A QRD-M/Kalman Filter-Based Detection and Channel Estimation Algorithm for MIMO-OFDM Systems

Kyeong Jin Kim, *Member, IEEE*, Jiang Yue, *Student Member, IEEE*, Ronald A. Iltis, *Senior Member, IEEE*, and Jerry D. Gibson, *Fellow, IEEE*

***Abstract*—The use of multiple transmit/receive antennas forming a multiple-input multiple-output (MIMO) system can significantly enhance channel capacity. This paper considers a V-BLAST-type combination of orthogonal frequency-division multiplexing (OFDM) with MIMO (MIMO-OFDM) for enhanced spectral efficiency and multiuser downlink throughput. A new joint data detection and channel estimation algorithm for MIMO-OFDM is proposed which combines the QRD-M algorithm and Kalman filter. The individual channels between antenna elements are tracked using a Kalman filter, and the QRD-M algorithm uses a limited tree search to approximate the maximum-likelihood detector. A closed-form symbol-error rate, conditioned on a static channel realization, is presented for the $M = 1$ case with QPSK modulation. An adaptive complexity QRD-M algorithm (AC-QRD-M) is also considered which assigns different values of M to each subcarrier according to its estimated received power. A rule for choosing M using subcarrier powers is obtained using a kernel density estimate combined with the Lloyd-Max algorithm.**

***Index Terms*—Frequency-division multiplexing (FDM), Kalman filtering (KF), maximum likelihood (ML), sequential detection.**

I. Introduction

VARIOUS multiple antenna systems have been introduced to increase wireless network capacity and spectral efficiency. Examples include space–time coding [1], [2] and V-BLAST [3], [4]. The focus in this paper is on a combination of V-BLAST modulation with orthogonal frequency-division multiplexing (OFDM), leading to a multiple-input multiple-output (MIMO)-OFDM system. In conventional V-BLAST, a serial data stream is converted to parallel, with each data symbol transmitted on a separate antenna. The MIMO-OFDM system considered here transmits an independent K-subcarrier OFDM signal on each of N_t antenna elements for KN_t symbols transmitted per symbol duration.

Previous combinations of MIMO with OFDM include Vector OFDM [2] and the space–time coded system in [5]. Here, the emphasis is on a spatially uncoded system motivated by V-BLAST, which allows for simpler detection and channel estimation methods. Note that the individual data streams in the proposed MIMO-OFDM system can nevertheless be temporally coded. The advantage of OFDM [6]–[8] in the MIMO application is that a frequency-selective fading channel is converted to parallel flat-fading channels, and intersymbol interference (ISI) is eliminated via use of a suitable time-guard band. The Bell Labs Layered Space–Time System (BLAST) [3], [9], [10] was originally developed for non-OFDM systems and allows for increased aggregate data rates by simultaneous transmission of independent data streams on multiple antennas. V-BLAST combines beamforming and successive interference cancellation (SIC) techniques [11], [12] in order to separate the parallel data streams.

The use of the QR decomposition (QRD)-M algorithm for MIMO-OFDM is motivated by previous work on joint data detection and channel estimation for DS-code-division mulitple access (CDMA) in [13]. In QRD-M, the QRD [14] is applied to the antenna outputs after the fast Fourier transform (FFT) operation. It is further shown that data detection can be performed independently on each of K OFDM carriers. The effective channel for N_r receive and N_t transmit antennas then reduces to K independent $N_r \times N_t$ upper triangular matrices. The maximum-likelihood (ML) decision rule based on the QRD corresponds to a full tree search [15]. However, the number of branch metrics in the tree grows exponentially with the number of transmit antennas and the size of the symbol alphabet. The M algorithm [15]–[17] combined with the QRD can greatly reduce the computational complexity of the tree search [13]. The resulting combination of QRD with MIMO was also proposed by [10] for a non-OFDM V-BLAST system. However, the approach of [10] was restricted to the $M = 1$ case. It is shown here that the more flexible search in the QRD-M algorithm will give better performance than the SIC $(M = 1)$ approaches in previous V-BLAST detection algorithms [3], [9], [10].

Channel estimation is required for accurate demodulation in the proposed MIMO-OFDM system. Here, the N_tN_r effective channels are modeled by finite-impulse response (FIR) filters. The Kalman filter (KF) is then employed in the sequel for joint estimation of channel coefficients in a manner similar to [18]. Note that to obtain a practical joint detection/estimation algorithm, the QRD-M step uses the channel estimate computed during the previous symbol interval.

An adaptive complexity QRD-M algorithm is proposed in which weaker subcarriers are assigned larger values of M in the

Manuscript received July 31, 2003; revised January 18, 2004; accepted January 20, 2004. The editor coordinating the review of this paper and approving it for publication is X. Wang. This work was supported in part by NSG Grant CCF-0429596. This paper was presented in part at the 2002 Asilomar Conference on Signals, Systems and Computers.

K. J. Kim is with the Nokia Research Center, Irving, TX 75039 USA (e-mail: kyeong.j.kim@nokia.com).

J. Yue is with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275 USA (e-mail: jyue@mail.smu.edu).

R. A. Iltis and J. D. Gibson are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560 USA (e-mail: iltis@ece.ucsb.edu; gibson@mat.ucsb.edu).

Digital Object Identifier 10.1109/TWC.2004.842951

approximate tree search. An empirical density (KDE) [19] of subcarrier estimated powers is first computed. The KDE is then quantized into $\hat{M}$ regions using the Lloyd-Max algorithm, where $\hat{M}$ is the maximum number of paths to search in the tree. The resulting look-up table (LUT) is used to assign appropriate values of M on the basis of subsequent subcarrier power estimates.

The remainder of this paper is organized as follows. In Section II, the signal and channel models are described. The joint data detection and channel estimation algorithms are described in Section III. Simulation results are provided in Section IV, and conclusions follow in Section V. The analytic SER for the $M = 1$ QRD-M algorithm is derived in Appendix A.

II. Signal and Channel Models

A low-pass equivalent model for a received MIMO OFDM signal is considered which incorporates a quasi-static multipath fading channel, in which the channel is time varying but constant over a symbol duration. Throughout this paper, K denotes the number of subcarriers, N_t the number of transmit antennas, N_r the number of receive antennas, $(\mathbf{A})_l$ is the lth column of the matrix $\mathbf{A}$, $(\mathbf{A})_{l,m}$ is the (l,m) element of the matrix $\mathbf{A}$, and $(\mathbf{x})_l$ is the lth element of the vector $\mathbf{x}$. The indexes p and q denote transmit and receive antenna, respectively, with $1 \le p \le N_t$, $1 \le q \le N_r$. The subcarriers are indexed by $0 \le k \le K-1$. It is assumed that $N_r \ge N_t$, which is a necessary condition for the QR decomposition to yield an upper triangular matrix and hence for implementation of the M algorithm.

During OFDM symbol epoch $[nT_d^g, (n+1)T_d^g)$, a K-point inverse FFT (IFFT) is computed using QPSK or quadrature amplitude modulation (QAM) data symbols $\left\{d_0^p(n), \ldots, d_{K-1}^p(n)\right\}$. The IFFT sequence corresponding to $d_k^p(n)$ is then transmitted by antenna p. A set of symbols $\{d_k^p(n)\}$ is assumed independent in indexes k, n, and p. The independence assumption is valid for either an uncoded system or in a temporally coded system with ideal interleaving. Note that we use the same signal constellation for all subcarriers and antennas. The IFFT output of the pth transmit element in the interval $nT_d^g \le t < (n+1)T_d^g$ with guard time interval T_g is

$$s^p(t) = \frac{1}{\sqrt{K}} \times \sum_{k=0}^{K-1} d_k^p(n) e^{j2\pi k\left(t - T_g - nT_d^g\right)/T_d} p\left(t - T_g - nT_d^g\right). \quad (1)$$

In (1), $T_d^g = (K + N_g)T_s$ is the OFDM symbol interval including the guard time interval, T_s is the sampling time, and $N_g = T_g/T_s$ is the number of samples in the guard interval. The OFDM subcarrier spacing is $1/T_d$ Hz with $T_d = KT_s$, and $p(t)$ is a pulse with support $[0, T_d)$. The channel between the pth transmit and qth receive antenna is modeled by a tapped delay line (TDL) with lth coefficient $f_l^{p,q}(n) \in \mathbb{C}$ and tap spacing of T_s s. [20, Ch. 7]. It is assumed that coarse OFDM symbol synchronization has been achieved. The residual timing error can then be approximately represented by the TDL channel model, using the interpolation formula arguments in [21].

The received signal at the qth receive antenna based on the TDL channel model is

$$r^q(t) = \sum_{p=1}^{N_t} \sum_{l=0}^{N_f-1} f_l^{p,q}(n) s^p(t - lT_s) + n^q(t) \quad (2)$$

for $t \in [nT_d^g, (n+1)T_d^g)$, where N_f is the number of multipaths. The total multipath spread satisfies $N_f T_s < T_g$ in order for the time guard band to eliminate ISI. The coefficients $\{f_l^{p,q}(n)\}$ are assumed to be constant over one OFDM symbol duration but vary from symbol to symbol [22]. The additive noise $n^q(t)$ is circular white Gaussian with spectral density $2N_0$.

The received signal sampled at instances $((m + N_g)T_s + nT_d^g)$ is given by

$$\begin{aligned} r^q(mT_s) &\triangleq r^q\left((m+N_g)T_s + nT_d^g\right) \\ &= \frac{1}{\sqrt{K}} \sum_{p=1}^{N_t} \sum_{l=0}^{N_f-1} f_l^{p,q}(n) \\ &\quad \times \sum_{k=0}^{K-1} d_k^p(n) e^{j2\pi k((m-l)T_s/T_d)} + n^q(mT_s). \end{aligned} \quad (3)$$

Note that the transmitter and receiver filters are modeled as ideal low pass, with passband $[0, 1/T_s]$. The pulse $p(t)$ is approximated as ideal rectangular, since its bandwidth is much smaller than $1/T_s$ Hz.

