Adaptive Frequency-Domain Equalization for Single-Carrier Multiple-Input Multiple-Output Wireless Transmissions

Justin Coon, *Student Member, IEEE*, Simon Armour, *Member, IEEE*, Mark Beach, *Associate Member, IEEE*, and Joe McGeehan

***Abstract*—Channel estimation and tracking pose real problems in broadband single-carrier wireless communication systems employing multiple transmit and receive antennas. An alternative to estimating the channel is to adaptively equalize the received symbols. Several adaptive equalization solutions have been researched for systems operating in the time domain. However, these solutions tend to be computationally intensive. A low-complexity alternative is to adaptively equalize the received message in the frequency domain. In this paper, we present an adaptive frequency-domain equalization (FDE) algorithm for implementation in single-carrier (SC) multiple-input multiple-output (MIMO) systems. Furthermore, we outline a novel method of reducing the overhead required to train the proposed equalizer. Finally, we address the issues of complexity and training sequence design. Other computationally efficient adaptive FDE algorithms for use in SC systems employing single transmit and receive antennas, receive diversity, or space-time block codes (STBC) can be found in the literature. However, the algorithm detailed in this paper can be implemented in STBC systems as well as in broadband spatial multiplexing systems, making it suitable for use in high data rate MIMO applications.**

***Index Terms*—Adaptive equalization, frequency-domain equalization (FDE), multiple-input multiple-output (MIMO) systems.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) architectures are very attractive solutions for high date rate wireless communication systems due to their enormous potential for capacity gains relative to single-antenna systems [1], [2]. Channel equalization in broadband MIMO systems can potentially be very complex due to the superposition of all of the transmitted streams at each receive antenna. The complexity of the equalization process can be mitigated somewhat by performing equalization in the frequency-domain at the receiver. There are two fundamental frequency-domain equalization (FDE) techniques: orthogonal frequency division multiplexing (OFDM) [3]–[5] and single-carrier (SC) transmission with FDE (SC-FDE) [6]–[10]. Although OFDM is currently the favored technique for use in MIMO systems, recent studies have shown that MIMO SC-FDE offers significant performance improvements over MIMO OFDM in certain environments [11]. In this paper, we present an adaptive equalization algorithm for implementation in MIMO SC-FDE systems.

To date, several adaptive solutions for SC MIMO systems have been studied. In [12], adaptive time-domain equalization for SC MIMO systems was addressed. The issue of MIMO channel tracking was explored for adaptive time-domain equalization in [13]. Adaptive algorithms based on the least mean squares (LMS) algorithm and the recursive least squares (RLS) algorithm were considered for SC-FDE systems employing receive diversity in [14], where the potential for significant reductions in complexity relative to time-domain LMS and RLS algorithms was shown. Another modification of the RLS algorithm was applied to a space-time coded SC-FDE system in [15], where the structure of the space-time block code was exploited in order to reduce the complexity of the algorithm beyond the reduction achieved through FDE alone.

The main drawback of the adaptive FDE algorithms mentioned above is that they cannot be applied to MIMO systems with spatial multiplexing (SM) architectures. As a result, they are not suitable for use in very high data rate applications where SM systems with an appropriate channel code, puncturing, and interleaving would be most useful. The algorithm proposed in this paper can be implemented in SM architectures as well as space-time block-coded architectures. In order to highlight the advantages of this algorithm, we focus on its implementation in an SM system in this paper. The application of the algorithm to space-time block-coded systems is straightforward.

This paper is organized as follows. In Section II, we introduce a mathematical model for an SM SC-FDE system. We present a detailed mathematical formulation of the proposed algorithm in Section III, which, to the best of our knowledge, cannot be found in the current literature. Convergence properties of the algorithm are discussed in Section IV. These properties are then exploited in Section V to devise a novel method of reducing the overhead required to train the adaptive equalizer. We address issues concerning the design of training sequences in Section VI, and the complexity of the proposed algorithm is compared to that of alternative techniques in Section VII. Finally, numerical

Manuscript received December 23, 2003; revised October 13, 2004. This paper was presented in part at the IEEE International Conference on Communications, Paris, France, 2004. This work was supported by Toshiba TRL Bristol. The associate editor coordinating the review of this manuscript and approvng it for publication was Dr. Petar M. Djuric.

J. Coon is with Toshiba Telecommunications Research Laboratory, Bristol BS1 4ND, U.K. (e-mail: justin.coon@toshiba-trel.com).

S. Armour and M. Beach are with the Centre for Communications Research, University of Bristol, Bristol BS8 1UB, U.K. (e-mail: simon.armour@bris.ac.uk; m.a.beach@bris.ac.uk).

J. McGeehan is with Toshiba Telecommunications Research Laboratory, Bristol BS1 4ND, U.K. He is also with the Centre for Communications Research, University of Bristol, Bristol, U.K. (e-mail: joe.mcgeehan@toshiba-trel.com).

Digital Object Identifier 10.1109/TSP.2005.851155

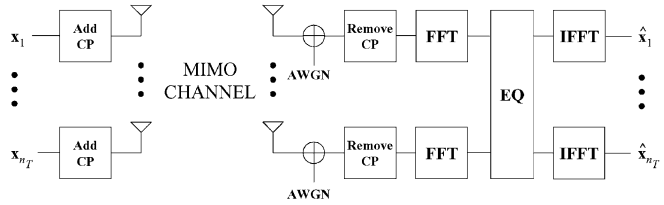

Fig. 1. Block diagram of baseband SM SC-FDE system.

results are illustrated in Section VIII, and conclusions are presented in Section IX.

Notation: We use a bold uppercase (lowercase) font to denote matrices (column vectors); frequency-domain variables are denoted by a tilde (e.g., $\tilde{a}$); $\mathbf{F}_m$ is the normalized $m \times m$ discrete Fourier transform (DFT) matrix, where its (k,i)th element is given by $F_{m;k,i} \triangleq (1/\sqrt{m})\exp(-j2\pi ki/m)$ for $k,i = 0,\ldots,m-1$; $\mathbf{I}_m$ is the $m\times m$ identity matrix; $\mathbf{0}_{m\times n}$ is an $m\times n$ all-zero matrix; $(\cdot)^*, (\cdot)^{-1}, (\cdot)^{\mathrm{T}}, (\cdot)^{\mathrm{H}}, (\cdot)_m$, and $|\cdot|$ denote the complex conjugate, inverse, transpose, conjugate transpose, modulo-m, and absolute value operations, respectively; $\otimes$ is the Kronecker product operator; $\mathrm{E}\{\cdot\}$ is the expectation operator; $\mathrm{Tr}\{\cdot\}$ is the trace operator; and $\mathrm{diag}\{x_0,\ldots,x_{m-1}\}$ denotes the $m \times m$ diagonal matrix with the elements $\{x_0,\ldots,x_{m-1}\}$ on the diagonal.

II. Spatial Multiplexing System Model for SC-FDE

Consider an SM system with n_T transmit antennas and n_R receive antennas, where $n_T \le n_R$ and the baseband sequence at each transmit antenna is modulated onto a single carrier waveform for transmission across a wireless channel. The received baseband sequences are equalized in the frequency domain. To facilitate FDE, a cyclic prefix is added to each sequence at the transmitter and removed from each of the received sequences. A block diagram of this system is illustrated in Fig. 1.

