v_N(i)\right]^T \sim \text{i.i.d. } \mathcal{CN}\left(\mathbf{0}, N_o\mathbf{I}_N\right), \end{aligned} \tag{36}$$

where

$$\begin{aligned} \tilde{\mathbf{A}} &= \begin{bmatrix} \mathbf{0}_{1\times D} & 0 \\ \mathbf{I}_D & \mathbf{0}_{D\times 1} \end{bmatrix}, \\ \mathbf{w}_t(i) &= \begin{bmatrix} x^t(i) & \mathbf{0}_{1\times D} \end{bmatrix}^T \end{aligned} \tag{37}$$

for $t = 1,\dots,M$. For this model, we assume that the scrambled transmitted signals are uncorrelated white processes so that

$$\mathbf{R}_{st}(k,l) \triangleq E\{\left[x^s(k)\right]\left[x^t(l)\right]^*\} = \delta(s-t)\delta(k-l)N_t, \tag{38}$$

where $\delta(\cdot)$ denotes the Kronecker delta function and $N_t = E[|x^t(i)|^2]$ which represents the power of the tth transmit antenna. With (38), the covariance matrix of the state noise process $\mathbf{w}(i)$ takes the form

$$\mathbf{Q}_w(i) \triangleq E\left[\mathbf{w}(i)\mathbf{w}^H(i)\right] = \begin{bmatrix} \mathbf{Q}_{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{Q}_{(M)} \end{bmatrix}, \tag{39}$$

where

$$\mathbf{Q}_{(t)} = E\left[\mathbf{w}_t(i)\mathbf{w}_t^H(i)\right] = \text{diag}\left(\left[N_t, \mathbf{0}_{1\times D}\right]\right). \tag{40}$$

By assuming that the scrambled signal driving each transmit antenna is white, we ignore its nonstationarity and, hence, forgo part of the tracking capability of the symbol-level model. In the FIR-equalizer framework (see, e.g., [3, 7, 8, 9]), the whiteness assumption is used extensively and is justified by viewing the scrambling sequence as pseudorandom. This assumption holds relatively well when the scrambling sequence is *persistently exciting*, that is, sufficiently white. The nonstationarity caused by the time variation of the channel is still retained in the model. The autocorrelation matrix of the measurement noise is given by

$$\mathbf{Q}_v = N_o\mathbf{I}_N. \tag{41}$$

3. REDUCED-COMPLEXITY KALMAN FILTER

As mentioned earlier, LMMSE equalization based on FIR filtering requires the solution to a linear matrix equation in order to obtain the FIR equalizer. For stability and reliability, FIR equalizers are usually implemented in a block-by-block manner. If the time span of each block is small compared to the coherence time of the channel, then the channel can be assumed constant over each block so that the equalizer needs to be recomputed only once per block. Hence, as the channel coherence time increases, the equalizer is updated less frequently so that the complexity of the block FIR equalizer decreases accordingly. By following this strategy, the operation of the Kalman filter can be modified slightly to trade away some tracking capability in exchange for a lower computational complexity in the presence of long coherence time. In the extreme case where the signal model is perfectly stationary, the FIR equalizer needs to be computed only once. Similarly, the Kalman-filter approaches the causal Wiener filter (see [12, page 254], [4, page 378]) and therefore its steady-state solution can be synthesized off-line and used to filter the entire observation record. The state prediction error covariance matrix $\mathbf{P}(k|k-1)$, the Kalman gain $\mathbf{K}(k)$, and the filtering error covariance matrix $\mathbf{P}(k|k)$ converge to their respective steady-state values.

Consider the state-space model

$$\begin{aligned} \mathbf{x}(k) &= \mathbf{A}\mathbf{x}(k-1) + \mathbf{w}(k), \\ \mathbf{y}(k) &= \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k). \end{aligned} \tag{42}$$

We assume that the channel remains relatively constant over a long time interval so that $\mathbf{H}(k) = \mathbf{H}$ over this time period. We assume also that the transmitted signal is stationary so that $\mathbf{Q}_w(k) = \mathbf{Q}_w$. This assumption is satisfied by the chip-level state-space model of Section 2.3. If approximation (27) is used, then the stationarity assumption is implied so that the symbol-level state-space model of Section 2.2 is considered stationary by noticing that the scrambling code has a constant modulus. The time index k stands for the chip and symbol index when applied to the chip-level and symbol-level state-space models, respectively. It is easy to verify that

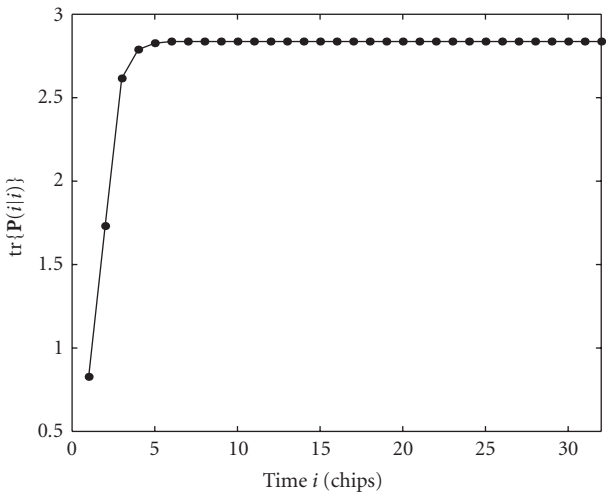


FIGURE 1: Convergence behavior of $\mathrm{tr}\{\mathbf{P}(i|i)\}$ produced by the CKF for a SISO 25-user system ($M = N = 1$, $U = 25$) with Veh-A channel model at vehicle speed $v = 50$ km/h.

all the eigenvalues of the $\mathbf{A}$-matrices of the state-space models of Section 2 are zero so that the models are stable. Then, given a nonnegative definite $\mathbf{P}(k_o|k_o - 1)$ for some k_o, as $k \to \infty$, $\mathbf{P}(k|k-1)$ converges to the steady-state value $\mathbf{P}$, independent of $\mathbf{P}(k_o|k_o - 1)$, that satisfies the well-known Riccati equation

$$\mathbf{P} = \mathbf{A}\Big[\mathbf{P} - \mathbf{P}\mathbf{H}^H\big(\mathbf{H}\mathbf{P}\mathbf{H}^H + \mathbf{Q}_v\big)^{-1}\mathbf{H}\mathbf{P}\Big]\mathbf{A}^H + \mathbf{Q}_w. \tag{43}$$

Therefore, $\mathbf{K}(k)$ and $\mathbf{P}(k|k)$ also converge to their respective steady-state solutions which follow directly from the Kalman filter recursion, that is, $\mathbf{K}(k) \to \mathbf{K} = \mathbf{P}\mathbf{H}^H(\mathbf{H}\mathbf{P}\mathbf{H}^H + \mathbf{Q}_v)^{-1}$ and $\mathbf{P}(k|k) \to \mathbf{P}^+ = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}$. In practice, it takes only a small number of steps for the Kalman filter to converge. The Figure 1 shows an example where the settling time is only about 5 chip periods (see also [24, Figure 2] and [32, Example 4.6]).

To reduce the complexity of the Kalman filter, let l_s denote the number of consecutive time steps for which the state-space model can be considered stationary. Let l_t represent the settling time of the KF for the given model. Since l_s, which on the order of the coherence time of the channel, is generally large and l_t is generally small, we have $l_t \ll l_s$. Then, in each data segment of l_s time steps, $\mathbf{P}(k|k-1)$, $\mathbf{K}(k)$, and $\mathbf{P}(k|k)$ need to be computed only during the first l_t steps. In other words, the subroutine for updating these quantities can be turned on for a period of l_t units of time and then off for a period of $l_s - l_t$ units of time, and the process is repeated periodically. The resulting recursion can be summarized as follows.

(1) Initialize $\mathbf{P}^+ = \mathbf{P}(0|0)$ and $\mathbf{x}(0|0) = \mathbf{0}$. For $k = 1, 2, \ldots$, execute the following recursion.