The received baseband OFDM signal vector for the nth symbol interval is written using (3) as

$$\mathbf{r}^q(n) = \sum_{p=1}^{N_t} \sum_{l=0}^{N_f-1} f_l^{p,q}(n) \sum_{k=0}^{K-1} d_k^p(n) \mathbf{s}_k e^{-j2\pi klT_s/T_d} + \mathbf{n}^q(n) \quad (4)$$

where

$$\begin{aligned} \mathbf{r}^q(n) &\triangleq [r^q(0), r^q(T_s), \ldots, r^q((K-1)T_s)]^T \\ \mathbf{n}^q(n) &\triangleq [n^q(0), n^q(T_s), \ldots, n^q((K-1)T_s)]^T \\ &\sim \mathcal{N}\left(\mathbf{n}^q(n); \frac{\mathbf{0}, 2N_0}{T_s \mathbf{I}}\right). \end{aligned} \quad (5)$$

In (5), $\mathcal{N}(\mathbf{x}; \bar{\mathbf{x}}, \mathbf{R})$ represents a circular Gaussian density with mean vector $\bar{\mathbf{x}}$ and covariance matrix $\mathbf{R}$, and $\mathbf{s}_k$ is defined by

$$\mathbf{s}_k = \frac{1}{\sqrt{K}}[1, e^{j2\pi kT_s/T_d}, e^{j2\pi k2T_s/T_d}, \ldots, e^{j2\pi k(K-1)T_s/T_d}]^T. \quad (6)$$

The received signal vector in (4) can be alternatively expressed as

$$\mathbf{r}^q(n) = \sum_{p=1}^{N_t} \mathbf{S}\mathbf{F}^{p,q}(n)\mathbf{d}^p(n) + \mathbf{n}^q(n) \quad (7)$$

where

$$\begin{aligned} \mathbf{S} &\triangleq [\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{K-1}] \in \mathbb{C}^{K\times K} \\ \mathbf{F}^{p,q}(n) &\triangleq \operatorname{diag}\left\{F_0^{p,q}(n), F_1^{p,q}(n), \ldots F_{K-1}^{p,q}(n)\right\} \in \mathbb{C}^{K\times K} \\ F_k^{p,q}(n) &\triangleq \mathbf{c}_k^T \mathbf{f}^{p,q}(n) \\ \mathbf{c}_k &\triangleq \left[1, e^{-j2\pi kT_s/T_d}, \ldots, e^{-j2\pi k((N_f-1)T_s)/T_d}\right]^T \\ \mathbf{f}^{p,q}(n) &\triangleq \left[f_0^{p,q}(n), f_1^{p,q}(n), \ldots, f_{N_f-1}^{p,q}(n)\right]^T \in \mathbb{C}^{N_f} \\ \mathbf{d}^p(n) &\triangleq \left[d_0^p(n), d_1^p(n), \ldots, d_{K-1}^p(n)\right]^T. \end{aligned} \quad (8)$$

Note that $\mathbf{r}^q(n)$ is a sufficient statistic, $\mathbf{S}$ is the DFT matrix satisfying $\mathbf{S}^H\mathbf{S} = \mathbf{I}$, and $\mathbf{c}_k \in \mathbb{C}^{N_f}$ is the truncated FFT vector. The demodulator output (a K-point FFT) is now given by

$$\mathbf{y}^q(n) = \mathbf{S}^H\mathbf{r}^q(n) = \sum_{p=1}^{N_t} \mathbf{F}^{p,q}(n)\mathbf{d}^p(n) + \mathbf{z}^q(n) \tag{9}$$

where $\mathbf{z}^q(n) \triangleq \mathbf{S}^H\mathbf{n}^q(n) \sim \mathcal{N}(\mathbf{z}^q(n); \mathbf{0}, 2N_0/T_s\mathbf{I})$.

The overall goal of this paper is to find an effective strategy to detect $\mathbf{d}^p(n)$ and estimate $\{\mathbf{f}^{p,q}(n)\}$ from the received samples. It is observed that conditioned on the channel vectors $\{\mathbf{f}^{p,q}(n)\}$, the detection of the data symbols $d_k^p(n)$ in (9) is separable in the subcarriers, since the components of the additive noise $\mathbf{z}^q(n)$ are independent. This will result in a greatly simplified QRD-M algorithm.

The following equation for the kth received subcarrier is obtained using the definition of the matrix $\mathbf{F}^{p,q}(n)$ in (8)

$$\begin{aligned}
y_k^q(n) &\triangleq (\mathbf{y}^q(n))_k = \sum_{p=1}^{N_t} F_k^{p,q}(n)d_k^p(n) + z_k^q(n)\\
\mathbf{y}_k(n) &\triangleq \left[y_k^1(n), y_k^2(n), \ldots, y_k^{N_r}(n)\right]^T\\
&= \mathbf{F}_k(n)\mathbf{d}_k(n) + \mathbf{z}_k(n)\\
\mathbf{F}_k(n) &\triangleq \begin{bmatrix} F_k^{1,1}(n) & F_k^{2,1}(n) & \ldots & F_k^{N_t,1}(n)\\ F_k^{1,2}(n) & F_k^{2,2}(n) & \ldots & F_k^{N_t,2}(n)\\ \vdots & \vdots & \cdots & \vdots\\ F_k^{1,N_r}(n) & F_k^{2,N_r}(n) & \ldots & F_k^{N_t,N_r}(n)\end{bmatrix}\\
\mathbf{d}_k(n) &\triangleq \left[d_k^1(n), d_k^2(n), \ldots, d_k^{N_t}(n)\right]^T\\
\mathbf{z}_k(n) &\sim \mathcal{N}\left(\mathbf{z}_k(n); \mathbf{0}, \frac{2N_0}{T_s}\mathbf{I}\right).
\end{aligned} \tag{10}$$

In the definition (10), $\mathbf{F}_k(n)$ represents frequency responses of all $N_t \times N_r$ channels at FFT frequency k.

The ML data detection for the kth subcarrier is performed using channel one-step predictions obtained from a KF

$$\hat{\mathbf{d}}_k(n)_{\mathrm{ML}} = \arg \min_{\mathbf{d}_k(n)\in\{\mathcal{S}\}^{N_t}} \|\mathbf{y}_k(n) - \hat{\mathbf{F}}_k(n|n-1)\mathbf{d}_k(n)\|^2 \tag{11}$$

where $\hat{\mathbf{F}}_k(n|n-1) \in \mathbb{C}^{N_r\times N_t}$ is the one-step KF prediction of the effective MIMO channel frequency response matrix and $\mathcal{S}$ is the symbol alphabet. The ML decision for the data $\mathbf{d}_k(n)$ (11) is separable in the subcarriers k when conditioned on the channel estimate. The separability follows from (8), where $\mathbf{F}^{p,q}(n)$ is diagonal.

As is well known, the computational complexity of (11) grows exponentially with the number of transmit antennas and the cardinality of the subcarrier modulation. To reduce this prohibitive complexity, the suboptimal M algorithm for detection is employed next.

III. New Joint Detection and Channel Estimation Algorithms

A. Joint Detection Algorithm

The QR decomposition [14] approach proposed in [13] for DS-CDMA systems is applied here to the estimated channel matrix $\hat{\mathbf{F}}_k(n) = \hat{\mathbf{F}}_k(n|n-1)$, and the M algorithm is used to efficiently approximate the ML decisions on $\mathbf{d}_k(n)$. Given the channel one-step prediction, the received signal power of the ith data symbol $d_k^i(n)$ is defined by $\hat{P}_2^i \triangleq \|(\hat{\mathbf{F}}_k(n))_i\|_2^2$, where the L_2 norm is $\|\mathbf{x}\|^2 = \mathbf{x}^H\mathbf{x}$.

Next, rearrange the channel estimates using the order statistics of the estimated powers as follows:

$$\hat{\mathbf{F}}_k(n) = [(\hat{\mathbf{F}}_k(n))_{(1)}, (\hat{\mathbf{F}}_k(n))_{(2)}, \ldots, (\hat{\mathbf{F}}_k(n))_{(N_t)}] \tag{12}$$

such that $\hat{P}_2^{(1)} \le \hat{P}_2^{(2)} \le \cdots \le \hat{P}_2^{(N_t)}$. The corresponding data vector $\mathbf{d}_k(n)$ is also reordered in terms of the corresponding powers as $\mathbf{d}_k(n) = \left[d_k^{(1)}(n), d_k^{(2)}(n), \ldots, d_k^{(N_t)}(n)\right]^T$. Note that in the interference cancellation structure, the detection order is crucial to the performance of the system [12], [23], and the estimated channel matrix is thus rearranged so that data $d_k^{(N_t)}$ will be detected first. In [4], a similar detection order strategy is motivated by minimizing the bit-error rate (BER) of the data associated with the weakest vector channel.