Adopting matrix notation, we can mathematically describe this system. Let $\mathbf{x}_q$ be a length-K vector of symbols that is transmitted from the qth transmit antenna. The symbols in $\mathbf{x}_q$ are drawn from an arbitrary constellation [e.g., binary phase-shift keying (BPSK)]. Typically, K is defined as a power of two to allow for the implementation of the fast Fourier transform (FFT) during equalization. A cyclic prefix of Q symbols is added to $\mathbf{x}_q$ prior to transmission, and the first Q symbols are removed from each of the received vectors. In the following discussion, we assume $Q \ge L$, where L is the memory order of each of the $n_T n_R$ channel impulse responses (CIRs). From this assumption, the exploitation of the cyclic prefix creates the illusion of periodicity in the transmitted message, thereby allowing the wireless channels to be expressed as circulant matrices, as given in the following expression for the vector $\mathbf{y}_p$ of symbols received at antenna p.

$$\mathbf{y}_p = \sum_{q=1}^{n_T} \mathbf{H}_{p,q}\mathbf{x}_q + \mathbf{n}_p. \tag{1}$$

In (1), $\mathbf{n}_p$ is a length-K vector of independent and identically distributed (i.i.d.) zero-mean complex Gaussian noise samples with variance $\sigma_n^2/2$ per dimension, and $\mathbf{H}_{p,q}$ is a $K \times K$ circulant matrix defined by the CIR between the qth transmit antenna and the pth receive antenna. Specifically, the first row of $\mathbf{H}_{p,q}$ is $(h_{p,q;0}, \mathbf{0}_{1\times K-(L+1)}, h_{p,q;L},\ldots,h_{p,q;1})$, where $h_{p,q;i}$ is the ith complex tap coefficient of the CIR between the qth transmit antenna and the pth receive antenna. It is assumed that the channels remain static for at least one block duration.

It is convenient to construct a length-$n_R K$ vector of received symbols $\mathbf{y} \triangleq (\mathbf{y}_1^{\mathrm{T}},\ldots,\mathbf{y}_{n_R}^{\mathrm{T}})^{\mathrm{T}}$. A length-$n_T K$ vector $\mathbf{x}$ of transmitted symbols can be constructed in a similar manner, where $\mathbf{x} \triangleq (\mathbf{x}_1^{\mathrm{T}},\ldots,\mathbf{x}_{n_T}^{\mathrm{T}})^{\mathrm{T}}$. Therefore, we have the linear system

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2}$$

where $\mathbf{n} \triangleq (\mathbf{n}_1^{\mathrm{T}},\ldots,\mathbf{n}_{n_R}^{\mathrm{T}})^{\mathrm{T}}$, and the (p,q)th submatrix of the $n_R \times n_T$ block matrix $\mathbf{H}$ is $\mathbf{H}_{p,q}$.

Equalization is performed on the received symbol vector $\mathbf{y}$ in the frequency domain. Consequently, the equalized symbols can be expressed by

$$\hat{\mathbf{x}} = \mathbf{D}_{n_T}^{-1}\mathbf{\Gamma}^{\mathrm{H}}\mathbf{D}_{n_R}\mathbf{y} \tag{3}$$

where $\mathbf{D}_m = \mathbf{I}_m \otimes \mathbf{F}_K$, and $\mathbf{\Gamma}^{\mathrm{H}}$ is the $n_T K \times n_R K$ equalizer matrix. Because each circulant submatrix of $\mathbf{H}$ is diagonalized by pre- and post-multiplication of a DFT and an indiscrete DFT (IDFT) matrix, respectively, it is convenient to express (3) as

$$\begin{aligned}\hat{\mathbf{x}} &= \mathbf{D}_{n_T}^{-1}\mathbf{\Gamma}^{\mathrm{H}}\tilde{\mathbf{y}} \\ &= \mathbf{D}_{n_T}^{-1}\mathbf{\Gamma}^{\mathrm{H}}(\tilde{\mathbf{H}}\tilde{\mathbf{x}} + \tilde{\mathbf{n}})\end{aligned} \tag{4}$$

where $\tilde{\mathbf{y}} = \mathbf{D}_{n_R}\mathbf{y}, \tilde{\mathbf{H}} = \mathbf{D}_{n_R}\mathbf{H}\mathbf{D}_{n_T}^{-1}, \tilde{\mathbf{x}} = \mathbf{D}_{n_T}\mathbf{x}$, and $\tilde{\mathbf{n}} = \mathbf{D}_{n_R}\mathbf{n}$. The (p,q)th diagonal $K \times K$ submatrix of $\tilde{\mathbf{H}}$ is given by $\tilde{\mathbf{H}}_{p,q} \triangleq \mathrm{diag}\{\tilde{h}_{p,q;0},\ldots,\tilde{h}_{p,q;K-1}\}$, where the discrete frequency response of the channel is given by $\tilde{h}_{p,q;k} = \sum_{i=0}^{L} h_{p,q;i}\exp(-j2\pi ki/K)$. This mathematical model for an SM SC-FDE system is used throughout this paper.

III. Adaptive Algorithm

The adaptive FDE algorithm described in this section is a version of the RLS algorithm. The goal is to adaptively update

$$\mathbf{\Gamma}^{\mathrm{H}} = \begin{pmatrix} \gamma_{0,0}^* & \cdots & \gamma_{n_RK-1,0}^* \\ \vdots & \ddots & \vdots \\ \gamma_{0,n_TK-1}^* & \cdots & \gamma_{n_RK-1,n_TK-1}^* \end{pmatrix}$$

utilizing training blocks (training mode) or detected data blocks (decision-directed mode). However, the update process requires the revision of $n_T n_R K^2$ elements, which can become computationally cumbersome as the block length K increases. We can reduce the number of elements that must be updated by observing that only the n_R elements of $\tilde{\mathbf{y}}$ corresponding to the kth frequency bin are required to recover the kth frequency component of the block transmitted from antenna q. It follows that

$$\gamma_{u,v}^* = \begin{cases} a_{u,v} + jb_{u,v}, & (|u-v|)_K = 0 \\ 0, & \text{otherwise} \end{cases}$$

for some real numbers $a_{u,b}$ and $b_{u,v}$. Thus, the number of elements in $\mathbf{\Gamma}^{\mathrm{H}}$ that must be updated is reduced to $n_T n_R K$.

To update the nonzero elements of $\mathbf{\Gamma}^{\mathrm{H}}$, consider the classical cost function defined by

$$J_v(t) = \sum_{\ell=1}^{t} \zeta(t,\ell)|E_v(\ell,t)|^2, \quad \forall v = 0,\ldots,n_T K-1 \quad (5)$$

where ℓ is a time index denoting a given block interval, t is a time index denoting the current block interval, and $\zeta(t,\ell) = \rho^{t-\ell}$ is the standard weighting factor, which is included for implementation of the algorithm in decision-directed mode. A *block interval* is defined here as the interval of time in which one block is sent from each transmit antenna. In (5), the error term $E_v(\ell,t)$ is given by

$$E_v(\ell,t) = \tilde{x}_v(\ell) - \sum_{\substack{u,\\ (|u-v|)_K=0}} \gamma^*_{u,v}(t)\tilde{y}_u(\ell). \quad (6)$$

The notation $\tilde{x}_v(\ell)$ denotes the vth element of the vector $\tilde{\mathbf{x}}$ at time ℓ.

The objective is to minimize $J_v(t)$ for each v. Taking the partial derivative of (5) with respect to $\gamma^*_{u,v}(t)$, setting the result equal to zero, and performing some algebraic manipulations yields

$$\mathbf{R}_v(t)\boldsymbol{\gamma}_v(t) = \mathbf{p}_v(t) \quad (7)$$

where

$$\mathbf{R}_v(t) = \sum_{\ell=1}^{t} \rho^{t-\ell}\boldsymbol{\psi}_v(\ell)\boldsymbol{\psi}_v^{\mathrm{H}}(\ell) \quad (8)$$

and

$$\mathbf{p}_v(t) = \sum_{\ell=1}^{t} \rho^{t-\ell}\tilde{x}^*_v(\ell)\boldsymbol{\psi}_v(\ell). \quad (9)$$

In (7) to (9), $\boldsymbol{\gamma}_v(t) \triangleq (\gamma_{u_0,v}(t),\ldots,\gamma_{u_{n_R-1},v}(t))^{\mathrm{T}}$, and $\boldsymbol{\psi}_v(t) \triangleq (\tilde{y}_{u_0}(t),\ldots,\tilde{y}_{u_{n_R-1}}(t))^{\mathrm{T}}$, where the index $u_m \in \{0,\ldots,n_R K-1\}$ such that $(|u_m - v|)_K = 0$. We may rewrite (8) as

$$\mathbf{R}_v(t) = \rho\mathbf{R}_v(t-1) + \boldsymbol{\psi}_v(t)\boldsymbol{\psi}_v^{\mathrm{H}}(t). \quad (10)$$

Similarly, we can rewrite (9) as

$$\mathbf{p}_v(t) = \rho\mathbf{p}_v(t-1) + \tilde{x}^*_v(t)\boldsymbol{\psi}_v(t). \quad (11)$$

Utilizing (7) to (11), it can be shown that the update equation for the nonzero elements in the vth column of $\mathbf{\Gamma}$ at time t is given by

$$\boldsymbol{\gamma}_v(t) = \boldsymbol{\gamma}_v(t-1) + \mathbf{R}_v^{-1}(t)\boldsymbol{\psi}_v(t)\varepsilon_v(t) \quad (12)$$

where $\varepsilon_v(t) = \tilde{x}^*_v(t) - \boldsymbol{\psi}_v^{\mathrm{H}}(t)\boldsymbol{\gamma}_v(t-1)$.