(2) Compute the time update

$$\mathbf{x}(k|k-1) = \mathbf{A}\mathbf{x}(k-1|k-1). \tag{44}$$

(3) Filter update: let $\tilde{l}_t$ and $\tilde{l}_s$ satisfy $l_t \le \tilde{l}_t \le \tilde{l}_s \le l_s$. If $|k|_{\tilde{l}_s} \le \tilde{l}_t$, compute

$$\begin{aligned}\mathbf{P}^- &= \mathbf{A}\mathbf{P}^+\mathbf{A}^H + \mathbf{Q}_w,\\ \mathbf{K} &= \mathbf{P}^-\mathbf{H}^H\big(\mathbf{H}\mathbf{P}^-\mathbf{H}^H + \mathbf{Q}_v\big)^{-1},\\ \mathbf{P}^+ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^-.\end{aligned} \tag{45}$$

Recall that $|\cdot|_l$ denotes the modulo-l reduction.

(4) Compute the measurement update

$$\mathbf{x}(k|k) = \mathbf{x}(k|k-1) + \mathbf{K}\big[\mathbf{y}(k) - \mathbf{H}(k)\mathbf{x}(k|k-1)\big]. \tag{46}$$

Note that the full-complexity KF ignores the condition $|k|_{\tilde{l}_s} \le \tilde{l}_t$ in step (3) and performs the filter update (45) at every time epoch. Thus, if we set $\tilde{l}_s = \tilde{l}_t = 1$, then the above recursion operates just like the original KF. We observe that the complexity of the routine for updating $\mathbf{P}(k|k-1)$, $\mathbf{K}(k)$, and $\mathbf{P}(k|k-1)$ is reduced by a factor of

$$f \triangleq \frac{\tilde{l}_s}{\tilde{l}_t}. \tag{47}$$

Thus, tradeoffs between complexity and tracking can be easily achieved by controlling the values of $\tilde{l}_s$ and $\tilde{l}_t$. Furthermore, the time update (44) involves only a shift operation and, hence, incurs essentially no computing cost. On the other hand, the key computing cost is incurred by the filter update (45) and this cost is reduced by a factor of f.

4. FIXED-LAG SMOOTHING

The signal estimate can be improved by employing fixed-lag smoothing which, for a fixed delay $k_o > 0$ and all k, yields the signal estimate $\mathbf{x}(k - k_o|k)$ of $\mathbf{x}(k - k_o)$ based on all past observations $\{\mathbf{y}(l) \mid l \le k\}$. The amount of improvement is increasing with k_o, though the marginal gain in general diminishes as k_o increases. For a general state-space model, fixed-lag smoothing can be done by applying the Kalman filter to the augmented state-space model of [12, page 176]. However, the state-space models of Section 2 have a special chip-delay structure which allows for a simpler augmentation.

To see this special feature of the state-space models, we first examine the chip-level model of Section 2.3. From definition (32) of the state vector $\mathbf{x}(i)$, we see from (5) that at time i the state $\mathbf{x}(i)$ contains the chip samples $x^t(i - D)$ for $t = 1, 2, \ldots, M$. Therefore, at time i the lag-D smoothed estimate of the vector $[x^1(i - D), x^2(i - D), \ldots, x^M(i - D)]^T$ is automatically available from $\mathbf{x}(i|i)$ without actually performing any smoothing.

Similarly, for the symbol-level state-space model, we see from (12) and (14) that during the nth symbol period the smoothed estimates $x^t(nF - 1|nF + F - 1)$, $x^t(nF - 2|nF + F - 1), \ldots, x^t(nF - D|nF + F - 1)$ of the respective chips $x^t(nF-1), x^t(nF-2), \ldots, x^t(nF-D)$ are automatically available from $\mathbf{x}(n|n)$. Note that they are smoothed estimates because they belong to one or more symbols transmitted *before* the nth symbol period while they depend on the future observations $\{y_r(l) \mid l = nF, \ldots, nF+F-1; r = 1, \ldots, N\}$ collected

Table 1: Per-chip complexity for SISO links ($M = N = 1$).

Algorithm	Parameter	Complex multiplications	Complex additions
CKF	$p = D + 1$	$\mathcal{O}\left(\frac{1.5p^2 + 1.5p}{f} + 2p\right)$	$\mathcal{O}\left(\frac{1.5p^2}{f} + 2p\right)$
SKF	$p = D + F$	$\mathcal{O}\left(\frac{1.5p^2 + 1.5pF + F^2}{f} + 2p\right)$	$\mathcal{O}\left(\frac{p^2 + 2pF + F^2}{f} + 2p\right)$
Recursive FIR	$p = L + 1$	$\mathcal{O}(2p^2 + Dp + 2p)$	$\mathcal{O}(1.5p^2 + Dp)$
Block LMMSE	$p = L + 1$	$\mathcal{O}\left(\frac{p^3 + Dp}{K} + 2p\right)$	$\mathcal{O}\left(\frac{p^3 + Dp}{K} + 2p\right)$

Table 2: Per-chip complexity for MIMO links.

Algorithm	Parameter	Complex multiplications	Complex additions
CKF	$p = (D + 1)M$	$\mathcal{O}\left(\frac{1.5Np^2 + 1.5N^2p}{f} + 2pN\right)$	$\mathcal{O}\left(\frac{1.5Np^2 + 1.5N^2p - 2Np}{f} + 2Np\right)$
Block LMMSE	$p = (L + 1)N$	$\mathcal{O}\left(\frac{p^3 + MNDp}{K} + 2Mp\right)$	$\mathcal{O}\left(\frac{p^3 + MNDp}{K} + Mp + N^2L\right)$

during the nth symbol period. These smoothed estimates can be substituted for the filtered estimates $x^t(nF - 1|nF - 1)$, $x^t(nF - 2|nF - 1), \ldots, x^t(nF - D|nF - 1)$ when making a decision on the $(n - 1)$th symbol. In general, they can be used to make decision on the $(n - \lceil D/F \rceil)$th symbol, where $\lceil \cdot \rceil$ denotes the ceiling function (rounding toward ∞). Thus, the decision lag can be as large as $\lceil D/F \rceil$ symbol periods.

From the above discussion, we see that a simple augmented state-space model can be obtained by replacing D with

$$\tilde{D} = D + \Delta, \tag{48}$$

where $\Delta > 0$, and padding the channel impulse response with Δ zeros for each pair of transmit-receive antennas. Specifically, (4) and (5) are replaced by

$$\mathbf{h}_r^t(i) = \left[h_{r,0}^t, \ldots, h_{r,D}^t, \mathbf{0}_{1\times\Delta}\right]^T, \tag{49}$$

$$\underline{x}^t(i) = \left[x^t(i), \ldots, x^t(i - \tilde{D})\right]^T, \tag{50}$$

respectively. When replacing D with $\tilde{D}$, one needs to make a simple note that $h_{r,l}^t = 0$ for $l = D + 1, \ldots, \tilde{D}$. Using this augmented state-space model endows us the lag-$\tilde{D}$ smoothed estimate of $x^t(i - \tilde{D})$ based on all observations up to time i.

The augmentation method described above is different in several aspects from that of Anderson and Moore (AM) [12, page 176] and [33]. First, the AM's augmentation increases the state dimension by an integer factor whereas our augmentation, by taking advantage of the signal delay structure, allows the state vector to be extended by an arbitrary number of chips. Second, if the chip-level state-space model above is augmented according to AM's approach, then each additional lag of Δ chips increases the state dimension by $(D + 1)M\Delta$ chips and the new state vector contains some redundancy; specifically, many elements of the new state vector occur multiple times. For the symbol-level state-space model, the amount of redundancy is small compared to the state dimension, and the state vector is extended by $(F+D)M$ chips for each additional lag of one symbol, which can be larger than necessary.