The following modified ML cost function is obtained by applying the QRD in (11). Let $\mathbf{Q}_k(n)$ be the unique unitary matrix such that $\hat{\mathbf{F}}_k(n) = \mathbf{Q}_k(n)\mathbf{R}_k(n)$, such that $\mathbf{R}_k(n)$ is upper triangular. The detailed structure of $\mathbf{R}_k(n)$ is

$$\mathbf{Q}_k(n)^H\hat{\mathbf{F}}_k(n) \triangleq \mathbf{R}_k(n) = \begin{bmatrix}\mathbf{R}_{N_t\times N_t}\\ \mathbf{0}_{(N_r-N_t)\times N_t}\end{bmatrix}. \tag{13}$$

Define $\mathbf{y}_k(n) = \mathbf{Q}_k(n)^H\mathbf{y}_k(n)$. Then

$$\mathbf{y}_k(n) = \mathbf{R}_k(n)\mathbf{d}_k(n) + \mathbf{v}_k(n) \tag{14}$$

where $\mathbf{v}_k(n) = \mathbf{Q}_k(n)^H\mathbf{z}_k(n)$. Note that the statistics of the noise vector $\mathbf{v}_k(n)$ are invariant under the unitary transformation $\mathbf{Q}_k(n)^H\mathbf{z}_k(n)$. In particular, the components of $\mathbf{v}_k(n)$ are independent identically distributed (i.i.d.) Gaussian, which will facilitate the closed-form BER analysis for the case $M = 1$ in Appendix A. Then, the ML decision (11) becomes (15), as shown at the bottom of the page. Recall that $N_r \ge N_t$ in (15).

$$\begin{aligned}
\hat{\mathbf{d}}_k(n)_{\mathrm{ML}} &= \arg \min_{\mathbf{d}_k(n)\in\{\mathcal{S}\}^{N_t}} \|\mathbf{y}_k(n) - \mathbf{Q}_k(n)^H\hat{\mathbf{F}}_k(n)\mathbf{d}_k(n)\|^2\\
&= \arg \min_{\mathbf{d}_k(n)\in\{\mathcal{S}\}^{N_t}} \Bigg\{ \left((\mathbf{y}_k(n))_{N_t} - (\mathbf{R}_k(n))_{N_t,N_t}d_k^{(N_t)}(n)\right)^2 + \cdots ((\mathbf{y}_k(n))_1 - (\mathbf{R}_k(n))_{1,1}d_k^{(1)}(n)\cdots\\
&\qquad - (\mathbf{R}_k(n))_{1,N_t}d_k^{(N_t)}(n))^2 + (\mathbf{y}_k(n))_{N_t+1}^2 + (\mathbf{y}_k(n))_{N_t+2}^2 + \cdots + (\mathbf{y}_k(n))_{N_r}^2 \Bigg\}
\end{aligned} \tag{15}$$

The ML cost function can be simplified by deleting the terms in (15) that do not depend on the data $\mathbf{d}_k(n)$. The cost function is then written in terms of states and branch metrics as follows:

$$\hat{\mathbf{d}}_k(n)_{\mathrm{ML}} = \arg \min_{\mathbf{d}_k(n)\in\{\mathcal{S}\}^{N_t}} \sum_{p=1}^{N_t} m_p\left((\mathbf{y}_k(n))_{(p)}; s_{(p)}\right) \quad (16)$$

where the metric $m_p\left((\mathbf{y}_k(n))_{(p)}; s_{(p)}\right)$ and the state $s_{(p)}$ are, respectively, defined by

$$\begin{aligned} & m_p\left((\mathbf{y}_k(n))_{(p)}; s_{(p)}\right) \quad (17)\\ & \quad \triangleq \Bigg\| (\mathbf{y}_k(n))_{N_t-p+1} \\ & \qquad - (\mathbf{R}_k(n))_{N_t-p+1,N_t-p+1} d_k^{(N_t-p+1)}(n) \\ & \qquad - \sum_{l=N_t-p+2}^{N_t} (\mathbf{R}_k(n))_{N_t-p+1,l} d_k^{(l)}(n) \Bigg\|^2 \\ & s_{(p)} \\ & \quad \triangleq \left\{ d_k^{(N_t-p+1)}(n), d_k^{(N_t-p+2)}(n), \ldots, d_k^{(N_t)}(n) \right\}. \quad (18) \end{aligned}$$

Note that at stage p, the state and the metric are determined by a set of data symbols corresponding to antenna $(N_t - p + 1)$ through (N_t).

Using these states and metrics, the ML detector can be equivalently represented as a tree search beginning at level (1) and concluding at level (N_t). A state in the tree is specified by

$$s_{(p)}(i_1, i_2, \ldots, i_p), \quad 0 \le i_j \le |\mathcal{S}| - 1, \ \ j = 1, \ldots, p. \quad (19)$$

However, the complexity of the tree search at the final stage (N_t) is $\mathcal{O}(|S|^{N_t})$, which is impractical for even moderate values of N_t. The M algorithm simply retains only the paths through the tree with the M smallest accumulated metrics [15]–[17]. It is worthwhile to note that the structure of the QRD-M algorithm with $M = 1$ is very similar to an interference canceler or a decision feedback detector. Furthermore, the structure of the QR decomposition permits a closed-form BER computation for QPSK signaling in the $M = 1$ case, as shown in Appendix A.

The authors in [13] show that the QRD-M algorithm has a better BER performance than an interference cancellation $(M = 1)$ method for the case of DS-CDMA. Similar results will be shown in Section IV for the MIMO-OFDM system. An example of the QRD-M algorithm is presented as follows to clarify the approach.

Example: Fig. 1 shows the tree for $N_t = 2$ transmit antennas and QPSK modulation on all subcarriers. The data associated with the strongest channel $d_k^{(N_t)}(n)$ corresponds to the root node in the tree diagram, and there are four branches denoted by $\{s_{(1)}(i_1)|i_1 = 0,1,2,3\}$ with four metrics $\{m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(i_1))|i_1 = 0,1,2,3\}$ from the root node. For the next strongest data, four metrics exist for each branch. That is, $\{s_{(2)}(i_1, i_2)|i_1 = 0,1,2,3, i_2 = 0,1,2,3\}$ and the metrics $\{m_2((\mathbf{y}_k(n))_{(2)}; s_{(2)}(i_1, i_2))|i_1 = 0,1,2,3, i_2 = 0,1,2,3\}$. We can readily modify these functions to accommodate higher order constellations such as $\{16, 64\}$-QAM.

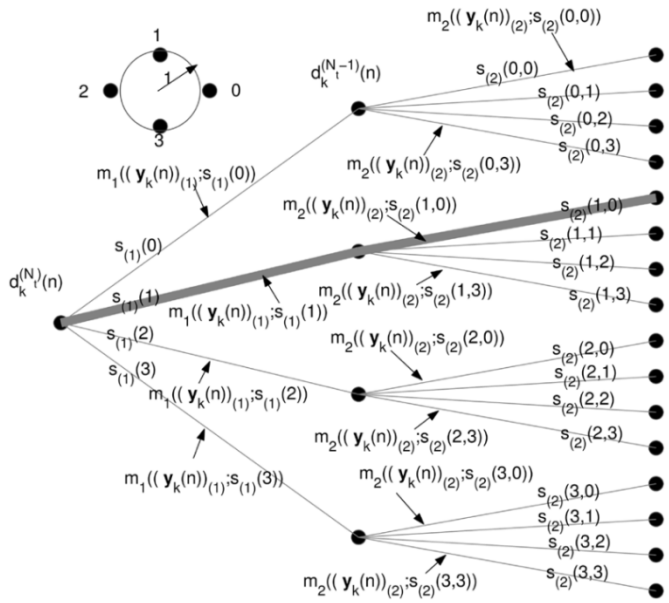

Fig. 1. Tree diagram showing $M = 1$ algorithm decisions with $N_t = 2$ and QPSK modulation.

Consider the case $M = 1$ with QPSK modulation. The QPSK constellation is rotated by $\pi/4$ to simplify the notation. Since we retain only one accumulated distance metric at each level of the tree, for the strongest data $d_k^{(N_t)}(n)$, only the metric $m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(1))$ and the state $s_{(1)}(1)$ are retained for the next data detection. This path corresponds to the signal point $S_1 = j$ in the signal constellation. For the next strongest data $d_k^{(N_t-1)}(n)$, first compute all possible accumulated branch metrics adding both the previous metric $m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(1))$ and the current metric, resulting in the sums

$$\begin{aligned} & m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(1)) + m_2((\mathbf{y}_k(n))_{(2)}; s_{(2)}(1,0)), \\ & m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(1)) + m_2((\mathbf{y}_k(n))_{(2)}; s_{(2)}(1,1)), \\ & m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(1)) + m_2((\mathbf{y}_k(n))_{(2)}; s_{(2)}(1,2)), \\ & m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(1)) + m_2((\mathbf{y}_k(n))_{(2)}; s_{(2)}(1,3)) \end{aligned}$$

and then sort these accumulated distance metrics.

Suppose that $m_1((\mathbf{y}_k(n))_{(1)}; s_{(1)}(1)) + m_2((\mathbf{y}_k(n))_{(2)}; s_{(2)}(1,0))]$ is the minimum in Fig. 1, then we can extend the survivor path at this stage. This survivor path specifies the decision $\hat{d}_k^{(N_t)} = S_1 = j$ and $\hat{d}_k^{(N_t-1)} = S_0 = 1$.

The computational efficiency of the QRD-M algorithm compared with ML is illustrated in Table I for 16-QAM. For example, the number of complex multiplies for $M = 16$ is only 11% of the number required for the full ML search, and similar savings are evident for the remaining arithmetic operations.