To this point, the derivation of the adaptive algorithm has more or less followed that of the standard RLS algorithm [16]. The key difference in our derivation here is the formulation of the RLS problem on a tonal basis in the context of a multidimensional, multiantenna system. It was important to step through this derivation to show the first of several interesting points that will be made about this algorithm: The time-average correlation matrix $\mathbf{R}_v(t)$ is an $n_R \times n_R$ matrix. Consequently, for small numbers of receive antennas, the inverse of $\mathbf{R}_v(t)$ can be computed directly with ease. In contrast, the size of the analogous correlation matrix in the standard time-domain RLS algorithm is dependent on the length of the time-domain filter, which grows large with increasing channel memory (i.e., as L increases) [16]. Therefore, in most broadband applications of interest, the time-domain RLS algorithm can only be implemented via the matrix inversion lemma, whereas the proposed FDE algorithm can be implemented easily for small n_R in any environment. Indeed, for MIMO systems with large numbers of receive antennas, the matrix inversion lemma can be used to compute $\mathbf{R}_v^{-1}(t)$ as well, which is given by

$$\mathbf{R}_v^{-1}(t) = \rho^{-1}\mathbf{R}_v^{-1}(t-1) - \frac{\rho^{-2}\mathbf{R}_v^{-1}(t-1)\boldsymbol{\psi}_v(t)\boldsymbol{\psi}_v^{\mathrm{H}}(t)\mathbf{R}_v^{-1}(t-1)}{1+\rho^{-1}\boldsymbol{\psi}_v^{\mathrm{H}}(t)\mathbf{R}_v^{-1}(t-1)\boldsymbol{\psi}_v(t)}. \quad (13)$$

The adaptive algorithm is summarized in Table I, where the matrix inversion lemma is used to calculate $\mathbf{R}_v^{-1}(t)$.

TABLE I
ADAPTIVE ALGORITHM FOR SM SC-FDE

Initialization:

$\mathbf{\Gamma}(0) = \mathbf{0}_{n_R K \times n_T K}$

$\mathbf{R}_v(0) = \delta\mathbf{I}_{n_R}, \quad \forall v = 0,\ldots,n_T K-1$ and for small δ

$\rho \leftarrow$ some number close to, but less than, 1

$t \leftarrow 1$

Recursion:

$\forall v = 0,\ldots,n_T K-1$

$\mathbf{R}_v^{-1}(t) = \rho^{-1}\mathbf{R}_v^{-1}(t-1) - \frac{\rho^{-2}\mathbf{R}_v^{-1}(t-1)\boldsymbol{\psi}_v(t)\boldsymbol{\psi}_v^{\mathrm{H}}(t)\mathbf{R}_v^{-1}(t-1)}{1+\rho^{-1}\boldsymbol{\psi}_v^{\mathrm{H}}(t)\mathbf{R}_v^{-1}(t-1)\boldsymbol{\psi}_v(t)}$

$\varepsilon_v(t) = \tilde{x}^*_v(t) - \boldsymbol{\psi}_v^{\mathrm{H}}(t)\boldsymbol{\gamma}_v(t-1)$

$\boldsymbol{\gamma}_v(t) = \boldsymbol{\gamma}_v(t-1) + \mathbf{R}_v^{-1}(t)\boldsymbol{\psi}_v(t)\varepsilon_v(t)$

$t \leftarrow t+1$

IV. CONVERGENCE PROPERTIES

The convergence properties of the proposed algorithm were studied in order to gain an understanding of the performance of the algorithm. The properties that are of greatest interest are the mean-square error (MSE) convergence and the rate of convergence. In general, the derivations of these convergence properties follow directly from the time-domain RLS algorithm, which can be found, for example, in [16]. However, the derivations included here contain additional steps, which provide insight into several aspects of this algorithm and allow the overhead involved in training the equalizer to be reduced from the general

case, as discussed in the next section. Attention is drawn here to the significance of each convergence property in the context of the basic algorithm presented in Section III. First, consider the MSE convergence from which the rate of convergence follows.

A. MSE Convergence

It can be shown that the adaptive equalizer converges to a steady-state solution [16]. The rate at which the adaptive equalizer converges to this solution can be examined through an MSE analysis. To derive an expression for the MSE convergence of the algorithm, we first define the vth weight-error vector by

$$\mathbf{e}_v(t) = \boldsymbol{\gamma}_v(t) - \bar{\boldsymbol{\gamma}}_v \tag{14}$$

where $\bar{\boldsymbol{\gamma}}_v = \mathrm{E}\{\boldsymbol{\gamma}_v(t)\}$ is the vth steady-state solution. The vth weight-error correlation matrix is given by

$$\mathbf{K}_v(t) = \mathrm{E}\left\{\mathbf{e}_v(t)\mathbf{e}_v^{\mathrm{H}}(t)\right\}. \tag{15}$$

The MSE of the vth equalizer vector relative to the mean solution $\bar{\boldsymbol{\gamma}}_v$ can be found by taking the trace of $\mathbf{K}_v(t)$. Assuming that

1) the vectors $\boldsymbol{\psi}_v(1), \ldots, \boldsymbol{\psi}_v(t)$ are i.i.d.;
2) $\boldsymbol{\psi}_v(1), \ldots, \boldsymbol{\psi}_v(t)$, where $t \geq n_R$ are drawn from a stochastic process with a zero-mean Gaussian distribution with an ensemble-average correlation matrix $\boldsymbol{\Phi}_v = \mathrm{E}\{\boldsymbol{\psi}_v\ \boldsymbol{\psi}_v^{\mathrm{H}}\}$;

the MSE of the vth equalizer vector can be expressed as

$$\begin{aligned}\mathrm{MSE}(t) &= \frac{\sigma^2}{t - n_R - 1}\mathrm{Tr}\left\{\boldsymbol{\Phi}_v^{-1}\right\} \\ &= \frac{\sigma^2}{t - n_R - 1}\sum_{p=1}^{n_R}\frac{1}{\lambda_{v,p}}, \quad t > n_R + 1\end{aligned} \tag{16}$$

where $\lambda_{v,1}, \ldots, \lambda_{v,n_R}$ are the eigenvalues of $\boldsymbol{\Phi}_v$ [16]. The term σ^2 denotes the variance of a zero-mean measurement error process $\epsilon_{0,v}(t)$ that is adopted from a multiple linear regression model, which is given by [16]

$$\tilde{x}_v(t) = \bar{\boldsymbol{\gamma}}_v^{\mathrm{H}}\boldsymbol{\psi}_v(t) + \epsilon_{0,v}(t). \tag{17}$$

Since $\boldsymbol{\Phi}_v$ is positive definite, the eigenvalues of $\boldsymbol{\Phi}_v$ are positive. Therefore, we may write

$$\begin{aligned}\mathrm{MSE}(t) &= \frac{\sigma^2}{t - n_R - 1}\sum_{p=1}^{n_R}\frac{1}{\lambda_{v,p}} \\ &\geq \frac{\sigma^2 n_R}{t - n_R - 1}\left(\prod_{p=1}^{n_R}\frac{1}{\lambda_{v,p}}\right)^{1/n_R}\end{aligned} \tag{18}$$

which is met with equality if and only if $\lambda_{v,1} = \lambda_{v,2} = \cdots = \lambda_{v,n_R}$. Furthermore, we have

$$\prod_{p=1}^{n_R}\frac{1}{\lambda_{v,p}} = \det\left(\boldsymbol{\Phi}_v^{-1}\right) = [\det(\boldsymbol{\Phi}_v)]^{-1}. \tag{19}$$