We note that for a general state-space model, the augmentation method of AM results in a stable and efficient fixed-lag smoother algorithm. Furthermore, our augmented model is guaranteed to be stable since all the eigenvalues of the dynamics matrix $\mathbf{A}$ lie within the unit circle; in fact they are all zero. This fact guarantees that, for a stationary signal model, the closed-loop gain $\mathbf{A} - \mathbf{AKH}$ of the prediction filter has all its eigenvalues inside the unit circle so that the Kalman filter for the given state-space model is asymptotically stable [12, page 77].

Another augmentation arrangement described by AM [12, page 184] can also be used in lieu of the previous one. This augmentation involves a delay structure for the model input, instead of the state; a chain of time-delayed copies of the state vector; and padding the measurement matrix with zeros. The resulting model has approximately the same computational complexity as the one described earlier.

5. COMPUTATIONAL COMPLEXITY

In Tables 1 and 2, we summarize the complexity of the Kalman and FIR filtering methods measured in terms of the number of complex multiplications (CMs) and complex additions (CAs) per-chip period. Typical complexity values are listed in Tables 3 and 4 for the Veh-A channel model. The complexity of the SKF and recursive FIR methods is shown for the single-input single-output (SISO) case only, since they are computationally more intensive than the CKF and the block LMMSE filter. However, we use the SKF as a reference for performance comparison. The performance of the KF techniques will be compared with the indicated FIR filtering methods which we briefly describe next. All FIR methods based on the LMMSE criterion employ linear FIR filtering.

Table 3: Per-chip complexity for a typical SISO link with $F = 32$, $D = 5$, $L = 15$, and $K = 256$.

Algorithm	CMs	CAs
CKF ($f = 240/5$)	$\mathcal{O}(13)$	$\mathcal{O}(13)$
SKF ($f = 5$)	$\mathcal{O}(1000)$	$\mathcal{O}(1000)$
Recursive FIR	$\mathcal{O}(624)$	$\mathcal{O}(464)$
Block LMMSE	$\mathcal{O}(48)$	$\mathcal{O}(48)$

Table 4: Per-chip complexity for a typical MIMO link with $M = N = 2$, $F = 32$, $D = 5$, $L = 15$, and $K = 256$.

Algorithm	CMs	CAs
CKF ($f = 240/5$)	$\mathcal{O}(59)$	$\mathcal{O}(58)$
CKF ($f = 100/5$)	$\mathcal{O}(73)$	$\mathcal{O}(71)$
Block LMMSE	$\mathcal{O}(259)$	$\mathcal{O}(255)$

Specifically, the underlying principle is to minimize the MSE objective function

$$J(i;\mathbf{G}_i) \triangleq E\left[\left|\left|\mathbf{x}_i - \hat{\mathbf{x}}_i\right|\right|^2\right], \tag{51}$$

where

$$\mathbf{x}_i \triangleq \left[x^1(i),\ldots,x^M(i)\right]^T, \tag{52}$$

and the estimate $\hat{\mathbf{x}}_i$ of $\mathbf{x}_i$ is obtained by passing the observations through a multidimensional linear FIR filter $\mathbf{G}_i$, that is,

$$\hat{\mathbf{x}}_i = \mathbf{G}_i\mathbf{y}_i \tag{53}$$

with

$$\mathbf{y}_i = \left[\mathbf{y}^T(i+\delta),\ldots,\mathbf{y}^T(i+\delta-L)\right]^T. \tag{54}$$

Note that $\mathbf{y}(i)$ is defined by (30). Here, L and δ are design parameters which indicate the equalizer order and the number of precursor taps, respectively. The filter $\mathbf{G}_i$ is selected to minimize the MSE criterion (51) so that

$$\mathbf{G}_i = \arg\min_{\mathbf{G}} J(i;\mathbf{G}) = \mathbf{R}_{xy}(i)\mathbf{R}_i^{-1}, \tag{55}$$

where

$$\begin{aligned}\mathbf{R}_{xy}(i) &\triangleq E\left[\mathbf{x}_i\mathbf{y}_i^H\right],\\ \mathbf{R}_i &\triangleq E\left[\mathbf{y}_i\mathbf{y}_i^H\right].\end{aligned} \tag{56}$$

Assuming that the scrambled transmitted sequence is white, obtaining $\mathbf{R}_{xy}(i)$ is straightforward, so we will not discuss it further (see, e.g., [27]). Methods based on FIR filtering differ essentially by the method in which the inverse autocorrelation matrix $\mathbf{R}_i^{-1}$ implicated in (55) is estimated. In the next section, we compare the error performance of the KF approach with the following cFIR methods.

Block LMMSE

In this approach, $\mathbf{R}_i$ is assumed to be blockwise constant and is estimated by the time average

$$\hat{\mathbf{R}}(b) = \frac{1}{K}\sum_{l=0}^{K-1}\mathbf{y}_{bK+l}\mathbf{y}_{bK+l}^H, \tag{57}$$

where K and b denote the block length and block index, respectively. Matrix inversion is thus required once per block in order to compute the filter (55).

Recursive FIR

This algorithm estimates $\mathbf{R}_i$ by the weighted average

$$\hat{\mathbf{R}}_i = \lambda\hat{\mathbf{R}}_{i-1} + (1-\lambda)\mathbf{y}_i\mathbf{y}_i^H \tag{58}$$

which yields the recursion

$$\hat{\mathbf{R}}_i^{-1} = \frac{1}{\lambda}\left[\hat{\mathbf{R}}_{i-1}^{-1} - \frac{\boldsymbol{\Gamma}_i\boldsymbol{\Gamma}_i^H}{\lambda/(1-\lambda) + \mathbf{y}_i^H\boldsymbol{\Gamma}_i}\right], \tag{59}$$

where $\boldsymbol{\Gamma}_i = \hat{\mathbf{R}}_{i-1}^{-1}\mathbf{y}_i$, $\lambda \in (0,1)$ is called the forgetting factor, and $\hat{\mathbf{R}}_0^{-1}$ can be initialized via direct matrix inversion or with a diagonal matrix.

Sliding-window RLS

The complexity of the sliding-window recursive least squares (swRLS) algorithm is not shown here (see [27]), but for performance comparison in Section 6, we briefly describe this approach whereby $\mathbf{R}_i$ is estimated by the sliding time average

$$\tilde{\mathbf{R}}_i \triangleq \frac{1}{K}\sum_{l=0}^{K-1}\mathbf{y}_{i+l}\mathbf{y}_{i+l}^H. \tag{60}$$

Then,

$$\tilde{\mathbf{R}}_{i+1} = \tilde{\mathbf{R}}_i + \frac{1}{K}\mathbf{E}_i\mathbf{D}\mathbf{E}_i^H, \tag{61}$$

where

$$\begin{aligned}\mathbf{E}_i &= \left[\mathbf{y}_i \;\; \mathbf{y}_{i+K}\right],\\ \mathbf{D} &= \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}.\end{aligned} \tag{62}$$

The inverse of (61) can be expressed in the recursive form

$$\tilde{\mathbf{R}}_{i+1}^{-1} = \tilde{\mathbf{R}}_i^{-1} - \tilde{\boldsymbol{\Gamma}}_i\left[K\mathbf{D} + \mathbf{E}_i^H\tilde{\boldsymbol{\Gamma}}_i\right]^{-1}\tilde{\boldsymbol{\Gamma}}_i^H, \tag{63}$$

where $\tilde{\boldsymbol{\Gamma}}_i \triangleq \tilde{\mathbf{R}}_i^{-1}\mathbf{E}_i$. As for $\hat{\mathbf{R}}_0^{-1}$, the initial value $\tilde{\mathbf{R}}_0^{-1}$ can be obtained from direct matrix inversion or approximated by a diagonal matrix.

6. NUMERICAL SIMULATIONS

In this section, we compare the performance of the Kalman filtering techniques developed in this paper with those of

Table 5: Simulation settings.