B. Channel Estimation

The nominal Doppler spread normalized to the OFDM symbol rate considered here is $f_d T_d \approx 10^{-3}$. Hence, it is reasonable to model the channel coefficient vectors $\mathbf{f}^{p,q}(n)$ as slowly varying autoregressive processes. A decision-directed

TABLE I
COMPARISON OF THE NUMBER OF FUNDAMENTAL OPERATIONS OF THE QRD-M AND THE ML ON A SINGLE SUBCARRIER FOR THE 16-QAM MODULATION WITH $N_t = 4$, $N_r = 4$

Number of operations	M=1	M=2	M=4	M=8	M=16	ML
Complex addition	67	115	211	403	787	69907
Complex subtraction	256	448	832	1600	3136	279616
Real multiplication	320	560	1040	2000	3920	349520
Complex multiplication	454	790	1462	2806	5494	489334

KF channel estimator is proposed here based on the following representation of the received signal at antenna q:

$$y_k^q(n) = (\mathbf{y}^q(n))_k = \sum_{p=1}^{N_t} \mathbf{c}_k^T \mathbf{f}^{p,q}(n) d_k^p(n) + z_k^q(n) \qquad (20)$$

where $\mathbf{c}_k$ is the truncated FFT vector defined in (8). The unknown data $d_k^p(n)$ are replaced by decisions from the M algorithm $\hat{d}_k^p(n)$. At time n, the M algorithm uses predicted channel values $\hat{\mathbf{f}}^{p,q}(n|n-1)$ from the KF to compute $\hat{d}_k^p(n)$.

The channel estimation problem is separable in the N_r receive antennas, since the $\mathbf{f}^{p,q}(n)$ are assumed to evolve independently, and the thermal noises at the antennas $z_k^q(n)$ are independent white Gaussian processes. However, all K subcarrier measurements $y_k^q(n)$, $k = 0, 1, \ldots, K-1$ depend on the same set of channel variables $\mathbf{f}^{p,q}(n)$; hence, N_r separate KFs are applied to the composite measurement vectors $\mathbf{y}^q(n) \in \mathcal{C}^K$. These vectors are then approximated using M-algorithm decisions by

$$\mathbf{y}^q(n) = \mathbf{H}(\hat{\mathbf{d}}(n))\mathbf{x}^q(n) + \mathbf{z}^q(n)$$

where the state vector $\mathbf{x}^q(n) \in \mathcal{C}^{N_t N_f}$ is defined by $\mathbf{x}^q(n) = [\mathbf{f}^{1,q}(n)^T, \ldots, \mathbf{f}^{N_t,q}(n)^T]^T$, and the measurement function $\mathbf{H}(\hat{\mathbf{d}}(n))$ is defined by

$$\begin{aligned}\mathbf{H}^q(n) &\triangleq \mathbf{H}(\hat{\mathbf{d}}(n)) \\ &= \left[\hat{\mathbf{D}}^1(n)\mathbf{C}^T, \hat{\mathbf{D}}^2(n)\mathbf{C}^T, \ldots, \hat{\mathbf{D}}^{N_t}(n)\mathbf{C}^T\right] \\ &\in \mathbb{C}^{K \times N_f N_t}. \end{aligned} \qquad (21)$$

In (21), $\mathbf{C} \in \mathcal{C}^{N_f \times K}$ is the truncated FFT matrix defined by $\mathbf{C} = [\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{K-1}]$. Also

$$\hat{\mathbf{D}}^p(n) = \operatorname{diag}\left\{\hat{d}_0^p(n), \hat{d}_1^p(n), \ldots, \hat{d}_{K-1}^p(n)\right\}. \qquad (22)$$

To derive the KF, an autoregressive model (AR) for the channel coefficients is assumed. The multipath coefficient vectors are assumed to be independent AR processes following

$$\mathbf{f}^{p,q}(n) = \mathbf{F}_f \mathbf{f}^{p,q}(n-1) + \mathbf{w}_f^{p,q}(n)$$

where $\mathbf{F}_f \in \mathcal{C}^{N_f \times N_f}$ is typically diagonal with nonzero elements computed according to a nominal Doppler spread [18]. The noise $\mathbf{w}_f^{p,q}(n)$ is independent circular Gaussian with variances $\mathcal{Q}_f$. In terms of the state vector, the AR model is then

$$\mathbf{x}^q(n) = \mathbf{F}_x \mathbf{x}^q(n-1) + \mathbf{w}_x^q(n) \qquad (23)$$

TABLE II
QRD-M-KF CHANNEL ESTIMATION ALGORITHM

Given $\{\hat{\mathbf{f}}^q(n|n-1)\}$
  Use QRD-M algorithm to compute decisions $\{\hat{d}_k^p(n)\}$
  (Use training symbols $\mathbf{d}^q(0)$ if $n = 0$)
For $q = 1, 2, \ldots, N_r$
  Compute measurement and KF updates
    $\mathbf{P}^q(n|n)^{-1} = \mathbf{P}^q(n|n-1)^{-1} + \frac{1}{\sigma_n^2}\mathbf{H}^q(n)^H\mathbf{H}^q(n)$
    $\hat{\mathbf{x}}^q(n|n) = \hat{\mathbf{x}}^q(n|n-1) + \mathbf{P}^q(n|n)\mathbf{H}^q(n)^H\frac{1}{\sigma_n^2}[\mathbf{y}^q(n) - \mathbf{H}^q(n)\hat{\mathbf{x}}^q(n|n)]$
  Compute one-step predictions
    $\mathbf{P}^q(n+1|n) = \mathsf{F}_x\mathbf{P}^q(n|n)\mathsf{F}_x^H + \mathcal{Q}_x$
    $\hat{\mathbf{x}}^q(n+1|n) = \mathsf{F}_x\hat{\mathbf{x}}^q(n|n)$
Next antenna q

where the one-step transition matrix and process noise $(\mathbf{w}_x^q(n))$ covariance, $\mathbf{F}_x$, $\mathcal{Q}_x \in \mathcal{C}^{N_t(N_f) \times N_t(N_f)}$, are

$$\begin{aligned}\mathbf{F}_x &= \text{ block diag}\{\mathbf{F}_f, \ldots, \mathbf{F}_f\} \\ \mathcal{Q}_x &= \text{ block diag}\{\mathcal{Q}_f, \ldots, \mathcal{Q}_f\}. \end{aligned} \qquad (24)$$

The complete QRD-M-KF algorithm is given in Table II. In practice, only one initial OFDM training symbol $\mathbf{d}(0)$ comprising KN_t QAM symbols was found to be required for convergence of the KF. The initial channel state vector $\hat{\mathbf{x}}^q(0|0)$ was set to $2 \times \mathbf{1}$, where $\mathbf{1}$ is the all-ones vector, and the covariance was initialized to $\mathbf{P}^q(n) = \mathbf{I}$ for all q.

C. Adaptive Complexity AC-QRD-M Algorithm

Due to frequency-selective fading, subcarriers in the OFDM system will have widely varying signal-to-noise ratios (SNRs). Recall that a separate QRD-M algorithm runs independently on each subcarrier. Thus, it is intuitively reasonable to use large values of M for those subcarriers k with low SNR and small values, e.g., $M = 1$, for those with high SNR. This observation leads us to the AC-QRD-M algorithm which maps estimated channel power to M. Since detector performance at the first stage of the M algorithm depends directly on $|(\mathbf{R}_k(n))_{N_t,N_t}|$, it is reasonable to use the metric $\Delta_R^{N_t} \triangleq |(\mathbf{R}_k(n))_{N_t,N_t}|^2$ for selecting M. Alternatively, the norm $\Delta_R \triangleq \sum_{i=1}^{N_t} |(\mathbf{R}_k(n))_{i,i}|^2$ will also be used to determine M.

It is not clear at first how to set M given a single observed value of $\Delta_R^{N_t}$ or Δ_R. The approach taken here is to use multiple observations of $\Delta_R^{N_t}$ to estimate the probability density function of this metric. The Lloyd-Max algorithm is then used to partition the pdf into $\hat{M}$ regions. Specifically, $\hat{M}$ corresponds to the maximum value of M which will be employed, determined for example by available computational resources.

For a set of training symbol sequences, the QRD is applied to the channel matrix, and then the KDE [19] forms the pdf for either $\Delta_R^{N_t}$ or Δ_R. Let $\left(\Delta_R^{N_t}\right)_i$ be the ith observation for $\Delta_R^{N_t}$ and N the total number of observations. The KDE is given by

$$\hat{f}_{\Delta_R^{N_t}}(r) = \frac{1}{N}\sum_{i=1}^{N} \mathcal{G}_h\left(r - \left(\Delta_R^{N_t}\right)_i\right) \qquad (25)$$

where $\mathcal{G}_h(r) = \mathcal{N}(\mathcal{G}(r); 0, h^2)$. The optimum width parameter h in terms of minimizing asymptotic mean squared error is $\hat{h} =$

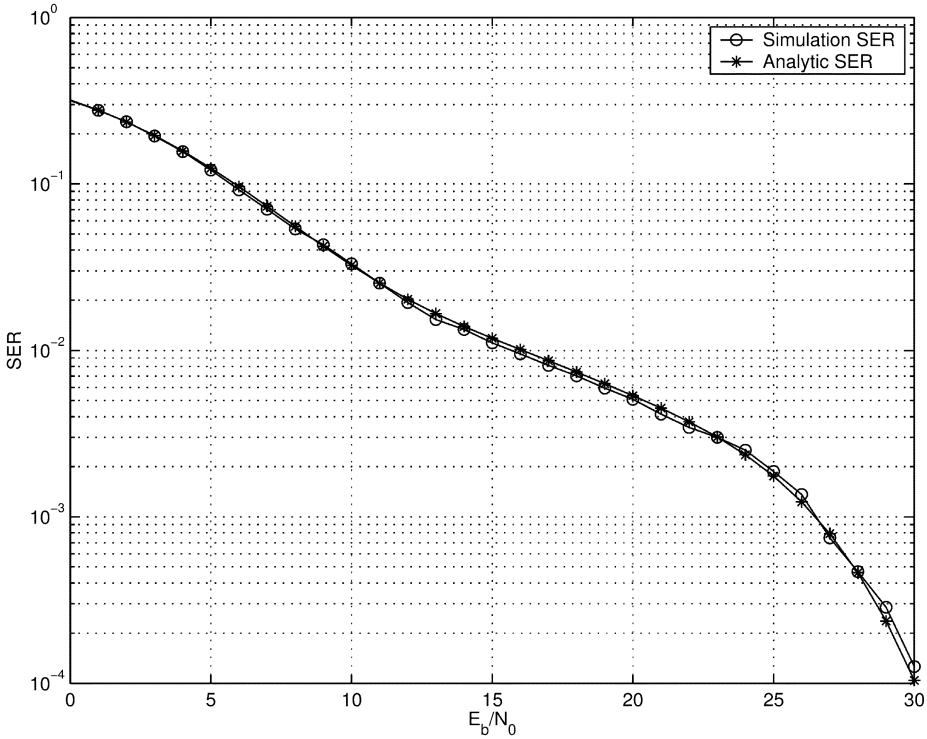

Fig. 2. Comparison between the analytic and simulated SER with $N_t = 2$, $N_r = 2$, $K = 64$, known $N_f = 5$ length channel, $\mathbf{f}^{p,q}(n) = [0.749, 0.502, 0.3365, 0.2256, 0.1512]^T$, $\forall\, p, q$, QPSK subcarrier modulation.