In order to minimize (18), we must maximize $\det(\boldsymbol{\Phi}_v)$. Applying Hadamard's inequality, we may write

$$\det(\boldsymbol{\Phi}_v) \leq \prod_{p=1}^{n_R}\phi_{v;p,p} \tag{20}$$

where $\{\phi_{v;p,p}\}_{p=1}^{n_R}$ are the diagonal elements of $\boldsymbol{\Phi}_v$. Equation (20) is met with equality if and only if $\boldsymbol{\Phi}_v$ is diagonal, in which case, $\phi_{v;p,p} = \lambda_{v,p}$ for $p = 1, \ldots, n_R$. Therefore, in order to maximize $\det(\boldsymbol{\Phi}_v)$, and thus minimize the MSE, $\boldsymbol{\Phi}_v$ must be a diagonal matrix with equal elements on the diagonal for all v. Consequently, improper design of the training sequences and/or ill-conditioned channels may lead to poor MSE convergence. We will return to this condition for achieving the minimum MSE during our discussion of training sequence design in Section VI.

B. Rate of Convergence

The rate at which the algorithm converges can also be measured by the error in the signal at the output of the equalizer prior to the equalizer update operation. This error is known as the *a priori* estimation error and is given by

$$\xi_v(t) = \epsilon_{0,v}(t) - \mathbf{e}_v^{\mathrm{H}}(t-1)\boldsymbol{\psi}_v(t). \tag{21}$$

The study of the mean-square *a priori* estimation error as a function of block intervals i, which is given by

$$J_v'(t) = \mathrm{E}\{|\xi_v(t)|^2\} \tag{22}$$

is a useful measure of the rate of convergence. This metric is similar to the MSE metric used to study the convergence of the equalizer to the steady-state solution discussed above.

Expanding (22) and evaluating the resulting expectations yields

$$J_v'(t) = \sigma^2 + \mathrm{Tr}\{\mathbf{K}_v(t-1)\boldsymbol{\Phi}_v\}. \tag{23}$$

Therefore, using (16), the mean-square *a priori* estimation error as a function of block intervals can be written as

$$J_v'(t) = \sigma^2\left(1 + \frac{n_R}{t - n_R - 2}\right), \quad t > n_R + 2. \tag{24}$$

Thus, the rate at which the proposed algorithm converges decreases as the number of receive antennas n_R increases. Furthermore, the rate of convergence is proportional to the variance of the measurement error process, which is an intuitively satisfying result since this error is related to the noise at the receiver. The expression for the rate of convergence given in (24) can be used to reduce the number of symbols required to train the equalizer, as will be shown in the next section.

V. Reducing Training Overhead

Equation (24) suggests that for $n_R \gg 2$, approximately $2n_R$ block intervals must be used for training before the mean-square *a priori* estimation error reaches within 3 dB of its final value. From this result, a more practical solution would be to implement a "channel sounding" technique in which the channel is estimated by sequentially transmitting one training block from each transmit antenna, while the others remain silent, which

would require n_T block intervals, then construct an equalizer from the estimate. However, we conclude from (24) that the rate of convergence is not dependent on the lengths of the transmitted symbol vectors. Therefore, the transmitted *time-domain* training blocks can theoretically have any length κ, as long as K frequency components can be obtained from each block. Three cases of interest emerge from this observation:

1) $\kappa > K$) Choosing $\kappa > K$ during equalizer training but using a K-point FFT during equalization is inefficient. Although the resolution of the training sequences in the frequency domain is higher, a larger FFT is needed for training than for transmission. Furthermore, this increased resolution is not exploited during equalization of the received data symbols.
2) $\kappa = K$) In this case, the full block length is used to train the equalizer. The algorithm converges to the minimum mean-square *a priori* estimation error floor, as described by (24).
3) $\kappa < K$) Frequency-domain interpolation can be used to obtain K frequency components from a sequence of length κ if $\kappa < K$. This is a standard technique and has been presented in the literature [15]. In this case, the eigenvalues in (16) are generally smaller than if the full block size were used because a length-κ sequence of constant-modulus training symbols has less energy than a length-K sequence. Consequently, the MSE floor is higher for $\kappa < K$ than for other values of κ.

Now, consider a system employing the proposed algorithm where κ is initially less than K. Specifically, let $\kappa = \kappa_0$ in the set $\mathcal{K} = \{\kappa_0, \kappa_1, \ldots, \kappa_{\theta-1}\}$, where $Q \leq \kappa_0 < \kappa_1 < \cdots < \kappa_{\theta-1} \leq K$, and θ is the cardinality of the set $\mathcal{K}$. After τ training block intervals, κ is incremented to $\kappa = \kappa_1$. Following another τ training block intervals, κ is incremented to $\kappa = \kappa_2$, and so on. By periodically incrementing κ, convergence to a high error floor is prevented, which is a problem with systems implementing frequency-domain interpolation with a nonvarying block size $\kappa < K$, as is the case in [15].

In general, τ can be varied to optimize convergence, but τ is defined as a constant here. As long as τ is not too large and the system is well-conditioned, the algorithm will continue to converge, in terms of training block intervals, as described by (24). However, the number of symbol intervals used for training will depend on the choice of τ and $\mathcal{K}$. In particular, if $\tau\theta$ *block* intervals are used for training, the number of *symbol* intervals used for training is given by

$$T_s = \tau\theta Q + \tau \sum_{\vartheta=0}^{\theta-1} \kappa_\vartheta. \tag{25}$$

Thus, the length κ of each training block can be varied such that the number of symbol intervals required to train the algorithm can be dramatically reduced from the case where $\kappa = K$.

VI. Training Sequence Design

To begin the discussion of training sequence design, we revisit Section IV-A and the condition required to achieve the minimum MSE, namely, that $\mathbf{\Phi}_v$ must be a diagonal matrix where all elements on the diagonal are positive, real-valued, and equal. In most practical systems, it is impossible to achieve this condition since no prior knowledge of the channel is available. To illustrate this point, consider the linear system

$$\boldsymbol{\psi}_v = \mathbf{G}_k \boldsymbol{\chi}_v + \boldsymbol{\nu}_v \tag{26}$$

where $\boldsymbol{\chi}_v \triangleq (\tilde{x}_{u_0,v}, \ldots, \tilde{x}_{u_{n_T-1},v})^{\mathrm{T}}, \boldsymbol{\nu}_v \triangleq (\tilde{n}_{u_0,v}, \ldots, \tilde{n}_{u_{n_R-1},v})^{\mathrm{T}}$, and the (p,q)th term of the $n_R \times n_T$ mixing matrix $\mathbf{G}_k$ is $\tilde{h}_{p,q;k}$, where $k = (v)_K$ is a frequency index. The time index t has been omitted for brevity. For a given channel, we note that the noise and the transmitted signal are uncorrelated and write

$$\mathbf{\Phi}_v = \mathbf{G}_k \mathrm{E}\left\{\boldsymbol{\chi}_v \boldsymbol{\chi}_v^{\mathrm{H}}\right\} \mathbf{G}_k^{\mathrm{H}} + \sigma_n^2 \mathbf{I}_{n_R}. \tag{27}$$

Without possessing knowledge of the channel, it is obvious that the training sequences should be designed such that they have a mean of zero and are spectrally white. Equivalently, the covariance matrix of $\boldsymbol{\chi}_v$ should be given by

$$\mathbf{C}_{\boldsymbol{\chi}} = \mathrm{E}\left\{\boldsymbol{\chi}_v \boldsymbol{\chi}_v^{\mathrm{H}}\right\} = \beta \mathbf{I}_{n_T} \tag{28}$$

for some positive $\beta \in \mathbb{R}$.