Standard	1X EV-DV	Data rate (QPSK)	163.2 kb/s
Channels models	ITU Veh-A, ITU Veh-B	Data rate (16 QAM)	316.8 kb/s
Modulation	QPSK, 16 QAM	FEC code rate (QPSK)	$0.7083/M$
Frame duration	1.25 ms	FEC code rate (16 QAM)	$0.5156/M$
Spreading factor	$F = 32$	Pilot power	10%
Total no. of Walsh codes	$U = 25$	User power	90%
Channel estimation	Ideal	User power distribution	Uniform

the cFIR filtering methods described in Section 5. For reference, we show also the performance of the RAKE receiver. We use bit error rate (BER) and block error rate (BLER) obtained from link level simulations as the criteria for comparison. The simulations were performed in accordance with the downlink packet data channel (F-PDCH) format of the CDMA2000 1X EV-DV standard [34]. The standard allows for several combinations of modulation schemes, coding rates, spreading code assignments, and frame durations based on packet size, channel conditions, and scheduling considerations. From these allowable combinations, we have chosen two formats, one using QPSK and another using 16 QAM. Note that the 1X EV-DV standard employs forward error correcting (FEC) codes of variable rates, all of which are achieved by puncturing a rate-1/5 parallel concatenated convolutional code (PCCC), also known as a turbo code. The simulations were performed under realistic channel conditions mandated by ITU. The settings and parameters used in the simulations are listed in Table 5. Of the 32 available Walsh codes, 25 are used by various users and 1 by the pilot. Hence the simulation represents an almost fully loaded situation. Out of the total of $U = 25$ user Walsh codes, the desired user utilizes 3 of them in the case of QPSK, and 4 in the case of 16 QAM. In the following plots, dotted and solid curves are used to indicate bit error rates and block error rates, respectively. The performance curves for the block LMMSE, recursive FIR, and swRLS methods are labelled respectively by "block," "adptv" with a λ value, and "swRLS." A fixed-lag value of $\Delta = 3$ is used for the CKF in the 16 QAM case, and $\Delta = 0$ for all other cases.

6.1. SISO links

Simulation results for SISO systems are shown in Figure 2 for QPSK and Figure 3 for 16 QAM. It is interesting to see from Figure 2a that the CKF and the SKF have similar performances and are about 2 dB better than the FIR filters. Furthermore, the SKF exhibits a small performance loss when the code correlation matrix $\mathbf{R}_{ss}$ of (23) is replaced by approximation (27). In Figures 2b and 3b, the simulation is performed for a high vehicle speed of $v = 120$ km/h, where the CKF has a substantial performance gain over the FIR filtering methods. In Figure 2a, a complexity reduction factor of $f = 240/5 = 48$ degrades the performance of the CKF by about 0.5 dB as compared to the case of no complexity reduction. Furthermore, the loss of performance of the CKF due to complexity reduction is also quite small in the other cases studied.

For the simulations results shown in Figure 3, where the CKF is used to estimate a 16 QAM signal transmitted over an SISO channel, the soft-demodulated bits (centered at ± 1) associated with the nth symbol estimate are scaled by $1/s_n$ prior to decoding. The quantity s_n is the symbol error variance estimate computed from the diagonal elements of $\mathbf{P}(i|i)$, $i = nF, \ldots, nF + F - 1$. In the soft decoding of the FEC code, the soft-demodulated bits sequence is treated as if it is the coded bits sequence observed in additive white Gaussian noise with time-dependent variance. This approximate noise model is adopted in consideration of the fact that the error variance of the signal estimate is a decreasing function of the channel strength which varies over time, and such variations are indicated by the magnitude of the diagonal elements of $\mathbf{P}(i|i)$. This noise model holds well as long as there is sufficient interleaving introduced between the encoder output and the input of the bits-to-symbol mapper. The aforementioned scale factor $1/s_n$ transforms the nonstationary noise process into a one of constant variance so that the ordinary soft decoder, which is designed for stationary noise, is applicable without modification. Note that the transition from symbols to bits is a soft-mapping demodulation operation.

6.2. MIMO links

Simulation results for MIMO systems are shown in Figures 4 and 5 with QPSK and in Figure 6 with 16 QAM. We see again that the SKF and the CKF perform comparably. Also, the swRLS and block LMMSE filters have almost the same performance. Note that the results in Figure 5 are for a multiple-input single-output (MISO) system with two transmit antennas ($M = 2$) and one receive antenna ($N = 1$). For this system, the FIR filtering methods exhibit a noise floor at a block error rate slightly above 10^{-2}, while the Kalman filtering techniques show no noise floor in the given BLER range. Comparing the KF performance curves in Figure 2a with their counterparts in Figure 5b indicates that with one receive antenna no diversity gain is seen when the number of transmit antennas is increased from one to two, even though in doing so the effective FEC code rate is reduced by a factor of 2 (the total transmit energy per information bit is unchanged). On the other hand, the diversity gain is large when two receive antennas are used instead of one, as evident from Figures 4 and 5.

Note, in particular, from Figure 4b that a complexity reduction factor of $f = 240/5$, results in negligible performance loss for the CKF. In Figure 4a, the performance loss

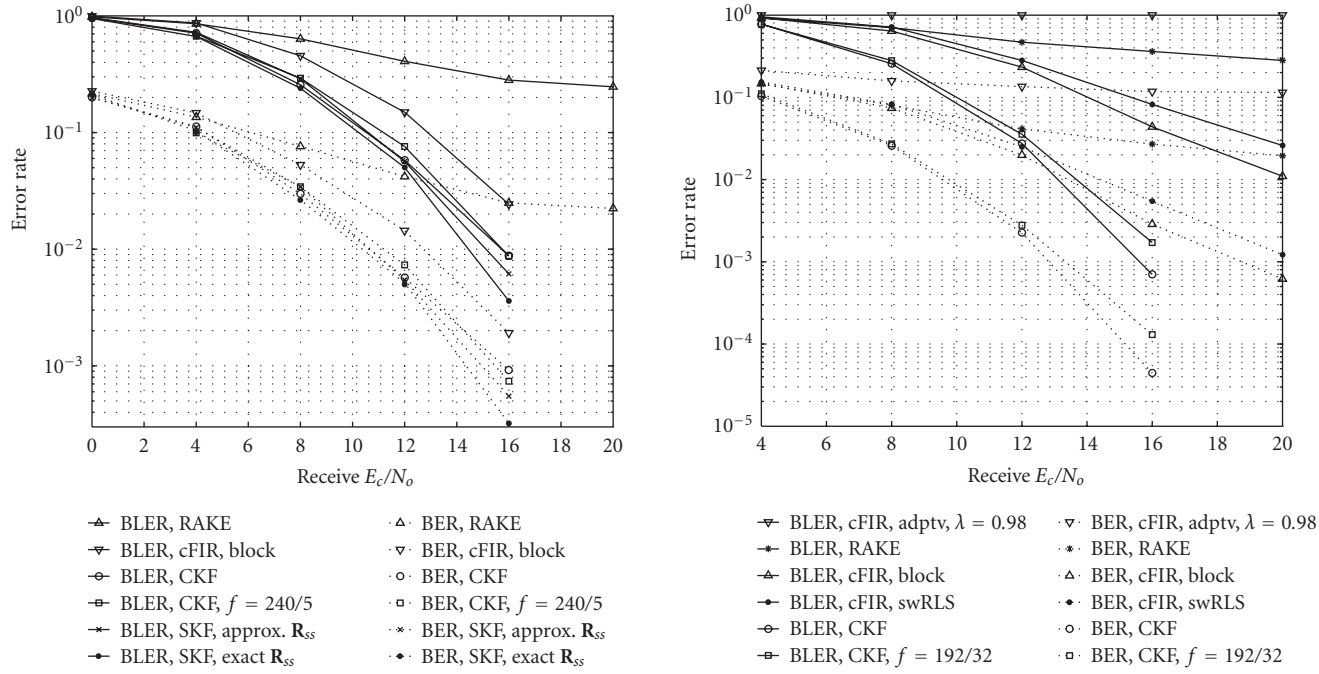


FIGURE 2: Error performance for a QPSK SISO link with $M = N = 1$: (a) $L = 15$, $K = 128$, $\delta = 7$, $f = 240/5$, Veh-A at $v = 50$ km/h and (b) $L = 33$, $K = 256$, $\delta = 16$, $\lambda = 0.98$, $f = 192/32$, Veh-B at $v = 120$ km/h.