$1.06 \min(\hat{\sigma}, \mathrm{IQR}(r)/1.34) \times N^{-0.2}$ [19]. The parameter $\hat{\sigma}^2$ is the sample variance defined as

$$\hat{\sigma}^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N}\left(\left(\Delta_R^{N_t}\right)_i\right)^2 - N\left(\frac{1}{N}\sum_{i=1}^{N}\left(\Delta_R^{N_t}\right)_i\right)^2\right] \tag{26}$$

and IQR is the distance between the upper quartile, 75th percentile, and the lower quartile, 25th percentile.

In order to assign values of M based on the observed $\Delta_R^{N_t}$, a partitioning strategy for the pdf is required. The Lloyd-Max optimization procedure [24] selects thresholds th_l to minimize the following quantization variance:

$$\min \sigma_q^2 \triangleq \min \sum_{l=1}^{\hat{M}} \int_{\mathrm{th}_l}^{\mathrm{th}_{l+1}} (r - y_l)^2 \hat{f}_{\Delta_R^{N_t}}(r) dr \tag{27}$$

for $\{\mathrm{th}_l, l = 2, 3, \ldots, \hat{M}\}$ and centroids $\{y_l, l = 1, 2, \ldots, \hat{M}\}$. The centroids are defined by [24] $y_l = \int_{\mathrm{th}_l}^{\mathrm{th}_{l+1}} r \hat{f}_{\Delta_R^{N_t}}(r) dr / \int_{\mathrm{th}_l}^{\mathrm{th}_{l+1}} \hat{f}_{\Delta_R^{N_t}}(r) dr$. The LUT is then defined by the following mappings of observed $\Delta_{R_k}^{N_t}$ to M

$$\begin{aligned}
&\Delta_{R_k}^{N_t} \in [0, \mathrm{th}_1) : M = \hat{M}, \\
&\Delta_{R_k}^{N_t} \in [\mathrm{th}_1, \mathrm{th}_2) : M = \hat{M} - 1, \ldots, \\
&\Delta_{R_k}^{N_t} \in [\mathrm{th}_{\hat{M}-1}\ \mathrm{th}_{\hat{M}}) : M = 1.
\end{aligned}$$

In all simulations, $N = 1000$ training composite OFDM symbols were used to form the KDE.

IV. Simulation Results

To verify the overall simulation approach and analytic SER for $M = 1$, the case of QPSK with $N_r = N_t = 2$ antennas was considered. Note that the analytic SER is conditioned on a time-invariant channel realization. Fig. 2 shows that the analytic SER computed using Appendix A closely matches the simulated SER, for the case of a time-invariant $N_f = 5$ tap channel and $K = 64$ subcarriers.

In subsequent simulations, QPSK or 16-QAM was used as the subcarrier modulation. The number of antennas in the transmitter and the receiver was set to $N_t = N_r = 4$. The number of multipaths is $N_f = 2$, with equal power strength. The packet size is one OFDM symbol, consisting of $K = 64$ subcarriers. The nominal Doppler spreads are $f_d T_d \approx 0.001/0.0032$, that is, $\mathbf{F}_f = 0.999\,99\mathbf{I}/0.999\mathbf{I}$, and $\mathbf{w}_f^{p,q}(n) \sim \mathcal{N}\left(\mathbf{w}_f^{p,q}(n); \mathbf{0}, 10^{-4}\mathbf{I}\right)$. We consider two cases.

[Case 1] Clairvoyant QRD-M (Known Channel): In this case, Fig. 3 shows the BER and packet-error rate (PER) performance for 16-QAM subcarrier modulation and different values of M. In the clairvoyant algorithm, only the QRD-M is run, that is, the channel is time varying but the true channel coefficients $\mathbf{f}^{p,q}(n)$ are used in place of KF estimates. When $M = 1$, the BER performance of the clairvoyant QRD-M system is slightly worse than that of the BLAST system. However, as the value of M increases, the BER becomes significantly lower than that of BLAST. For example, at a BER of 2×10^{-3}, the QRD-M ($M = 8$) system yields 5-dB gain over the BLAST system. QRD-M with $M = 16$ yields performance very close to that of the ML algorithm. The advantage of QRD-M is quite clear in that optimum ML detection for 16-QAM, $N_t = 4$ requires computation of 65 536 metrics. In contrast, QRD-M with $M = 16$ requires a total of 784 branch metric computations in the $N_t = 4$ level tree. The PER performance is also significantly improved compared with the BLAST system. Fig. 3 suggests that the advantage of QRD-M over BLAST

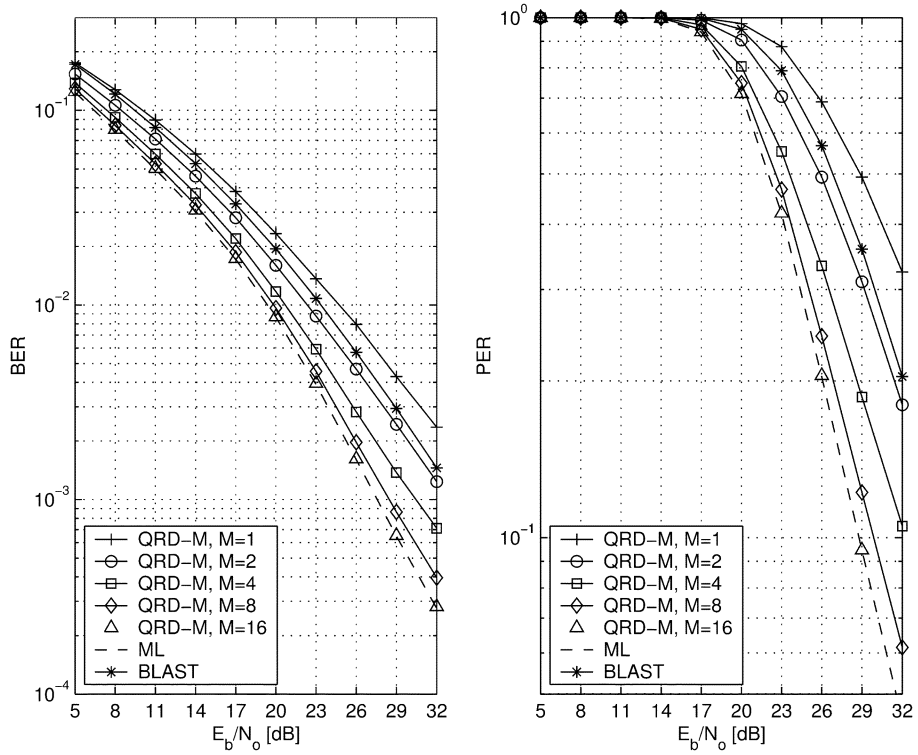

Fig. 3. BER/PER performances of the QRD-M algorithm [16-QAM, one OFDM symbols, $N_t = 4$, $N_r = 4$, $N_f = 2$, $f_d T_d = 0.001$].

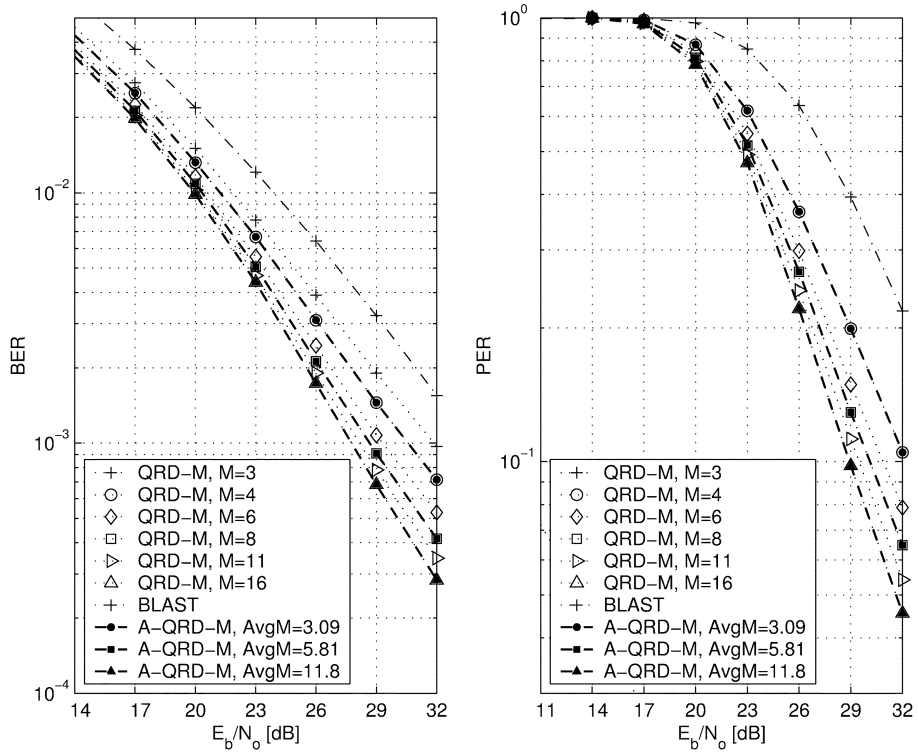

Fig. 4. BER/PER performances of the QRD-M system versus AC-QRD-M system [16-QAM, one OFDM symbols, $N_t = 4, N_r = 4, N_f = 2, f_d T_d = 0.0032$].

becomes very significant for higher order modulations and larger numbers of antennas.