The condition shown in (28) can be satisfied as follows. Consider a length-κ vector $\mathbf{a}_q$ of training symbols at the qth transmit antenna, where $\kappa \leq K$. The symbols in $\mathbf{a}_q$ are taken from a finite constellation such as BPSK. Let $\mathbf{a} \triangleq (\mathbf{a}_1^{\mathrm{T}}, \ldots, \mathbf{a}_{n_T}^{\mathrm{T}})^{\mathrm{T}}$. The K-element normalized DFT of $\mathbf{a}_q$ is given by

$$\tilde{\mathbf{a}}_q = (1/\sqrt{K})\mathbf{V}\mathbf{a}_q \tag{29}$$

where $\mathbf{V}$ is a $K \times \kappa$ Vandermonde matrix with its (k,i)th element $V_{k,i} \triangleq \exp(-j2\pi ki/K)$. Consequently, we have

$$\tilde{\mathbf{a}} = (1/\sqrt{K})\mathbf{L}\mathbf{a} \tag{30}$$

where $\mathbf{L} = \mathbf{I}_{n_T} \otimes \mathbf{V}$. We can construct a vector $\boldsymbol{\alpha}_v$ from $\tilde{\mathbf{a}}$, where $\boldsymbol{\alpha}_v \triangleq (\tilde{a}_{u_0}, \ldots, \tilde{a}_{u_{n_T-1}})^{\mathrm{T}}$, and $u_m \in \{0, \ldots, n_T K - 1\}$ such that $(|u_m - v|)_K = 0$ for all m.

Assuming $\boldsymbol{\alpha}_v$ has a mean of zero, the covariance matrix of $\boldsymbol{\alpha}_v$ is given by $\mathbf{C}_{\boldsymbol{\alpha}} = \mathrm{E}\{\boldsymbol{\alpha}_v\ \boldsymbol{\alpha}_v^{\mathrm{H}}\}$, which can be diagonalized to give $\mathbf{C}_{\boldsymbol{\alpha}} = \mathbf{\Sigma}\Delta\Sigma^{\mathrm{H}}$, where $\mathbf{\Delta}$ is a diagonal matrix with the eigenvalues of $\mathbf{C}_{\boldsymbol{\alpha}}$ on the diagonal, and $\mathbf{\Sigma}$ is a unitary matrix composed of the eigenvectors corresponding to the eigenvalues in $\mathbf{\Delta}$. It is important to note that a covariance matrix is positive semi-definite, which implies that its eigenvalues are real and non-negative. Indeed, a covariance matrix is positive definite if and only if its eigenvalues are positive [17]. Therefore, assuming the eigenvalues of $\mathbf{C}_{\boldsymbol{\alpha}}$ are positive, we can construct prewhitened training sequences by letting

$$\boldsymbol{\chi}_v = \sqrt{\beta}\mathbf{\Sigma}\mathbf{\Delta}^{-\frac{1}{2}}\mathbf{\Sigma}^{\mathrm{H}}\boldsymbol{\alpha}_v. \tag{31}$$

Consequently, the condition shown in (28) is met.

The time-domain training sequences $\{\mathbf{x}_q\}_{q=1}^{n_T}$ can be constructed by first computing $\boldsymbol{\chi}_v$ for all v as in (31) to obtain $\tilde{\mathbf{x}} \triangleq (\tilde{\mathbf{x}}_1^{\mathrm{T}}, \ldots, \tilde{\mathbf{x}}_{n_T}^{\mathrm{T}})^{\mathrm{T}}$ and then computing $\mathbf{x}_q = \sqrt{K}\mathbf{V}^{\dagger}\tilde{\mathbf{x}}_q$ for each of the n_T resulting frequency-domain vectors, where $\mathbf{V}^{\dagger} = (\mathbf{V}^{\mathrm{H}}\mathbf{V})^{-1}\mathbf{V}^{\mathrm{H}}$ is the pseudoinverse of $\mathbf{V}$. For the case where $\kappa = K$, this step can be simplified by calculating the direct inverse of $\mathbf{V}$ rather than the pseudoinverse. It should be noted that this step corresponds to performing a truncated IDFT.

Thus, training sequences constructed in this manner provide good performance only if they are relatively long (i.e., they are close to length-k).

From this analysis, it is seen that the parameter β can be explicitly defined. However, in practice, β is limited by the maximum transmission power of the system. If the restriction on the maximum transmission power is relaxed, β can be used to increase the power of the short training sequences in order to mitigate the problem of low training sequence energy, which leads to a high MSE, as discussed in the previous section.

In addition, it should be noted that this method of training sequence construction is susceptible to peak-to-average power ratio (PAPR) problems. Another solution to the problem of training sequence design is to manually construct the time-domain sequences $\{\mathbf{x}_q\}_{q=1}^{n_T}$ from a constant-modulus symbol alphabet such that $\mathbf{C}_{\boldsymbol{\chi}} = \beta \mathbf{I}_{nT}$, which would not cause a PAPR problem. However, the difficulty of this method is highlighted by the fact that the sequences are designed in the time domain, whereas the condition for optimality lies in the frequency domain.

VII. Complexity

The complexity of the proposed algorithm was compared with that of adaptive time-domain equalization for systems employing receive diversity in [14]. It was shown that a significant reduction in complexity can be obtained by adaptively equalizing the received message in the frequency domain as opposed to in the time domain. In this section, the complexity of the proposed algorithm is compared with that of the channel sounding technique mentioned in Section V and a channel estimation/equalizer construction technique based on the least squares (LS) criterion. The LS method was first developed for OFDM systems with transmit diversity in [18] and has been adapted for use in SC-FDE systems in this study.

Complexity is measured in terms of the number of complex multiplications that are performed in each method. It is assumed here that the frequency/time-domain transforms are carried out with FFT/IFFT operations. The number of multiplications executed in a matrix inversion and in an FFT vary slightly, depending on the implementation of the operations. Typically, an $m \times m$ matrix inversion requires m^3 multiplications, and an m-point FFT requires $m \log_2(m)$ multiplications. To preserve generality in this study, the numbers of multiplications required to compute an $m \times m$ matrix inversion and an m-point FFT are denoted by $(\cdot)_m^{-1}$ and FFT_m, respectively.

Equations defining the complexities of the proposed algorithm, the channel sounding technique, and the LS technique are given in Table II for training mode, whereas Table III depicts equations defining the complexities of the LS technique with equalizer construction and the proposed algorithm for decision-directed mode. Since the channel sounding technique is, by default, a technique that utilizes training symbols rather than detected/decoded data symbols, the complexity of this method was not studied for decision-directed mode. The complexity of determining the channel memory order L for the LS technique is not included in these calculations. The terms $\eta_{\text{cs}}, \eta_{\text{ad}}$, and η_{ls} are the numbers of block intervals used for channel estimation and/or equalizer training for the channel sounding technique, the proposed algorithm, and the LS technique, respectively. It should be noted that the equations presented in Tables II and III that relate to the proposed algorithm assume that the matrix inversion lemma is used to compute $\mathbf{R}_v^{-1}(t)$, as discussed in Section III.

