Finite-Alphabet Based Channel Estimation for OFDM and Related Multicarrier Systems

Shengli Zhou, *Student Member, IEEE,* and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Novel blind channel estimators based on the finite alphabet property of information symbols are derived in this paper for OFDM and related multicarrier code-division multiple access (MC-CDMA) systems. The resulting algorithms are applicable not only to standard OFDM transmitters with cyclic prefix, but also to the recently proposed zero padded OFDM transmissions that improve symbol recovery at the expense of altering the transmitter and complicating the equalizer. Based on FFT-processed received data, channel identifiability is guaranteed regardless of channel zero locations and various channel estimation algorithms become available by trading off complexity for performance. Unlike existing blind channel estimators, the proposed alternatives require short data records especially for PSK transmissions. The inherent scalar ambiguity is easily resolved because it has unit amplitude and phase values drawn from a finite set. Decoupling channel from symbol estimation enables a phase-directed operation that improves upon decision-directed schemes that are known to suffer from error propagation. Practical issues are also addressed including the presence of frequency guard intervals, constellation and power loading, various frame designs, coded transmissions as well as semi-blind and online implementations for systems with training sequences. The algorithms are tested with simulations and also compared with existing alternatives in a realistic HIPERLAN/2 setting.**

***Index Terms*—Blind channel estimation, multicarrier transmission, OFDM.**

I. Introduction

HOLDING great promise for high rate transmissions, orthogonal frequency division multiplexing (OFDM) has been adopted and is currently being considered by many standards, including digital audio and video broadcasting (DAB, DVB) in Europe and high-speed DSL modems over twisted pairs in the United States [7]. OFDM has also been proposed for local area mobile wireless broad-band standards [23] including IEEE802.11a [6], MMAC, and HIPERLAN/2 [4].

In order to avoid intersymbol interference (ISI) arising due to channel memory, conventional OFDM systems first take the inverse Fourier transform (IFFT) of data symbols and then insert redundancy in the form of a cyclic prefix (CP) of length larger than the FIR channel order [1]. CP is discarded at the receiver and the remaining part of the OFDM symbol is FFT processed. A combination of IFFT and CP at the transmitter with the FFT at the receiver converts the frequency-selective channel to separate flat-fading subchannels (see, e.g., [1], [24]). Frequency-domain channel equalization is then applied by dividing the FFT output by the corresponding channel frequency response. However, symbol recovery is not assured if the channel has nulls on (or close to) some subcarriers. If Channel State Information (CSI) is available at the transmitter, channel nulls can be avoided by not transmitting symbols on those subcarriers as in DMT [19]. But CSI is not always available or it may be costly to acquire at the transmitter when the channel is time-varying. Recently, zero padding (ZP) has been proposed in [10], [20] to replace the CP, in order to guarantee (even blind) symbol recovery regardless of channel zeros at the cost of modifying the transmitter and complicating slightly the equalizer.

Because they save bandwidth and are capable of tracking slow channel variations, blind channel estimation and equalization methods are well motivated as they avoid use of training sequences [13], [21]. A number of blind channel estimators have been developed for OFDM. Some of them are geared toward multicarrier systems without CP [8], while others are based on cyclostationarity [12], and subspace decompositions that exploit the CP or the ZP structure in OFDM transmissions [15], [20].

In this paper, we develop and analyze a new channel estimation method that is applicable to both CP- and ZP-OFDM transmissions and relies on the finite alphabet property of information-bearing symbols. Unlike subspace methods developed for CP-OFDM, channel identifiability is guaranteed here even when the channel has nulls on subcarriers. Although ZP-OFDM guarantees channel identifiability and symbol recovery, it is not applicable to existing OFDM standards where the CP is inserted. Moreover, subspace-based methods for both CP- and ZP-OFDM require collection of data records that are sufficiently long to render the data covariance matrix full rank [15], [20]. In contrast, PSK transmissions with the novel method enable channel estimation even from a single OFDM symbol at high SNR, which is also impossible for statistical methods such as the one in [12]. Reducing the size of required samples equips the proposed method with the ability to track even fast channel variations. Furthermore, by exploiting the finite alphabet property, the scalar ambiguity (inherent to all blind algorithms) is restricted to have unit amplitude and phase values belonging to a finite set, which can be resolved easily. Practical issues such as presence of frequency guard bands, constellation and power loading, variable frame designs, coded transmissions, and application to semi-blind scenarios are also addressed.

Paper approved by C. Tellambura, the editor for Multicarrier Systems of the IEEE Communication Society, Manuscript received March 10, 2000; revised December 24, 2000. This work was supported by the National Science Foundation Wireless Initiative under Grant 99-79443 and the ARL under Grant DAAL01-98-Q-0648. This paper was presented in part at the Conference on Information Sciences and Systems (CISS), Princeton, NJ, March 15-17, 2000.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: szhou@ece.umn.edu; georgios@ece.umn.edu).

Publisher Item Identifier S 0090-6778(01)06939-2.