The comparative performance of AC-QRD-M and QRD-M is shown in Fig. 4. These versions of the algorithms are again clairvoyant with time-varying coefficients $\mathbf{f}^p(n)$ known for all n. The normalized Doppler spread is $f_d T_d \approx 0.032$. The values of $\hat{M}$ (maximum number of paths used for each carrier) are 4/8/16. The corresponding average values of M, denoted by $\bar{M}$, from the AC-QRD-M algorithm are 3.09/5.81/11.8. The improvement over BLAST at a 10^{-3} BER is 1 dB for AC-QRD-M with $\hat{M} = 4$. Furthermore, the performance of AC-QRD-M with $\hat{M} = 16$, $\bar{M} = 11.8$ is almost identical to that of the

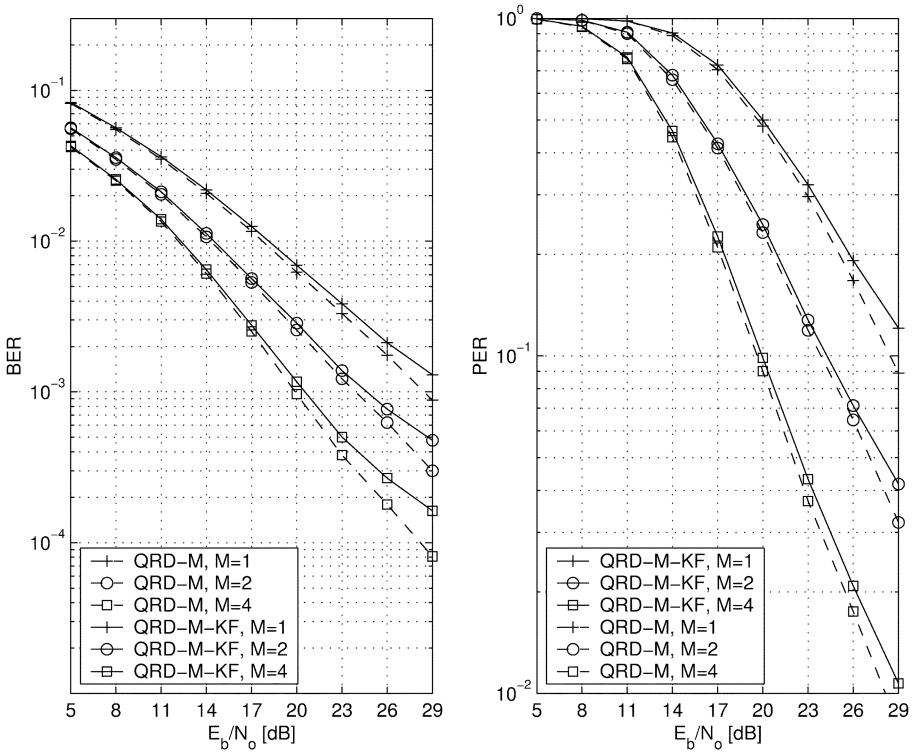

Fig. 5. BER/PER performances of the QRD-M-KF system versus QRD-M system [QPSK, one OFDM symbols, $N_t = 4$, $N_r = 4$, $N_f = 2$, $f_d T_d = 0.001$].

QRD-M system with fixed complexity $M = 16$. Thus, the average complexity of AC-QRD-M in terms of the number of branch metrics computed is significantly less than that of the fixed complexity QRD-M algorithm.

[Case 2] KF Based Channel Estimation: The impact on BER/PER performances for varying Doppler spread is evaluated using QRD-M with KF channel estimation. The nominal Doppler spreads considered are $f_d T_d \approx 10^{-3}$, $\mathbf{F}_f = 0.999\,99\mathbf{I}$, and $f_d T_d \approx 0.032$, $\mathbf{F}_f = 0.9999\mathbf{I}$. Also, $\mathbf{w}_f^{p,q}(n) \sim \mathcal{N}\left(\mathbf{w}_f^{p,q}(n); \mathbf{0}, 10^{-4}\mathbf{I}\right)$. Fig. 5 is the result for the QRD-M system combined with the KF-based channel estimator at $f_d T_d \approx 0.001$. Compared with the *a priori* known channel case, a clear performance degradation occurs as shown in Fig. 5. For example, at 10^{-3} BER, QPSK subcarrier modulation, $N_t = 4$, and one OFDM symbol, there is about a 1-dB loss compared with a clairvoyant (i.e., known channel) QRD-M algorithm-based receiver.

Fig. 6 compares QRD-M-KF for the unknown channel with clairvoyant QRD-M for 16-QAM. The channel is time varying with $f_d T_d = .001$. From this figure, we observe the same trends as in Fig. 3 with approximately a 1-dB loss in BER performance using the Kalman channel estimator.

In Figs. 7 and 8, the performance of the QRD-M-KF with 16-QAM subcarrier modulation is evaluated at $f_d T_d \approx 0.001/0.0032$ and compared with clairvoyant QRD-M. As expected, the larger Doppler spread leads to a performance degradation. However, the performance degradation from .001 to .0032 Doppler spread is still less than 1 dB for the QRD-M-KF.

V. Conclusion

A new computationally efficient joint detection and channel estimation algorithm for MIMO-OFDM system was proposed, which combines the QRD-M algorithm with Kalman channel estimator. The performance was evaluated by analysis for $M = 1$ and by simulation for larger values of M. It was shown that the QRD-M and QRD-M-KF significantly outperform the BLAST system except for the $M = 1$ case. Furthermore, significantly less computation is incurred by the M algorithm for higher order signal constellations and larger numbers of transmit antennas. Computational efficiency was also improved by using an adaptive complexity AC-QRD-M algorithm, which assigns different values of M to subcarriers depending on their estimated channel power. The QRD-M-KF algorithm appears to be robust even at large normalized Doppler spreads and hence may be a good candidate for implementation in MIMO-OFDM systems.

Appendix
Analytic Conditional SER for $M = 1$

The analytic SER is derived for $M = 1$ conditioned on a static channel realization. For convenience, reorder the indexes in the QRD-M algorithm and represent the pth element of the vector $\mathbf{y}$ as

$$\mathbf{y}_p = R_{p,p} d_p + \sum_{l=p+1}^{N_t} R_{p,l} d_l + v_p \tag{A.1}$$

where $\{v_p\}$ are i.i.d. circular Gaussian, zero mean, and with variance $\sigma_n^2 = 2N_0/T_s$.

Consider first the case of BPSK, with $d_p = \pm 1$, where the $\{d_l\}$ are assumed i.i.d. either through uncoded modulation or perfect interleaving. Although $R_{p,p} \in \mathbb{R}$ is real valued, there is a π rad. phase ambiguity unless training data is used to estimate the channel. Thus for known channels, or when training is feasible, it is assumed that the decision variable is premultiplied to yield $m'_p = \mathrm{Re}\{\mathrm{sgn}(R_{p,p}) m_p\}$. For the case $M = 1$, the

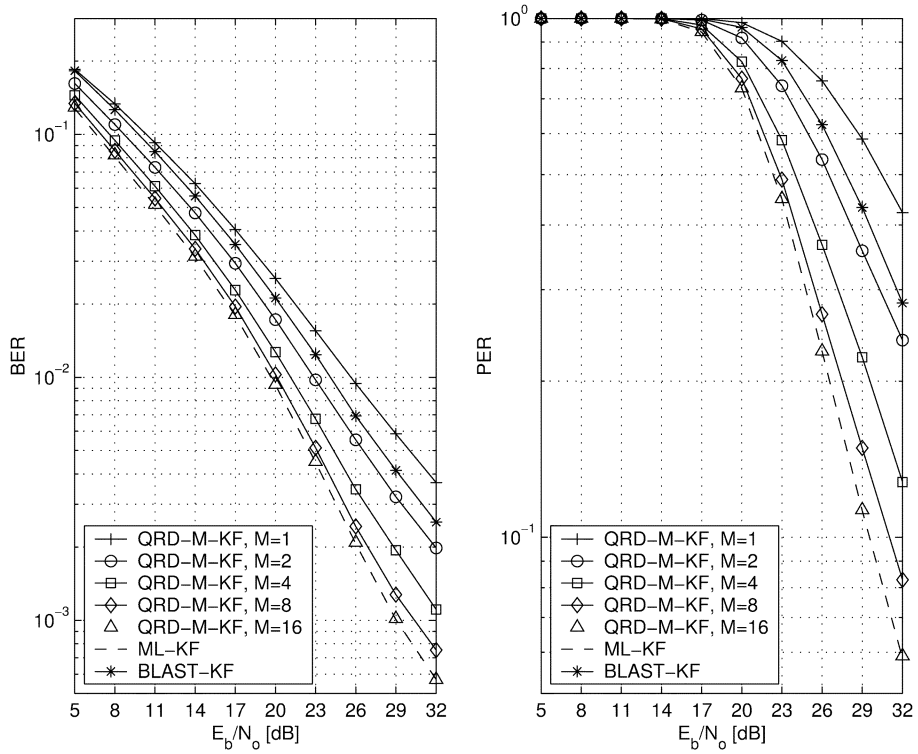

Fig. 6. BER/PER performances of the QRD-M-KF system [16-QAM, one OFDM symbols, $N_t = 4, N_r = 4, N_f = 2, f_dT_d = 0.001$].