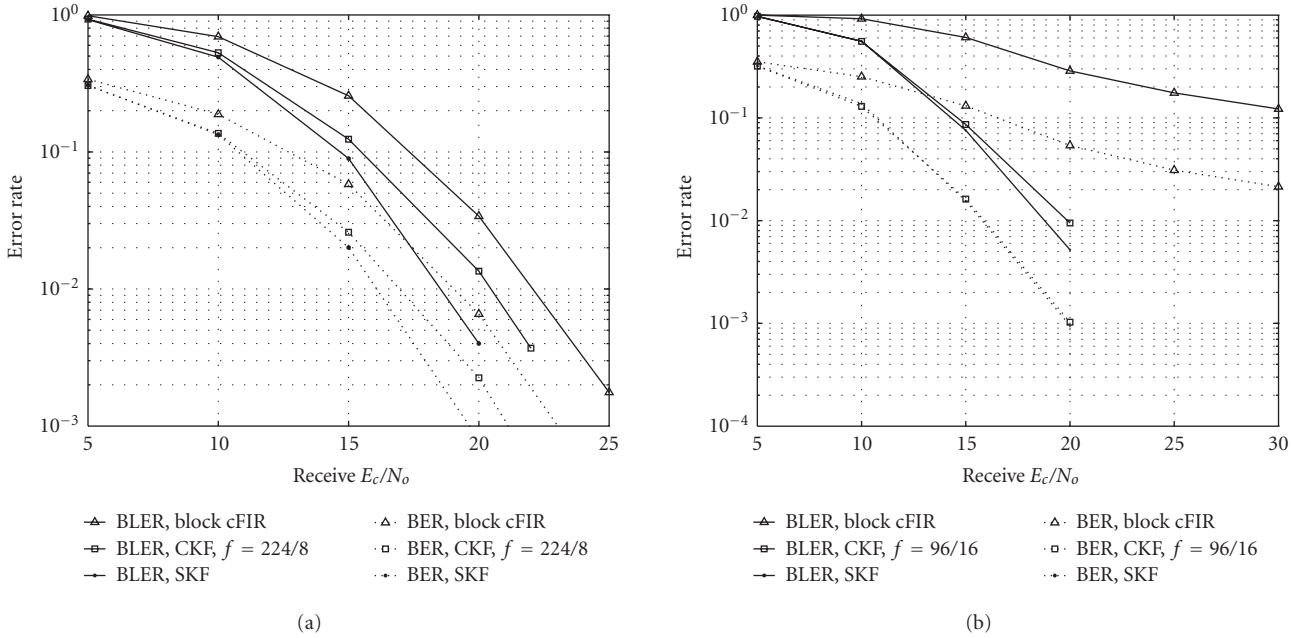


FIGURE 3: Error performance for a 16 QAM SISO link with $M = N = 1$: (a) Veh-A at 50 km/h, $L = 15$, $K = 256$, $\delta = 7$, $f = 224/8$ and (b) Veh-B at 120 km/h, $L = 41$, $K = 512$, $\delta = 20$, $f = 96/16$.

is almost unnoticeable for $f = 256/32$, indicating that the reduction factor can be increased to a higher value such as the one used in Figure 4b.

It is not surprising to see that the CKF performs as well as the SKF based on the estimated code correlation matrix $\widehat{\mathbf{R}}_{ss}$ given by (27). It can be concluded by examining

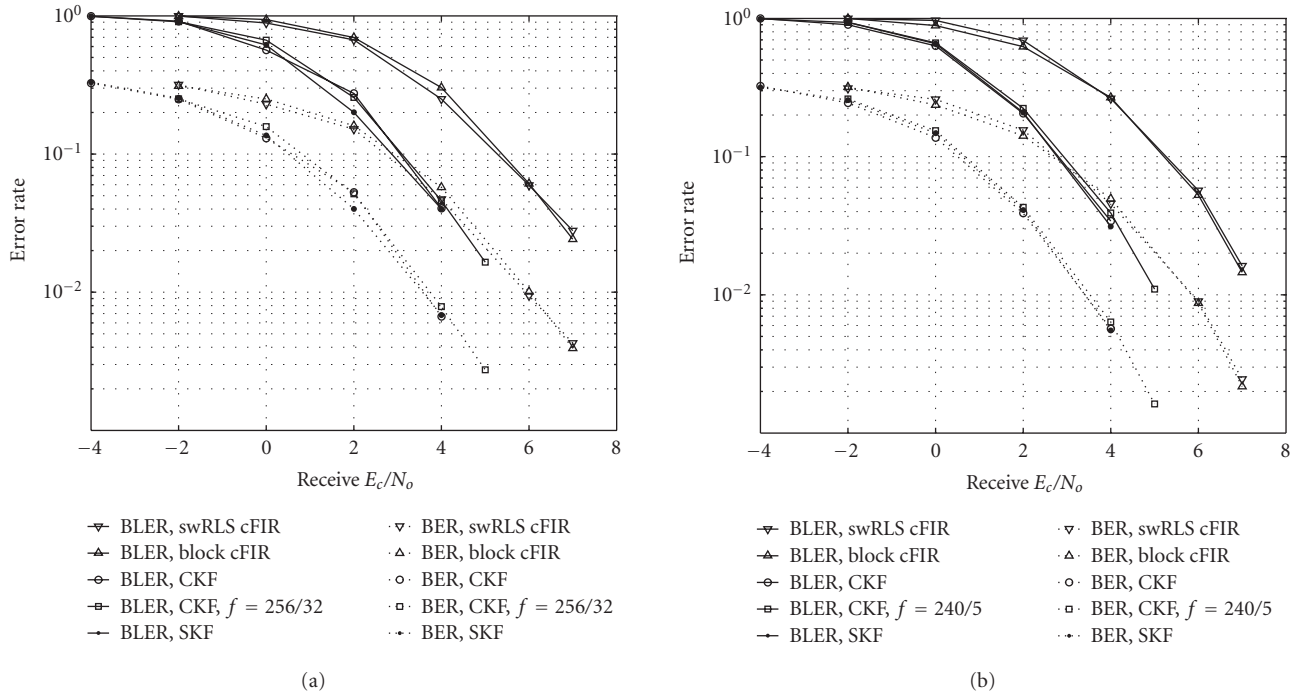


FIGURE 4: Error performance for a QPSK MIMO link with $M = N = 2$, $L = 10$, $K = 128$, $\delta = 5$, Veh-A channel model: (a) $v = 30$ km/h, $f = 256/32$ and (b) $v = 50$ km/h, $f = 240/5$.