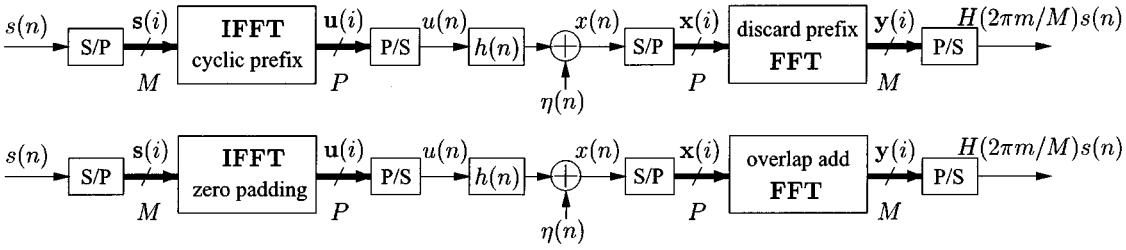


Fig. 1. Transmitter/channel/receiver for CP-OFDM (top) and ZP-OFDM (bottom).

II. Matrix Formulation of CP- and ZP-OFDM

The block diagram in Fig. 1 (top) describes the conventional CP-OFDM system. First, the information stream $s(n)$ is parsed into blocks: $\mathbf{s}(i):=[s(iM), s(iM+1), \ldots, s(iM+M-1)]^{\mathrm{T}}$ of length M, where the block index $i=\lfloor n/M \rfloor$ and $\lfloor\ \rfloor$ denotes integer floor. For each symbol index n, we can write $n=iM+m$, with $m \in [0, M-1]$ henceforth indicating the position inside the symbol block $\mathbf{s}(i)$ which constitutes the ith (so called) OFDM symbol. The $M \times 1$ symbol block $\mathbf{s}(i)$ is then mapped to a block $\mathbf{u}(i):=[u(iP), u(iP+1), \ldots, u(iP+P-1)]^{\mathrm{T}}$ of length $P > M$ by taking the M-point IFFT of $\mathbf{s}(i)$ and replicating the last $P-M$ elements in the front to form the CP.

Let $\mathbf{I}_{\mathrm{cp}}$ denote the last $P-M$ rows of the $M \times M$ identity matrix $\mathbf{I}_M$ and $\mathbf{F}_M$ be the $M \times M$ FFT matrix with (m,n) entry $(1/\sqrt{M})\exp(-j2\pi mn/M)$. We describe the CP insertion using the $P \times M$ transmit-matrix $\mathbf{T}_{\mathrm{cp}}:=\left[\mathbf{I}_{\mathrm{cp}}^{\mathrm{T}}, \mathbf{I}_M^{\mathrm{T}}\right]^{\mathrm{T}}$. Matrix $\mathbf{T}_{\mathrm{cp}}$ along with the IFFT matrix $\mathbf{F}_M^{\mathcal{H}}$ allows one to express the ith transmitted block as: $\mathbf{u}(i)=\mathbf{T}_{\mathrm{cp}}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}(i)$, where $^{\mathcal{H}}$ denotes Hermitian transpose.

The resulting linearly precoded "OFDM chip sequence" $u(n)$ is then pulse-shaped (not shown in Fig. 1) to the corresponding continuous time signal $u_c(t)=\sum_{n=-\infty}^{\infty}u(n)\varphi_c(t-nT)$, where T is the chip period and $\varphi_c(t)$ is the chip pulse (subscript $_c$ denotes continuous-time signals). The transmitted waveform $u_c(t)$ propagates through a (possibly *unknown*) dispersive channel $h_c(t)$ and is filtered by the receive filter $\bar{\varphi}_c(t)$, that is matched to $\varphi_c(t)$ and has spectrum $|\Phi(f)|^2$ with Nyquist characteristics and bandwidth $B \geq 1/(2T)$. Let $\Gamma_{\varphi\bar{\varphi}}(t)$ be the convolution of transmit with receive filters, and with $\star$ denoting convolution let $h(l):=(\varphi_c \star h_c \star \bar{\varphi}_c)(t)|_{t=lT}=\int_{-\infty}^{\infty}h_c(\tau)\Gamma_{\varphi\bar{\varphi}}(lT-\tau)d\tau$ be the equivalent discrete time channel impulse response. The received signal $x_c(t)$ sampled at the chip rate can then be written as

$$\begin{aligned} x(n):=x_c(t)|_{t=nT} &= h(n)\star u(n)+\eta(n) \\ &= \sum_{l=0}^{L}h(l)u(n-l)+\eta(n) \end{aligned} \tag{1}$$

where $\eta(n):=\eta_c(t)|_{t=nT}$ is filtered additive Gaussian noise (AGN) and L is the order of the FIR channel. To avoid ISI, we select the CP length to be larger than the channel order and state it formally as

a0) $P-M \geq L$.