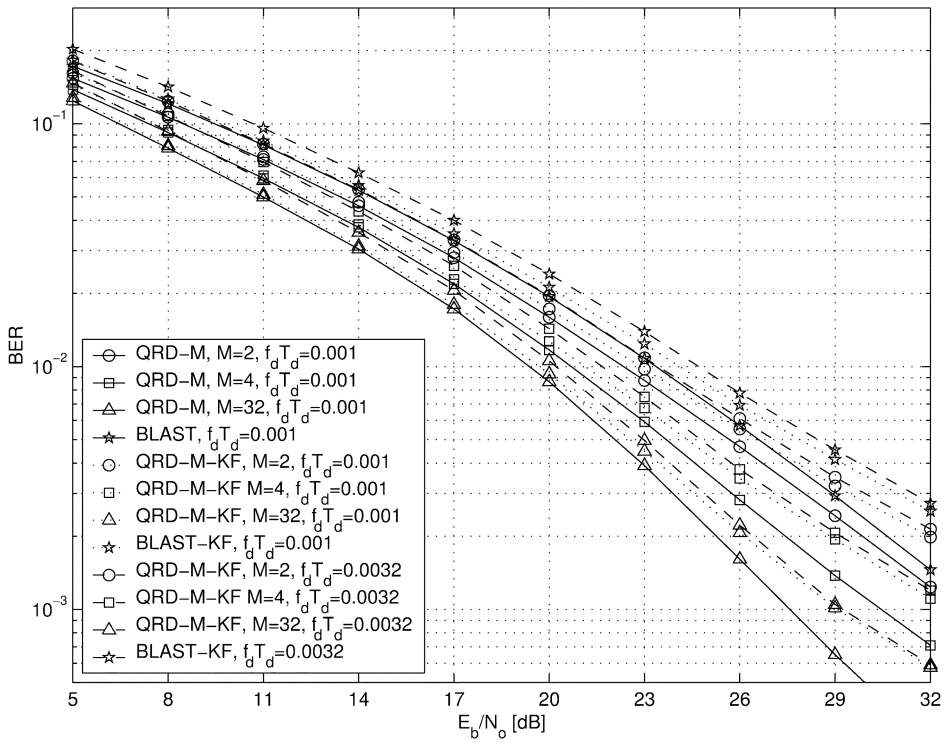

Fig. 7. BER performance of the QRD-M-KF system versus QRD-M system [16-QAM, one OFDM symbols, $N_t = 4, N_r = 4, N_f = 2, f_dT_d = 0.001/0.0032$].

phase-corrected branch metric (17) can then be written in simplified form as

$$m'_p = |R_{p,p}|d_p + \sum_{l=p+1}^{N_t} \mathrm{Re}\{R'_{p,l}\}(d_l - \hat{d}_l) + n_p \qquad \text{(A.2)}$$

where $R'_{p,l} = \mathrm{sgn}(R_{p,p})R_{p,l}$ and $n_p = \mathrm{Re}\{v_p\}$.

Define an error sequence $e_l = d_l - \hat{d}_l \in \{-2, 0, 2\}$ and a conditional probability mass function $P(e_l|e_{l+1}, \ldots, e_{N_t})$. Note that the noise terms n_p are i.i.d. Gaussian with variance N_0/T_s, and thus the branch metric m_p is conditionally Gaussian given d_p and $\{e_p, \ldots, e_{N_t}\}$ from (A.2). The conditional error

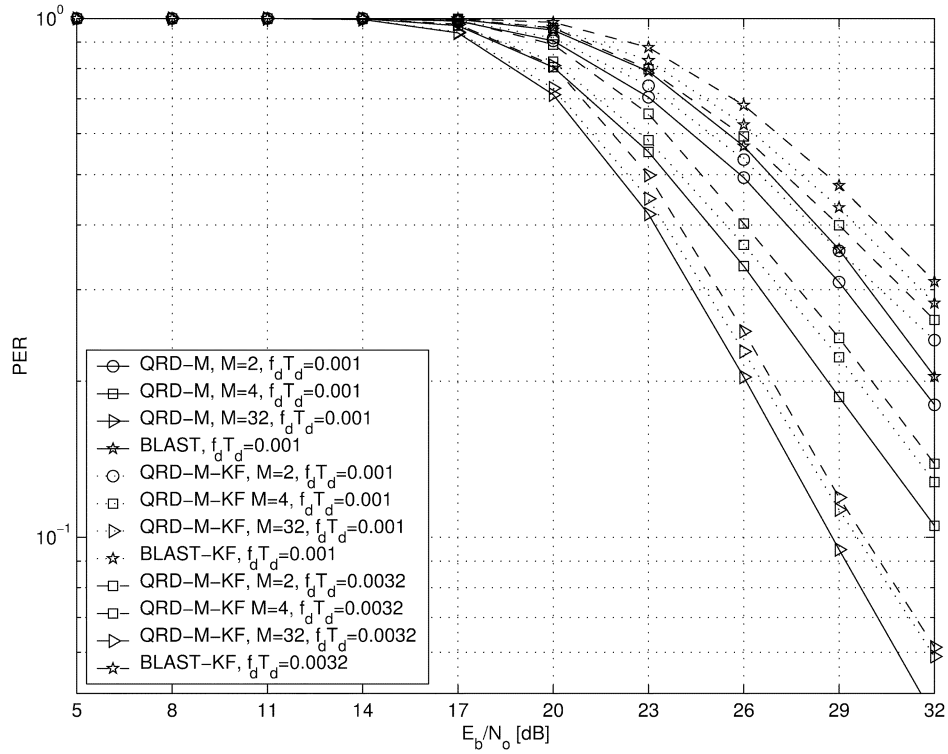

Fig. 8. PER performance of the QRD-M-KF system versus QRD-M system [16-QAM, one OFDM symbols, $N_t = 4, N_r = 4, N_f = 2, f_dT_d = 0.001/0.0032$].

sequence probability is then

$$\begin{aligned} &P(e_p = 2(-2)|e_{p+1}, \ldots, e_{N_t}) \\ &\quad = P(m_p < (>)0|d_p = 1(-1), e_{p+1}, \ldots, e_{N_t}) P(d_p = \pm 1) \\ &\quad = \frac{1}{2} Q\left(\frac{|R_{p,p}| + (-) \sum_{l=p+1}^{N_t} \mathrm{Re}\{R'_{p,l}\} e_l}{\sqrt{\frac{N_0}{T_s}}} \right). \end{aligned} \tag{A.3}$$

The conditional pmf for e_p is finally written as

$$\begin{aligned} &p(e_p|e_{p+1}, \ldots, e_{N_t}) \\ &\quad = \delta_{e_p,2} \frac{1}{2} Q\left(\frac{|R_{p,p}| + \sum_{l=p+1}^{N_t} \mathrm{Re}\{R'_{p,l}\} e_l}{\sqrt{\frac{N_0}{T_s}}} \right) \\ &\qquad + \delta_{e_p,-2} \frac{1}{2} Q\left(\frac{|R_{p,p}| - \sum_{l=p+1}^{N_t} \mathrm{Re}\{R'_{p,l}\} e_l}{\sqrt{\frac{N_0}{T_s}}} \right) \\ &\qquad + \delta_{e_p,0} \left(1 - \frac{1}{2} Q\left(\frac{|R_{p,p}| + \sum_{l=p+1}^{N_t} \mathrm{Re}\{R'_{p,l}\} e_l}{\sqrt{\frac{N_0}{T_s}}} \right) \right. \\ &\qquad \left. - \frac{1}{2} Q\left(\frac{|R_{p,p}| - \sum_{l=p+1}^{N_t} \mathrm{Re}\{R'_{p,l}\} e_l}{\sqrt{\frac{N_0}{T_s}}} \right) \right) \end{aligned} \tag{A.4}$$

where $\delta_{k,j}$ is the Kronecker delta. The unconditional probability of error for d_p is then found according to

$$\begin{aligned} P_b^p = &\sum_{e_p = \pm 2} \sum_{e_{p+1} = -2,0,2} \cdots \sum_{e_{N_t} = -2,0,2} p(e_p|e_{p+1}, \ldots, e_{N_t}) \\ &\times p(e_{p+1}|e_{p+2}, \ldots, e_{N_t}) \ldots p(e_{N_t}). \end{aligned} \tag{A.5}$$

The BPSK error rate (A.5) can be extended to higher order modulations by determining the appropriate conditional error sequence pmfs. For example, for QPSK, e_p takes on the nine values $e_p = \{-2, 0, 2\} + j\{-2, 0, 2\}$. The unconditional SER becomes

$$\begin{aligned} P_s^p = &\sum_{\substack{e_p \neq 0, e_p = \\ \{-2,0,2\} + j\{-2,0,2\}}} \sum_{\substack{e_{p+1} = \\ \{-2,0,2\} + j\{-2,0,2\}}} \cdots \\ &\sum_{\substack{e_{N_t} = \\ \{-2,0,2\} + j\{-2,0,2\}}} (p(e_p|e_{p+1}, \ldots, e_{N_t}) \\ &\times p(e_{p+1}|e_{p+2}, \ldots, e_{N_t}) \ldots p(e_{N_t})). \end{aligned} \tag{A.6}$$