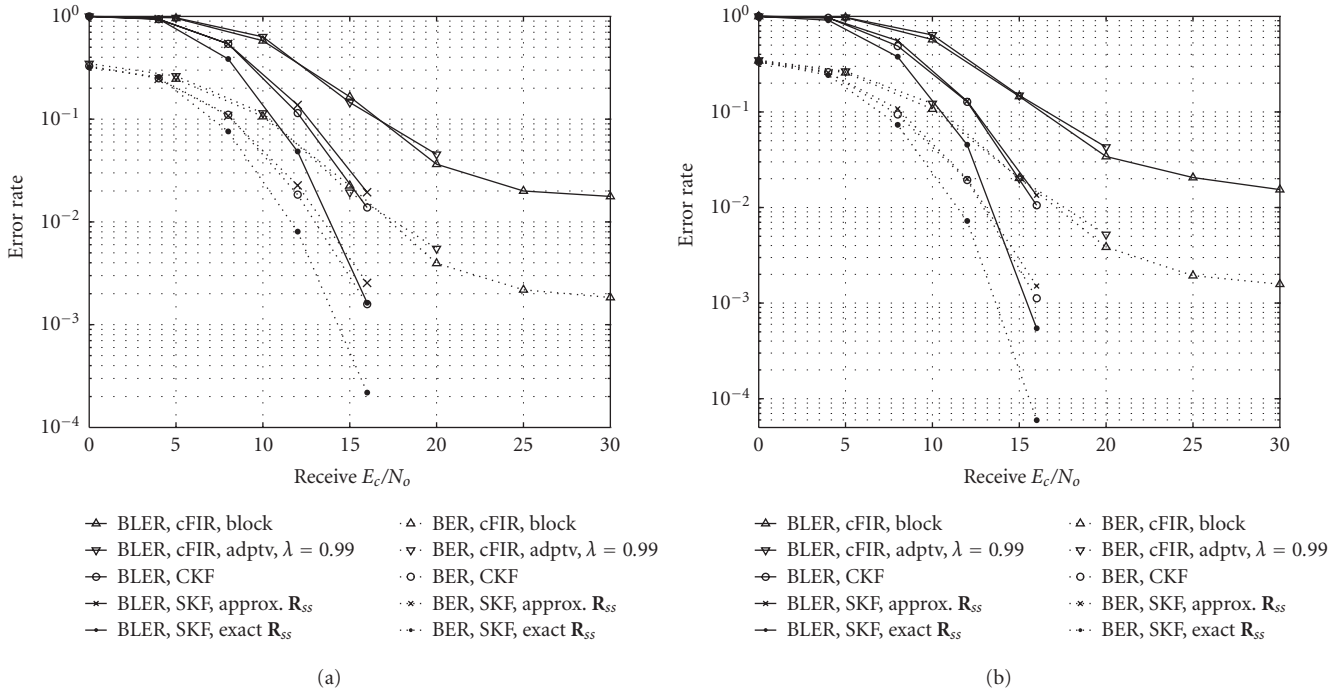


FIGURE 5: Error performance for a QPSK MISO link with $M = 2$, $N = 1$, $L = 9$, $K = 192$, $\delta = 5$, $\lambda = 0.99$, Veh-A channel model: (a) $v = 30$ km/h and (b) $v = 50$ km/h.

the state-space models that the two techniques should perform exactly the same since, with the use of (27), the nonstationary behavior of the transmitted signal is masked out by the constant-modulus property of the scrambling sequence as can be seen from (22), yielding a time-invariant $\mathbf{Q}_w(n)$.

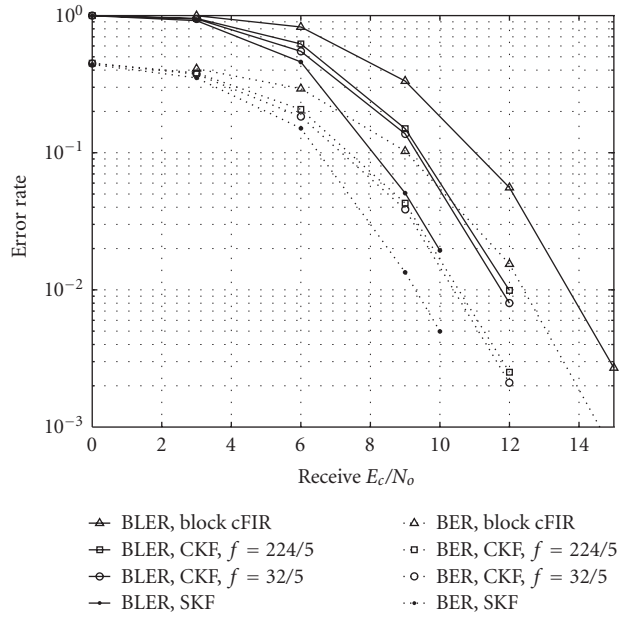


Figure 6: Error performance for a 16 QAM MIMO link with $M = N = 2$, $L = 10$, $K = 192$, $\delta = 5$, Veh-A at $v = 50$ km/h.

7. CONCLUSIONS

In this paper, we have developed symbol-level and chip-level state-space models for downlink CDMA channels. The Kalman filter is employed to estimate the transmitted signal and restore the code orthogonality so that single-user detection is feasible in multiple-access channels. Simulations indicate that the proposed Kalman filtering techniques outperform the FIR filtering approaches by 1–2 dBs. Due to the tracking capability of the KF, its performance gain over FIR methods tends to increase with the channel fading rate. Like methods based on FIR filtering which can take advantage of slow fading to reduce the computational complexity via block-by-block filter synthesis, the KF can periodically skip the update of the Kalman gain and the error covariance matrices. For moderate reduction factors, the periodic-update operation reduces the KF complexity but essentially retains the performance of the full-complexity KF.

Due to the signal delay structure, the KF automatically produces fixed-lag smoothed estimates of the transmitted signal corresponding to all lags up to the channel order. We have demonstrated that an augmented state-space model can be obtained by taking advantage of the chip-delay structure of the channel. Therefore, an arbitrary fixed lag larger than the channel order can be achieved.

We have assumed that the channel information is available. In actual receiver operation, however, the channel is estimated from the pilot tones embedded in the downlink signals and is subject to estimation error due to multiple-access interference. The channel estimate can be improved by feeding the soft signal estimate at the output of the KF back to the channel estimator and combining it with the pilots, thereby giving rise to a stronger effective pilot. The soft feedback thus accounts for all traffic interference. On the other hand, hard decision feedback from the output of the decoder through a reencoder is also possible, which, however, accounts only for the desired user traffic.

The KF can be implemented in the square-root form. This implementation is better at keeping the error covariance matrices nonnegative definite, improves the matrix condition number, and enhances the numerical stability.

Finally, the error covariance matrix $\mathbf{P}(k|k)$ contains useful information. Since the diagonal elements of the matrix are a time-varying reliability measure for the chip estimates and, hence, the symbol estimates, incorporating this information into the soft demodulator/decoder may further improve the error performance.

APPENDIX

COLOR NOISE MODEL FOR FRACTIONAL SAMPLING

Let $g_c(t)$ and P denote the impulse response of the (continuous-time) receive filter and the integer oversampling factor, respectively. Here, we consider only the case where $P > 1$, since the observation noise is automatically white when $P = 1$ if the deterministic autocorrelation of $g_c(t)$ is a Nyquist pulse. To avoid out-of-band interference, the two-sided *nominal* bandwidth of $g_c(t)$ is limited to the chip rate $1/T_c$. Therefore, for all practical purposes, $g_c(t)$ is considered band-limited. Since the two-sided bandwidth of $g_c(t)$ is smaller than the sampling rate $1/T_s$, with $T_s = T_c/P$, from the sampling theorem we can represent $g_c(t)$ by its discrete-time (DT) samples $g(i) \triangleq g_c(iT_s)$, that is,

$$g_c(t) = \sum_i g(i)\,\mathrm{sinc}\left(\frac{t}{T_s} - i\right), \tag{A.1}$$

where $\mathrm{sinc}(x) \triangleq \sin(\pi x)/(\pi x)$. Assume that the DT measurement noise results from the zero-mean additive white Gaussian noise (WGN) process $w_c(t)$ present at the input of the receive filter. By definition, the autocorrelation of $w_c(t)$ is $E[w_c(t+\tau)w_c^*(t)] = N_o\delta_c(\tau)$ where $\delta_c(\tau)$ is the Dirac delta function. The response of $g_c(t)$ to the input noise is given by

$$\begin{aligned} v_c(t) \triangleq g_c(t) * w_c(t) &= \int_{-\infty}^{\infty} g_c(\tau) w_c(t-\tau)\,d\tau \\ &= \int_{-\infty}^{\infty} \sum_i g(i)\,\mathrm{sinc}\left(\frac{\tau}{T_s} - i\right) w_c(t-\tau)\,d\tau, \end{aligned} \tag{A.2}$$

where $*$ denotes the convolution operator. When the filtered signal is oversampled by an integer factor of P, the DT output noise is

$$\begin{aligned} v(k) &\triangleq v_c(kT_s + t_o) \\ &= \sum_i g(i) \int_{-\infty}^{\infty} \mathrm{sinc}\left(\frac{\tau}{T_s}\right) w_c(kT_s - iT_s + t_o - \tau)\,d\tau \end{aligned} \tag{A.3}$$

for all integer k, where t_o is an arbitrary timing offset. We will show that $v(k)$ can be modelled as the output of a DT filter driven by WGN. We define the following:

$$\tilde{w}(k) \triangleq \int_{-\infty}^{\infty} \operatorname{sinc}\left(\frac{\tau}{T_s}\right) w_c(kT_s + t_o - \tau)\, d\tau. \tag{A.4}$$