TABLE II
Equations for Numbers of Complex Multiplications Performed During Training Mode

Algorithm	Number of multiplications
Channel sounding	$K\ (2n_R n_T^2 + \eta_{cs} n_R + (\cdot)_{n_R}^{-1})$
Proposed algorithm	$\eta_{ad} K n_T (5n_R^2 + 4n_R + 3)$
LS	$n_T^2 n_R (L+1)^2 + 2\eta_{ls} K n_T n_R (L+1) +$ $2K n_R n_T^2 + K(\cdot)_{n_R}^{-1} + n_T n_R \text{FFT}_K$

TABLE III
Equations for Numbers of Complex Multiplications Performed During One Training Iteration in Decision-Directed Mode

Algorithm	Number of multiplications
Proposed algorithm	$K n_T (5n_R^2 + 4n_R + 3) + n_T \text{FFT}_K$
LS	$n_T^2 n_R (L+1)^2 + 2K n_T^2 (2L+1) +$ $2K n_R n_T^2 + 2K n_T n_R (L+1) +$ $(\cdot)_{n_T(L+1)}^{-1} + K(\cdot)_{n_R}^{-1} + n_T (n_R+1) \text{FFT}_K$

A. Training Mode

In training mode, the complexity of the proposed algorithm is dominated by the number of iterations that are required for the algorithm to converge. This fact is illustrated in Fig. 2, where the numbers of complex multiplications used by the three techniques in a 2×2 system are plotted against the memory order of the channel L and the number of training blocks η_{ad} processed by the adaptive algorithm. In this analysis, we let $(\cdot)_m^{-1} = m^3$ and $\text{FFT}_m = m \log_2(m)$, and $\mathbf{R}_v^{-1}(t)$ is computed directly since it is only a 2×2 matrix. As observed in Fig. 2, the proposed algorithm becomes much more attractive from a complexity viewpoint as L increases, which follows from the fact that the length of the channel impulse response has no direct effect on the complexity of the proposed algorithm, as shown by the equations in Table II. However, for channels with low excess delay spread, the LS technique provides a good low-complexity method of estimating the channel and constructing the frequency-domain equalizer. This result follows from the fact that most of the computationally intensive operations in the LS technique can be precomputed if training data is used rather than detected data. It should, of course, be noted that channel sounding is the least complex of the three techniques that are

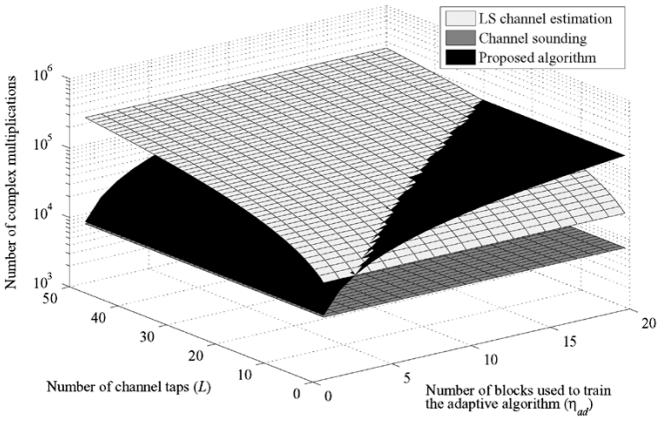

Fig. 2. Complexity of techniques for a 2×2 system operating in training mode where $\eta_{\rm ls} = \eta_{\rm cs} = 2$ and $K = 256$.

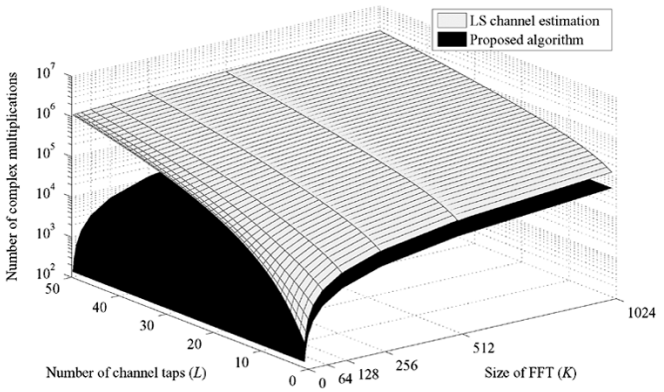

Fig. 3. Complexity of techniques for a 2×2 system operating in decision-directed mode, where $\eta_{\rm ls} = \eta_{\rm ad} = 1$.

studied here. However, this method suffers from convergence problems, which are addressed in the next section.

B. Decision-Directed Mode

The proposed algorithm shows much promise in terms of complexity when operating in decision-directed mode. Fig. 3 illustrates the numbers of complex multiplications used by the LS technique and the proposed algorithm as a function of the memory order of the channel L and the FFT size K. The number of transmit antennas for this example is $n_T = 2$, and the number of receive antennas is $n_R = 2$. Again, $\mathbf{R}_v^{-1}(t)$ is computed directly. As shown in Fig. 3, the complexity of the proposed algorithm is lower than that of the LS technique for all values of L and K when operating in decision-directed mode, although this advantage is only marginal when the channel memory order is low. It should be noted that this advantage is not limited to the case where $n_T = n_R = 2$ but in fact extends to systems in which more antennas are employed, even when the matrix inversion lemma is used to compute $\mathbf{R}_v^{-1}(t)$.

VIII. Simulation Results

Computer simulations were used to observe the rate of convergence of the proposed algorithm and the packet error rate (PER) of a system employing the algorithm. The channel

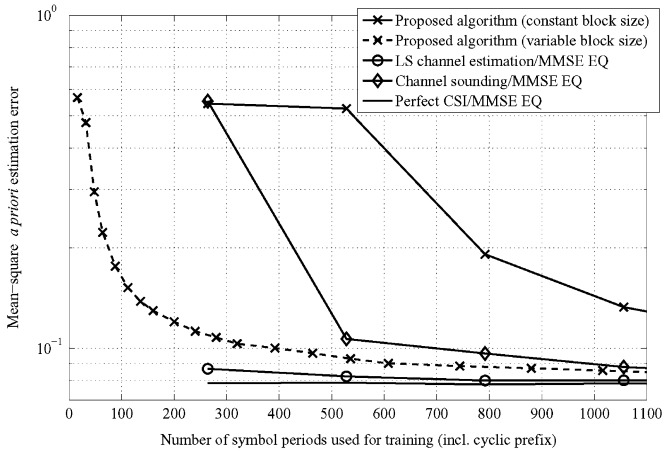

Fig. 4. Rate of convergence curves for FDE techniques employed in a 2×2 SC-FDE system operating in the ETSI BRAN A channel. (SNR per RX antenna $= 15$ dB, $K = 256, \mathcal{K} = \{8, 16, 32, 64, 128\}, Q = 8, \tau = 4$).

models used in the simulated systems are the ETSI BRAN A and E models [19]. The ETSI BRAN A model is specified for nonline-of-sight (NLOS) transmission in a typical office environment and has an RMS delay spread of 50 ns. The ETSI BRAN E model is specified for NLOS transmission in a large open environment and has an RMS delay spread of 250 ns. Root-raised cosine filters, each with a roll-off factor of 0.4, were employed at each transmitter and receiver, resulting in the utilization of 20-MHz bandwidth to support 14.3 Mbaud. Consequently, the memory order of the ETSI BRAN A channel is $L_A = 5$, and the memory order of the ETSI BRAN E channel is $L_E = 25$. A Doppler spread of 50 Hz was assumed. Therefore, the channel remained static throughout the transmission of the training blocks and for one packet in the case of the PER study.

A. Convergence Results

Figs. 4–6 illustrate the rate of convergence of the proposed algorithm implemented with both a constant training block size and a variable training block size. As a comparison, the convergence curves of two techniques that use an estimate of the channel to construct a linear MMSE frequency-domain equalizer are also depicted. The first of these techniques is the LS technique discussed in Section VII. The second technique is the aforementioned channel sounding technique. Chu sequences are used for training in the latter technique to provide a good estimate with n_T blocks [20]. These sequences are optimal for the channel sounding technique since they have constant-modulus elements in both the time domain and the frequency domain. Random training sequences are used in the LS technique, and optimal sequences are implemented in the proposed algorithm. Both of the channel estimation techniques utilize full-size training blocks for channel estimation. As a reference, the mean-square *a priori* estimation error produced by an MMSE frequency-domain equalizer constructed with perfect channel state information (CSI) and knowledge of the noise power is illustrated for each system in Fig. 4 to Fig. 6. It should be noted that the markers in these figures mark the ends of training blocks for the respective techniques. Thus, the utilization of short blocks to train the equalizer, as discussed

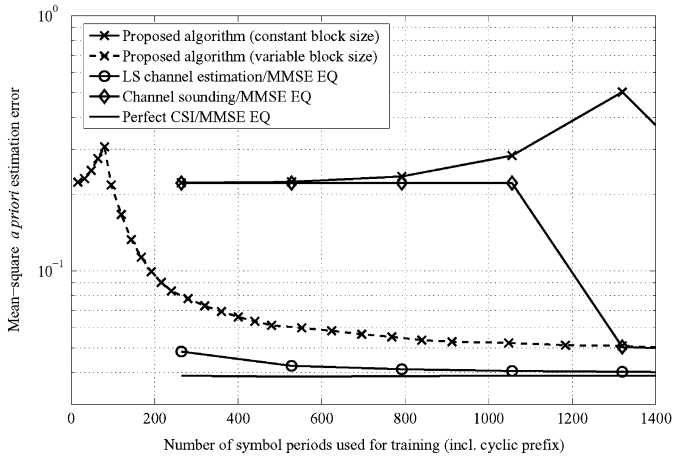

Fig. 5. Rate of convergence curves for FDE techniques employed in a 5×5 SC-FDE system operating in the ETSI BRAN A channel. (SNR per RX antenna $= 15$ dB, $K = 256, \mathcal{K} = \{8, 16, 32, 64, 128\}, Q = 8, \tau = 6$).