Furthermore, since the real and imaginary parts of v_p are independent, the real and imaginary parts of e_p are likewise independent when conditioned on $e_{p+1}, \ldots, e_{N_t}$. The conditional SER factors as

$$\begin{aligned} &P(e_p|e_{p+1}, \ldots, e_{N_t}) \\ &\quad = P(e_{p,r}|e_{p+1}, \ldots, e_{N_t}) P(e_{p,i}|e_{p+1}, \ldots, e_{N_t}). \end{aligned} \tag{A.7}$$

The pmf of $e_{p,i}$ is found in a manner similar to (A.4) as

$$\begin{aligned} &p(e_{p,i}|e_{p+1},\ldots,e_{N_t}) \\ &\quad = \delta_{e_{p,i},2}\frac{1}{2}Q\left(\frac{|R_{p,p}|+\sum_{l=p+1}^{N_t}\mathrm{Im}\{R'_{p,l}e_l\}}{\sqrt{\frac{N_0}{T_s}}}\right) \\ &\qquad + \delta_{e_{p,i},-2}\frac{1}{2}Q\left(\frac{|R_{p,p}|-\sum_{l=p+1}^{N_t}\mathrm{Im}\{R'_{p,l}e_l\}}{\sqrt{\frac{N_0}{T_s}}}\right) \\ &\qquad + \delta_{e_{p,i},0}\left(1-\frac{1}{2}Q\left(\frac{|R_{p,p}|+\sum_{l=p+1}^{N_t}\mathrm{Im}\{R'_{p,l}e_l\}}{\sqrt{\frac{N_0}{T_s}}}\right)\right. \\ &\qquad\qquad \left.-\frac{1}{2}Q\left(\frac{|R_{p,p}|-\sum_{l=p+1}^{N_t}\mathrm{Im}\{R'_{p,l}e_l\}}{\sqrt{\frac{N_0}{T_s}}}\right)\right). \end{aligned} \tag{A.8}$$

The pmf of $e_{p,r}$ is identical to (A.8) with $\mathrm{Re}\{\}$ operators replacing $\mathrm{Im}\{\}$.

REFERENCES

[1] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, no. 1, pp. 744–765, Mar. 1998.

[2] G. Raleigh and V. Jones, "Multivariate modulation and coding for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 17, no. 5, pp. 851–866, May 1999.

[3] G. D. Golden, C. J. Foschini, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, pp. 14–16, Jan. 1999.

[4] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. 1998 URSI Int. Symp. Signal, Systems, and Electronics*, 1998, pp. 295–300.

[5] Y. Li, J. H. Winters, and N. R. Sollenberger, "MIMO-OFDM for wireless communications: Signal detection with enhanced channel estimation," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1471–1477, Sept. 2002.

[6] L. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COM-33, no. 7, pp. 665–675, Jul. 1985.

[7] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, no. 1, pp. 119–124, Jan. 1989.

[8] R. J. Castle, A. E. Jones, and T. A. Wilkinson, "An experiment study of OFDM a 525 GHz in an office environment," *IEEE J. Select. Areas Commun.*, vol. 19, no. 11, pp. 2279–2289, Nov. 2001.

[9] S. Baro, G. Bauch, A. Pavlic, and A. Semmler, "Improving BLAST performance using space-time block codes and turbo decoding," in *Proc. 2000 IEEE Global Telecommunications Conf.*, 2000, pp. 1067–1071.

[10] M. O. Damen, K. Abed-Meraim, and S. Burykh, "Iterative QR detection for BLAST," *Wireless Personal Commun.*, vol. 19, pp. 179–191, 2001.

[11] Y. Cho and J. Lee, "Analysis of an adaptive SIC for near-far resistant DS-CDMA," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1429–1431, Nov. 1998.

[12] K. Lai and J. J. Shynk, "Steady-state analysis of the adaptive successive inteference canceler for DS/CDMA signals," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2345–2362, Oct. 2001.

[13] 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, May 2002.

[14] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.

[15] L. Wei, L. Rasmussen, and R. Wyrwas, "Near optimum tree-search detection schemes for bit-synchronous multiuser CDMA systems over Gaussian and two-path Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 45, no. 6, pp. 691–700, Jun. 1997.

[16] Z. Xie, C. K. Rushforth, R. T. Short, and T. K. Moon, "Joint signal detection and parameter estimation in multiuser communications," *IEEE Trans. Commun.*, vol. 41, no. 8, pp. 1208–1216, Aug. 1993.

[17] T. M. Aulin, "Breadth-first maximum likelihood sequence detection: Basics," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 208–216, Feb. 1999.

[18] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[19] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. London, U.K.: Chapman and Hall, 1986.

[20] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.

[21] J. Fuchs, "Multipath time-delay detection and estimation," *IEEE Trans. Signal Process.*, vol. 47, pp. 237–243, Jan. 1999.

[22] B. Yang, Z. Cao, and K. B. Letaief, "Analysis of low-complexity windowed DFT-based MMSE channel estimator for OFDM systems," *IEEE Trans. Commun.*, vol. 49, pp. 1977–1987, Nov. 2001.

[23] T. Tung, "Adaptive space time signal processing for wireless communication and sensor systems," Ph.D. dissertation, Electrical Engineering Dept., Univ. California, Los Angeles, 2001.

[24] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 3, pp. 129–137, Mar. 1982.

Kyeong Jin Kim (S'98–M'00) was born in Korea. He received the M.S. degree from the Korea Advanced Institute of Science and Technology (KAIST) in 1991 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, Santa Barbara, in 2000.

From 1991 to 1995, he was a Research Engineer at the video research center, Daewoo Electronics, Ltd., Korea. In 1997, he joined the Data Transmission and Networking Laboratory, University of California, Santa Barbara. After receiving his degrees, he joined the Nokia Research Center, Dallas, TX, as a Senior Research Engineer. His research interests include development of joint detection and channel estimation algorithms for code-division multple access, multi-input multi-output orthogonal frequency-division multiplexing, and 4G systems.

Jiang Yue (S'02) received the B.S. degree from Southeast University, China, in 1995, and the M.S. degree from Nanyang Technological University, Singapore, in 1999. He is pursuing the Ph.D. degree in electrical engineering at Southern Methodist University, Dallas, TX.

Since 1999, he has been a Research Assistant with the Department of Electrical Engineering, Southern Methodist University. His research interests include wireless communications, multi-input multi-output signal processing, error control coding, and channel estimation and tracking.

Ronald A. Iltis (S'83–M'84–SM'91) received the B.A. degree in biophysics from The Johns Hopkins University, Baltimore, MD, in 1978, the M.Sc. degree in engineering from Brown University in 1980, and the Ph.D. degree in electrical engineering from the University of California, San Diego, in 1984.

Since 1984, he has been with the University of California, Santa Barbara, where he is currently a Professor in the Department of Electrical and Computer Engineering. His current research interests include code-division multple access, software radio, radiolocation, and nonlinear estimation. He has also served as a consultant to government and private industry in the areas of adaptive arrays, neural networks, and spread-spectrum communications.

Dr. Iltis was previously an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. In 1990, he received the Fred W. Ellersick award for best paper at the IEEE MILCOM conference.

Jerry D. Gibson (S'73–M'73–SM'81–F'92) currently serves as Professor of Electrical and Computer Engineering and Professor of Media Arts and Technology, University of California, Santa Barbara. He was Chairman of the Department of Electrical Engineering, Southern Methodist University, Dallas, TX, from 1997 to 2002, and from 1987 to 1997, he held the J. W. Runyon, Jr., Professorship in the Department of Electrical Engineering, Texas A&M University. He is coauthor of the books *Digital Compression for Multimedia* (San Diego, CA: Morgan-Kaufmann, 1998) and *Introduction to Nonparametric Detection with Applications* (New York: Academic, 1975, and New York: IEEE Press, 1995) and author of *Principles of Digital and Analog Communications* (Englewood Cliffs, NJ: Prentice-Hall, 1993). He is Editor-in-Chief of *The Mobile Communications Handbook* (Boca Raton, FL: CRC, 1999), Editor-in-Chief of *The Communications Handbook* (Boca Raton, FL: CRC, 2002), and Editor of the book, *Multimedia Communications: Directions and Innovations* (New York: Academic, 2000). His research interests include data, speech, image, and video compression, multimedia over networks, wireless communications, information theory, and digital signal processing.

Dr. Gibson was Associate Editor for Speech Processing for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1981 to 1985 and Associate Editor for Communications for the IEEE TRANSACTIONS ON INFORMATION THEORY from 1988 to 1991. He has served as a member of the Speech Technical Committee of the IEEE Signal Processing Society (1992–1995), as a member of the IEEE Information Theory Society Board of Governors (1990–1998), and as a member of the Editorial Board for the PROCEEDINGS OF THE IEEE. He was President of the IEEE Information Theory Society in 1996. He served as Technical Program Chair of the 1999 IEEE Wireless Communications and Networking Conference, Technical Program Chair of the 1997 Asilomar Conference on Signals, Systems, and Computers, Finance Chair of the 1994 IEEE International Conference on Image Processing, and General Co-Chair of the 1993 IEEE International Symposium on Information Theory. Currently, he serves on the Steering Committee for the Wireless Communications and Networking Conference. In 1990, he received the Fredrick Emmons Terman Award from the American Society for Engineering Education. He was a Co-Recipient of the 1993 IEEE Signal Processing Society Senior Paper Award for the Speech Processing area.