It is easy to show that the DT process $\tilde{w}(k)$ is wide-sense stationary and admits the autocorrelation

$$R_{\tilde{w}}(l) \triangleq E\left[\tilde{w}(k+l)\tilde{w}^*(k)\right] = N_o q_c(lT_s) = N_o T_s \delta(l), \tag{A.5}$$

where $\delta(\cdot)$ is the Kronecker delta function and

$$q_c(\tau) \triangleq \int_{-\infty}^{\infty} \operatorname{sinc}\left(\frac{\tau + t}{T_s}\right) \operatorname{sinc}\left(\frac{t}{T_s}\right) dt = T_s \operatorname{sinc}\left(\frac{\tau}{T_s}\right). \tag{A.6}$$

Equation (A.5) clearly indicates that $\tilde{w}(k)$ is a WGN process. Thus, we obtain from (A.3) and (A.4) the noise model

$$v(k) = \sum_{i=0}^{D_g} g(i)\tilde{w}(k-i), \tag{A.7}$$

where we have assumed that $g_c(t)$ has a finite time span from $t = 0$ to $t = D_g T_s$ for some integer D_g. Equation (A.7) shows that $v(k)$ is the output of a DT filter driven by WGN.

We now demonstrate how (A.7) can be incorporated into the state-space model for the downlink CDMA signal. We revisit the chip-level state-space model given by (34) and (35). For simplicity, consider the case of one receive antenna sampled at the rate of P/T_c, so N is replaced by P and the channel structure remains the same. The only difference is that the noise vector $\mathbf{v}(i)$ in (35) is no longer a white process with respect to i because it is $v(k)$ demultiplexed into P parallel channels. That is, during the ith chip interval,

$$\mathbf{v}(i) = \left[v(iP), v(iP+1), \ldots, v(iP+P-1)\right]^T \tag{A.8}$$

which can be rewritten as

$$\mathbf{v}(i) = \mathbf{G}\tilde{\mathbf{w}}(i), \tag{A.9}$$

where

$$\mathbf{G} \triangleq \begin{bmatrix} 0 & \cdots & 0 & 0 & g(0) & g(1) & \cdots & g(D_g) \\ 0 & \cdots & 0 & g(0) & g(1) & \cdots & g(D_g) & 0 \\ \vdots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \vdots \\ 0 & g(0) & g(1) & \cdots & g(D_g) & 0 & \cdots & 0 \\ g(0) & g(1) & \cdots & g(D_g) & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{A.10}$$

$$\tilde{\mathbf{w}}(i) \triangleq \left[\tilde{w}(iP+P-1), \tilde{w}(iP+P-2), \ldots, \tilde{w}(iP), \ldots, \tilde{w}(iP-D_g)\right]^T. \tag{A.11}$$

By defining

$$\overline{\mathbf{w}}(i) \triangleq \left[\tilde{w}(iP+P-1), \tilde{w}(iP+P-2), \ldots, \tilde{w}(iP), \mathbf{0}_{1\times D_g}\right]^T, \qquad \mathbf{A}_w \triangleq \begin{bmatrix} \mathbf{0}_{P\times D_g} & \mathbf{0}_{P\times P} \\ \mathbf{I}_{D_g} & \mathbf{0}_{D_g\times P} \end{bmatrix}, \tag{A.12}$$

we have the following chip-level state-space model:

$$\begin{bmatrix} \mathbf{x}(i) \\ \tilde{\mathbf{w}}(i) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_w \end{bmatrix} \begin{bmatrix} \mathbf{x}(i-1) \\ \tilde{\mathbf{w}}(i-1) \end{bmatrix} + \begin{bmatrix} \mathbf{w}(i) \\ \overline{\mathbf{w}}(i) \end{bmatrix}, \tag{A.13}$$

$$\mathbf{y}(i) = \begin{bmatrix} \mathbf{H}(i) & \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{x}(i) \\ \tilde{\mathbf{w}}(i) \end{bmatrix}. \tag{A.14}$$

Note that the upper part of (A.13) comes directly from (34) and the lower part is obvious from the delay structure of (A.11). In addition, (A.14) is obtained by substituting (A.9) into (35) and, from the KF viewpoint, contains no measurement noise. Instead, the source of measurement noise is contained in the new state vector $[\mathbf{x}^T(i), \tilde{\mathbf{w}}^T(i)]^T$. Several observations are in order. First, the above development introduces no unknown parameters as $g(i)$ is completely known since it is the DT-sampled version of the front-end filter. Second, the above noise model allows $g_c(t)$ to be an arbitrary band-limited filter. Third, the new state-dynamics matrix has all its eigenvalues inside the unit circle; in fact, all the eigenvalues are zero. Therefore, the model is stable. Fourth, if the noise processes at different antenna front ends are uncorrelated, the above state-space model can be cascaded for multi-antenna receivers. Finally, this model admits the ordinary KF as the optimal linear unbiased estimator since the state noise process $\overline{\mathbf{w}}(i)$ is white over i.

ACKNOWLEDGMENTS

This research was performed while the first author was a Research Intern with Nokia Research Center, Dallas, Tex, August–December 2002. Part of the results contained in this paper has been presented at the 37th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, Calif, November 9–12, 2003.

REFERENCES

[1] M. Latva-aho, "Bit error probability analysis for FRAMES WCDMA downlink receivers," *IEEE Trans. Veh. Technol.*, vol. 47, no. 4, pp. 1119–1133, 1998.

[2] K. Hooli, M. Latva-aho, and M. Juntti, "Multiple access interference suppression with linear chip equalizers in WCDMA downlink receivers," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '99)*, vol. 1a, pp. 467–471, Rio de Janeireo, Brazil, December 1999.

[3] J. Zhang, T. Bhatt, and G. Mandyam, "Efficient linear equalization for high data rate downlink CDMA signaling," in *Proc. 37th Asilomar Conference on Signals, Systems & Computers*, vol. 1, pp. 141–145, Pacific Grove, Calif, USA, November 2003.

[4] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons, New York, NY, USA, 1996.

[5] A. E. Yagle, "A new multichannel split levinson algorithm for block Hermitian-Toeplitz matrices," *IEEE Trans. Circuits Syst.*, vol. 36, no. 6, pp. 928–931, 1989.

[6] R. R. Joshi and A. E. Yagle, "Split versions of the Levinson-like and Schur-like fast algorithms for solving block-slanted-Toeplitz systems of equations," *IEEE Trans. Signal Processing*, vol. 46, no. 7, pp. 2027–2030, 1998.

[7] T. P. Krauss, W. J. Hillery, and M. D. Zoltowski, "MMSE equalization for forward link in 3G CDMA: symbol-level versus chip-level," in *Proc. IEEE 10th Workshop on Statistical Signal and Array Processing*, pp. 18–22, Pocono Manor, Pa, USA, 2000.

[8] P. Komulainen, M. J. Heikkila, and J. Lilleberg, "Adaptive channel equalization and interference suppression for CDMA downlink," in *Proc. IEEE 6th International Symposium on Spread Spectrum Techniques and Applications*, vol. 2, pp. 363–367, Parsippany, NJ, USA, 2000.