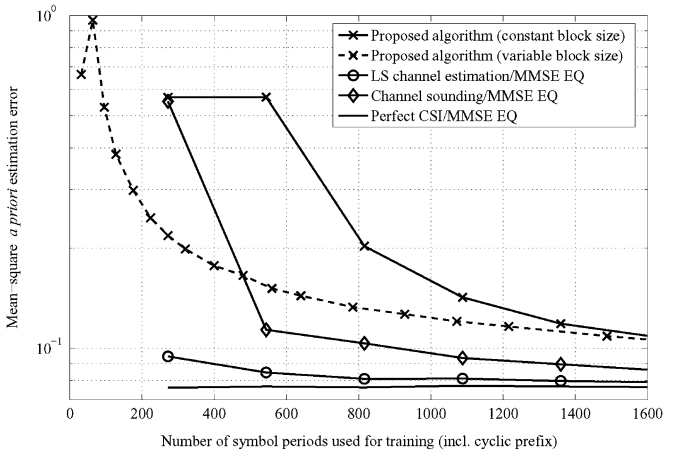

Fig. 6. Rate of convergence curves for FDE techniques employed in a 2×2 SC-FDE system operating in the ETSI BRAN E channel. (SNR per RX antenna $= 15$ dB, $K = 256, \mathcal{K} = \{16, 32, 64\,128\,256\}, Q = 16, \tau = 4$).

in Section V, is clearly visible from the dashed curves in these examples.

The benefits that can be gained by utilizing a variable training block size as opposed to a constant block size with the proposed algorithm are evident in Fig. 4. In this example, it is shown that for the 2×2 system operating in the ETSI BRAN A channel, the proposed algorithm with a variable block size and the channel sounding technique reach the same mean-square *a priori* estimation error after 280 training symbol intervals and 528 training symbol intervals, respectively. This difference corresponds to a 47% decrease in training overhead in favor of the proposed algorithm, although this decrease comes at the expense of an increase in complexity. Additionally, the difference in the estimation errors of these two methods after approximately 528 training symbol intervals is 0.6 dB. Not surprisingly, the LS technique converges very quickly, nearly reaching the reference curve after the first training block interval. A difference of approximately 1 dB is measured between the mean-square *a priori* estimation errors of the proposed algorithm and the LS technique at this point. Furthermore, the LS technique and the proposed algorithm require similar numbers of computations during the training process, as illustrated in Fig. 2. However, the proposed algorithm can be used to continually update the equalizer after training for a much smaller computational cost than is required by the LS technique, as depicted in Fig. 3.

Fig. 5 illustrates the rate-of-convergence curves for a 5×5 system operating in the ETSI BRAN A channel. It is important to note that although the proposed algorithm with a variable block size converges to the same estimation error floor as the channel sounding technique, it begins to converge to this point much more quickly. For example, the proposed algorithm and the channel sounding technique converge to the same estimation error after 1320 training symbol intervals. However, the proposed algorithm reaches within approximately 0.6 dB of this value after only 624 symbol intervals and within 0.3 dB of this value after 840 training symbol intervals.

The rate-of-convergence curves for a 2×2 system operating in the ETSI BRAN E channel are illustrated in Fig. 6. From this figure, it is observed that although the proposed algorithm with a variable block size converges, it would be more economical in terms of complexity to let $\kappa = K$ for all training blocks. Specifically, the proposed algorithm with a variable block size converges after 16 blocks in this example, whereas only 6 blocks are required to reach the same point if a constant block size is used. Furthermore, the proposed algorithm requires nearly an order of magnitude fewer computations to train the equalizer than the LS technique in this case, according to Fig. 2.

Figs. 4–6 suggest that $\kappa_{\theta-1}$ can be much less than K for channels with large coherence bandwidth. Indeed, the length of the ETSI BRAN A CIR is 390 ns. Therefore, the coherence bandwidth of the channel is approximately 2.6 MHz, which is much greater than the bandwidth of 78 kHz occupied by each of the 256 frequency bins used during the equalization process. In fact, Fig. 4 shows the algorithm converging after 16 training blocks. With $\tau = 4$ and $\mathcal{K} = \{8, 16, 32, 64, 128\}$, the 16th block interval used for training in this example comprises 64 training symbol intervals. Similarly, Fig. 5 depicts the algorithm converging after 24 training blocks when, again, $\kappa = 64$. However, Fig. 6 shows the proposed algorithm converging after 16 training blocks, at which point, $\kappa = 128$. A larger value of κ is required in this example due to the smaller coherence bandwidth of the ETSI BRAN E channel.

As a final note on the convergence curves depicted in Figs. 4–6, observe that the curves related to the proposed algorithm appear to diverge in the first few training blocks; however, this divergence reverses after n_T training blocks are received. This behavior is simply due to the underdetermined nature of the adaptive system prior to the reception of an adequate number of training blocks, from which point, the system becomes determined and can be solved. In addition, it should be noted that the sudden jump in each of the channel sounding curves occurs after n_T training blocks, which is simply due to the fact that all required training information has been received at this point, and a full channel estimate can be made.

B. Packet Error Rate Results

The PERs were simulated for four systems. Each system has $n_T = 2$ transmit antennas and $n_R = 2$ receive an-

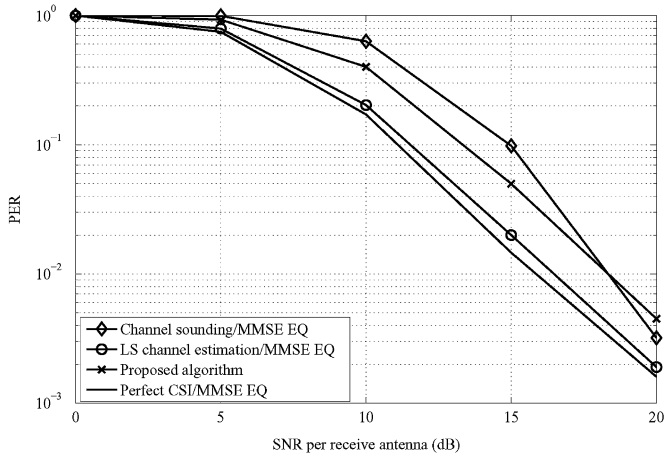

Fig. 7. PER curves for 2×2 SM SC-FDE systems operating in the ETSI BRAN A channel.

tennas. The first system devotes two block intervals to channel sounding. The second system uses two training block intervals for LS channel estimation. The third system adaptively computes the equalizer matrix with 15 training blocks, where $\mathcal{K} = \{8, 16, 32, 64\}$, and $\tau = 4$. The fourth system is assumed to have perfect knowledge of the channel. Each of the three nonadaptive systems employ a linear MMSE equalizer at the receiver. The ETSI BRAN A model was used to simulate the channel, where each tap was faded according to a Rayleigh distribution, and 10 000 independent channel realizations were constructed for each SNR point.