[9] M. J. Heikkila, "A novel blind adaptive algorithm for channel equalization in WCDMA downlink," in *Proc. 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, pp. A41–A45, San Diego, Calif, USA, 2001.

[10] R. A. Iltis, "Joint estimation of PN code delay and multipath using the extended Kalman filter," *IEEE Trans. Commun.*, vol. 38, no. 10, pp. 1677–1685, 1990.

[11] R. A. Iltis and L. Mailaender, "An adaptive multiuser detector with joint amplitude and delay estimation," *IEEE J. Select. Areas Commun.*, vol. 12, no. 5, pp. 774–785, 1994.

[12] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood cliffs, NJ, USA, 1979.

[13] T. J. Lim and L. K. Rasmussen, "Adaptive symbol and parameter estimation in asynchronous multiuser CDMA detectors," *IEEE Trans. Commun.*, vol. 45, no. 2, pp. 213–220, 1997.

[14] R. A. Iltis, "A DS-CDMA tracking mode receiver with joint channel/delay estimation and MMSE detection," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1770–1779, 2001.

[15] K. J. Kim and R. A. Iltis, "Joint detection and channel estimation algorithms for QS-CDMA signals over time-varying channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 845–855, 2002.

[16] T. T. Tjhung, X. Xue, Y. Dai, K. N. Wong, and P. He, "An adaptive asynchronous CDMA multiuser detector for frequency-selective Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 51, no. 4, pp. 788–793, 2002.

[17] M. Yan and B. D. Rao, "A coherent minimum-output-energy receiver with a Kalman channel predictor for DS-CDMA systems," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02)*, vol. 3, pp. III.2301–III.2304, 2002.

[18] P. H.-Y. Wu and A. Duel-Hallen, "Multiuser detectors with disjoint Kalman channel estimators for synchronous CDMA mobile radio channels," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 752–756, 2000.

[19] K.-J. Song, S.-K. Hong, S-Y. Jung, and D.-J. Park, "Novel channel estimation algorithm using Kalman filter for DS-CDMA Rayleigh fading channel," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 4, pp. IV.429–IV.432, 2003.

[20] S. H. Isabelle and G. W. Wornell, "Efficient multiuser detectors for intersymbol interference channels," in *Proc. 1st IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, pp. 277–280, Paris, France, 1997.

[21] T. J. Lim and Y. Ma, "The Kalman filter as the optimal linear minimum mean-squared error multiuser CDMA detector," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2561–2566, 2000.

[22] L.-M. Chen, B.-S. Chen, and W.-S. Hou, "Adaptive multiuser DFE with Kalman channel estimation for DS-CDMA systems in multipath fading channels," *Elsevier Signal Processing*, vol. 81, pp. 713–733, 2001.

[23] Y. Ma and T. J. Lim, "Linear and nonlinear chip-rate minimum mean-squared-error multiuser CDMA detection," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 530–542, 2001.

[24] T. J. Lim, L. K. Rasmussen, and H. Sugimoto, "An asynchronous multiuser CDMA detector based on the Kalman filter," *IEEE J. Select. Areas Commun.*, vol. 16, no. 9, pp. 1711–1722, 1998.

[25] Z. Xu and T. Wang, "Blind detection of CDMA signals based on Kalman filter," in *Proc. IEEE 35th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1545–1549, Pacific Grove, Calif, USA, 2001.

[26] H. Gerlach, D. Dahlhaus, M. Pesce, and W. Xu, "Joint Kalman channel estimation and equalization for the UMTS FDD downlink," in *Proc. IEEE 58th Vehicular Technology Conference (VTC '03)*, vol. 2, pp. 1263–1267, Orlando, Fla, USA, 2003.

[27] H. Nguyen, J. Zhang, and B. Raghothaman, "Linear minimum MSE equalization of SISO and MIMO CDMA downlink channels via the Kalman and FIR Filters," Nokia Research Center Internal Report, December 2002.

[28] H. Nguyen, J. Zhang, and B. Raghothaman, "Equalization of CDMA downlink channel via Kalman filtering," in *Proc. 37th IEEE Asilomar Conference on Signals, Systems & Computers*, vol. 1, pp. 1128–1132, Pacific Grove, Calif, USA, 2003.

[29] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Mag.*, vol. 11, no. 3, pp. 18–60, 1994.

[30] P. Castoldi, *Multiuser Detection in CDMA Mobile Terminals*, Artech House, Boston, Mass, USA, 2002.

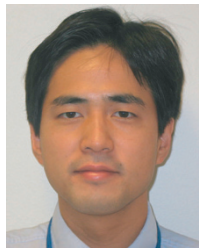
[31] J. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.

[32] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications*, Springer-Verlag, New York, NY, USA, 2000.

[33] J. B. Moore, "Discrete-time fixed-lag smoothing algorithms," *Automatica*, vol. 9, pp. 163–174, 1973.

[34] 3rd Generation Partnership Project 2, "IS-2000-2-C: Physical Layer Standard for CDMA2000 Spread Spectrum Systems - Release C," .

Hoang Nguyen received his B.S. degree (summa cum laude) in 1999 from the University of Missouri-Columbia/Kansas City (UMC/KC) and the M.S. and Ph.D. degrees in 2002 and 2003, respectively, from the University of California-Davis (UC-Davis,) all in electrical engineering. From November 1995 to August 1999, he was a member of the mathematics tutoring staff at Penn Valley Community College, Kansas City, Mo. From August to December 2002, he was a Research Intern at Nokia Research Center, Dallas, Tex, where he developed efficient equalization algorithms for downlink CDMA receivers. From January to September 2003, he held a Research Assistantship position in electrical engineering at UC-Davis. He is currently with Nokia Research Center, Dallas, Tex. His research interests include statistical signal processing, equalization, estimation, detection, and coding. He is a Member of the IEEE (Institute of Electrical and Electronics Engineers) and of the Phi Kappa Phi, Tau Beta Pi, and Eta Kappa Nu honor societies. He was the 1998 recipient of the Outstanding Junior Award of the Electrical and Computer Engineering Department, UMC/KC, and a 1997 Barry M. Goldwater Scholarship nominee. He holds a Missouri-registered Engineer-in-Training license. He is a graduate research fellow of the US National Science Foundation; he spent the three-year fellowship tenure at UC-Davis.



Jianzhong Zhang received his B.S. degrees in both electrical engineering and applied physics from Tsinghua University, Beijing, China, in 1995, the M.S. degree in electrical engineering from Clemson University in 1998, and the Ph.D. degree in electrical engineering from the University of Wisconsin, Madison, in 2003. His research has focused on the application of statistical signal processing methods to wireless communication problems. Since 2001, he has been with Nokia Research Center, Irving, Texas, where he worked on the transceiver designs for both EDGE and CDMA2000/WCDMA cellular systems.

Balaji Raghothaman was born in Coimbatore, India, in 1973. He received his B.S. degree in engineering from Coimbatore Institute of Technology in 1994. He went on to complete the M.S. and Ph.D. degrees from the University of Texas, Dallas, in 1997 and 1999, respectively, both in electrical engineering with an emphasis on signal processing for communications. During his graduate studies, he held summer internships at Texas Instruments, Dallas, in 1996 and 1997, in their wireless products groups. He has been with Nokia Research Center, Irving, Texas, from 1999 onwards, where he has been involved in research and standardization efforts in multiple antenna technologies for CDMA and WCDMA. He has about 20 papers to his credit in conferences and referees journals. He also has 2 granted patents and 13 pending patent applications. His research interests include algorithms for MIMO transmission and reception, advanced receiver algorithms for wireless communications, and spatial channel modeling for multiple antennas. He has held the positions of Program Chairman and Chairman of the Dallas Chapter of the IEEE Signal Processing Society from 2000 till 2002. He received the TxTEC Fellowship Award while being at the University of Texas, Dallas. He was also recognized as the Outstanding Engineering Student in the city of Coimbatore, India, in 1994.