For each system, once the channel is estimated and the equalizer is trained, a packet of 1024 data bits is encoded with a half-rate convolutional encoder with generator polynomials $g_1(D) = D^6 + D^4 + D^3 + D + 1$, and $g_2(D) = D^6 + D^5 + D^4 + D^3 + 1$ [21]. The encoded bits are randomly interleaved and mapped to quadrature phase shift keying (QPSK) symbols that are then arranged into blocks, each with $K = 256$ symbols. A cyclic prefix of $Q = 8$ symbols is added to each block prior to transmission. At the receiver, the equalized symbols are mapped to soft bits that are then deinterleaved and passed through a standard Viterbi decoder to obtain hard decoded data bits. The PERs of the four simulated systems are shown in Fig. 7. As observed, the system employing the proposed algorithm performs within 3 dB of the system with perfect channel knowledge.

The crossover of the curves for the proposed algorithm and the channel sounding technique results from the use of frequency-domain interpolation at high SNR. Since the performance of the channel sounding technique is primarily affected by noise, as SNR increases, the technique obviously performs better. However, the adaptive algorithm in this example is affected by noise and the error floor imposed by frequency-domain interpolation. Therefore, the crossover point can be viewed as the point where noise no longer affects the adaptive algorithm as much as frequency-domain interpolation. It can be argued, however, that as long as the system is operating below the desired PER when this crossover occurs, the fact that the channel sounding technique performs better than the adaptive algorithm is irrelevant. Furthermore, the adaptive algorithm can now be used to track variations in the channel, whereas the channel sounding technique cannot. Note that since tracking is performed on full-size data blocks, no frequency-domain interpolation is needed to update the equalizer once it has been initially trained. Thus, error floors are prevented by switching to decision-directed mode following equalizer training, as long as the data decisions that are fed back to update the equalizer are reliable.

IX. Conclusion

In this paper, we presented a new method for adaptively equalizing multiantenna SC transmissions in the frequency domain. We applied this algorithm to a high data rate SM system. Furthermore, we examined convergence properties of the algorithm and subsequently used these properties to develop a method by which the required training overhead can be reduced. We derived conditions for optimality in training sequence design and detailed one approach to meeting these conditions. The complexity of the proposed algorithm is significantly lower than that of a channel estimation technique based on the LS criterion when both methods operate in the decision-directed mode. This advantage in complexity is greater when the channel has a large excess delay spread.

We used computer simulations to examine the convergence of the algorithm and the PER of a system employing the proposed algorithm. The results of these simulations show that the algorithm tends to converge very quickly in channels possessing a large coherence bandwidth. The results also show that the PER of an SM system employing the proposed algorithm is only 3 dB higher than the PER of a system that has perfect knowledge of the channel.

Acknowledgment

The authors wish to thank the anonymous reviewers for their helpful comments and the fruitful discussions with their collegues at the University of Bristol and the Toshiba Research Europe Limited: Telecommunications Research Laboratory.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] B. Holter, "On the Capacity of the Mimo Channel—A Tutorial Introduction," Dept. Telecommun., Norwegian Univ. Sci. Technol., Stavenger, Norway, Tech. Rep., 2001.

[3] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, vol. COM-19, no. 5, pp. 628–634, Oct. 1971.

[4] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 29–48, May 2000.

[5] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, First ed. Boston, MA: Artech House, 2000.

[6] T. Walzman and M. Schwartz, "Automatic equalization using the discrete frequency domain," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 1, pp. 56–68, Jan. 1973.

[7] H. Sari, G. Karam, and I. Jeanclaude, "Frequency-domain equalization of mobile radio and terrestrial broadcast channels," in *Proc. IEEE Global Telecommun. Conf.*, vol. 1, Nov. 1994, pp. 1–5.

[8] ——, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, Feb. 1995.

[9] A. Czylwik, "Comparison between adaptive OFDM and single carrier modulation with frequency domain equalization," in *Proc. IEEE Veh. Technol. Conf.*, vol. 2, May 1997, pp. 865–869.
[10] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.
[11] J. Coon, J. Siew, M. Beach, A. Nix, S. Armour, and J. McGeehan, "A comparison of MIMO-OFDM and MIMO-SCFDE in WLAN environments," in *Proc. IEEE Global Telecommun. Conf.*, vol. 6, Dec. 2003, pp. 3296–3301.
[12] A. Maleki-Tehrani, B. Hassibi, and J. M. Cioffi, "Adaptive equalization of multiple-input multiple-output (MIMO) channels," in *Proc. IEEE Int. Conf. Commun.*, vol. 3, 2000, pp. 1670–1674.
[13] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multiple-input multiple-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.
[14] M. V. Clark, "Adaptive frequency-domain equalization and diversity combining for broadband wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1385–1395, Oct. 1998.
[15] W. M. Younis, N. Al-Dhahir, and A. H. Sayed, "Adaptive frequency-domain equalization of space-time block-coded transmissions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, May 2002, pp. 2353–2356.
[16] S. Haykin, *Adaptive Filter Theory*, Third ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[17] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 2000.
[18] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461–471, Mar. 1999.
[19] *Channel Models for HIPERLAN/2 in Different Indoor Scenarios*, 1998. 3ER1085B.
[20] D. C. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 4, pp. 531–532, Jul. 1972.
[21] *Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) layer*, 2000.

Justin Coon (S'00) received the B.S. degree (with distinction) in electrical engineering from the Calhoun Honors College, Clemson University, Clemson, SC, in December 2000. In November 2001, he joined the Centre for Communications Research, University of Bristol, Bristol, U.K., from which he received the Ph.D. degree in wireless communications.

He then joined the Toshiba Telecommunications Research Laboratory, Bristol, where he is currently a research engineer. His primary research interests include low-complexity single-carrier and multicarrier techniques, channel estimation and tracking, multiple-antenna transmission, and signal processing in communications.

Simon Armour (M'01) received the B.Eng. degree from the University of Bath, Bath, U.K., in 1996 and the Ph.D. degree from the University of Bristol, Bristol, U.K., in 2001.

Following a period of post-doctoral research in the area of advanced WLAN technologies, he was appointed to the position of lecturer in software radio at the University of Bristol in 2001. His research interests include multicarrier modulation, WLANs/PANs, link adaptation and medium access control. He has published over 50 papers on these subjects.

Dr. Armour is a member of IEE.

Mark Beach (A'90) received the Ph.D. degree for research addressing the application of Smart Antennas to GPS from the University of Bristol, Bristol, U.K., in 1989, where he subsequently joined as a member of academic staff.

His interest in smart antenna techniques has continued with the application of dual array techniques, or Multiple-Input Multiple-Output architectures, to high-performance wireless networks. In particular, he has conducted research in the area of double-directional channel measurements and analysis. Furthermore, he has conducted novel research in the area of analog RF technologies for software-definable or reconfigurable radio. Examples include linearized mixers and tunable RF filters under E.U. projects such as TRUST and SCOUT. He currently holds the post of Professor of Radio Systems Engineering at the University of Bristol.

Dr. Beach is an active member of the IEE Professional Network on Antennas and Propagation.

Joe McGeehan received the B.Eng. and Ph.D. degrees in electrical and electronic engineering from University of Liverpool, Liverpool, U.K., in 1967 and 1971, respectively. In 2003, he received the D.Eng. degree from the University of Liverpool for his significant contribution to the field of mobile communications research.

He is presently Professor of communications engineering and Director of the Centre for Communications Research at the University of Bristol, Bristol, U.K. He is concurrently Managing Director of Toshiba Research Europe Limited: Telecommunications Research Laboratory, Bristol. He has been actively researching spectrum-efficient mobile radio communication systems since 1973 and has pioneered work in many areas, including linear modulation, linearized power amplifiers, smart antennas, propagation modeling/prediction using ray-tracing, and phase-locked loops.

Prof. McGeehan is a Fellow of both the Royal Academy of Engineering and the Institute of Electrical Engineers. He has served on numerous international committees and standards bodies and was advisor to the UK's first DTI/MOD "Defence Spectrum Review Committee" in the late 1970s. He was the joint recipient of the IEEE Vehicular Technology Transactions "Neal Shepherd Memorial Award" (for work on SMART Antennas) and the IEE Proceedings Mountbatten Premium (for work on satellite tracking and frequency control systems). In June 2004, he was made a Commander of the Order of the British Empire (CBE) in the Queen's Birthday Honors List for services to the Communications Industry.