To cast (1) from a serial to a convenient matrix-vector form, we define the $P \times 1$ vectors: $\mathbf{x}(i):=[x(iP), x(iP+1), \ldots, x(iP+P-1)]^{\mathrm{T}}$ and $\eta(i):=[\eta(iP), \eta(iP+1), \ldots, \eta(iP+P-1)]^{\mathrm{T}}$. Let $\mathbf{H}_0$ be the $P \times P$ lower triangular Toeplitz matrix with first column $[h(0), \ldots, h(L), 0, \ldots, 0]^{\mathrm{T}}$, and $\mathbf{H}_1$ be the $P \times P$ upper triangular Toeplitz matrix with first row $[0, \ldots, 0, h(L), \ldots, h(1)]$.

Relying on a0) and the fact that $h(l)=0, \forall\, l \notin [0, L]$, we can then write (1) as

$$\mathbf{x}(i)=\mathbf{H}_0\mathbf{u}(i)+\mathbf{H}_1\mathbf{u}(i-1)+\boldsymbol{\eta}(i) \tag{2}$$

where the second term denotes inter block (and thus inter-OFDM-symbol) Interference (IBI). At the receiver, the CP of length $P-M$ is removed first and FFT is performed on the remaining $M \times 1$ vector. In matrix form, this is accomplished with the receive-matrix $\mathbf{R}_{\mathrm{cp}}:=[0_{M\times(P-M)}, \mathbf{I}_M]$ which removes the first $P-M$ entries from a $P \times 1$ vector. According to a0), $\mathbf{R}_{\mathrm{cp}}\mathbf{H}_1=0$ removes IBI (and thus ISI) among OFDM symbols to obtain

$$\begin{aligned} \mathbf{y}_{\mathrm{cp}}(i) &= \mathbf{F}_M\mathbf{R}_{\mathrm{cp}}\mathbf{x}(i)=\mathbf{F}_M\mathbf{R}_{\mathrm{cp}}\mathbf{H}_0\mathbf{u}(i)+\mathbf{F}_M\mathbf{R}_{\mathrm{cp}}\boldsymbol{\eta}(i) \\ &= \mathbf{F}_M\mathbf{R}_{\mathrm{cp}}\mathbf{H}_0\mathbf{T}_{\mathrm{cp}}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}(i)+\tilde{\boldsymbol{\eta}}_{\mathrm{cp}}(i) \\ &= \mathbf{F}_M\tilde{\mathbf{H}}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}(i)+\tilde{\boldsymbol{\eta}}_{\mathrm{cp}}(i) \end{aligned} \tag{3}$$

where $\tilde{\boldsymbol{\eta}}_{\mathrm{cp}}(i):=\mathbf{F}_M\mathbf{R}_{\mathrm{cp}}\boldsymbol{\eta}(i)$ is the filtered noise vector and $\tilde{\mathbf{H}}=\mathbf{R}_{\mathrm{cp}}\mathbf{H}_0\mathbf{T}_{\mathrm{cp}}$ is the resulting channel matrix. It can be easily verified that $\tilde{\mathbf{H}}$ is an $M \times M$ circulant matrix with its (k,l)th entry given by $h((k-l)\mathrm{mod}M)$. Because $\tilde{\mathbf{H}}$ is a circulant matrix, it is well known that $\mathbf{F}_M\tilde{\mathbf{H}}\mathbf{F}_M^{\mathcal{H}}$ is a diagonal matrix $\mathbf{D}_H:=\mathrm{diag}[H(e^{j0}), H(e^{j(2\pi/M)}), \ldots, H(e^{j(2\pi/M)(M-1)})]$, where the diagonal elements are values of the channel frequency response $H(z):=\sum_{l=0}^{L}h(l)z^{-l}$ evaluated at the subcarriers $z=e^{j(2\pi/M)m}$ for each $m \in [0, M-1]$ (see e.g., [11, p. 202] for a proof). Therefore, we can rewrite (3) as

$$\mathbf{y}_{\mathrm{cp}}(i)=\mathbf{D}_H\mathbf{s}(i)+\tilde{\boldsymbol{\eta}}_{\mathrm{cp}}(i). \tag{4}$$

An alternative approach to removing IBI is to replace the CP insertion by zero-padding, as shown at the bottom of Fig. 1 (see also [10], [20]). To implement zero padding (ZP) at the transmitter, the mapping matrix must be changed from $\mathbf{T}_{\mathrm{cp}}$ to $\mathbf{T}_{\mathrm{zp}}:=\left[\mathbf{I}_M^{\mathrm{T}}, \mathbf{0}_{M\times(P-M)}^{\mathrm{T}}\right]^{\mathrm{T}}$. Note that IBI is canceled now at the transmitter because $\mathbf{H}_1\mathbf{T}_{\mathrm{zp}}=\mathbf{0}$. Therefore, $\mathbf{u}(i)=\mathbf{T}_{\mathrm{zp}}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}(i)$, and

$$\begin{aligned} \mathbf{x}(i) &= \mathbf{H}_0\mathbf{u}(i)+\mathbf{H}_1\mathbf{u}(i-\mathbf{1})+\boldsymbol{\eta}(i) \\ &= \mathbf{H}_0\mathbf{T}_{\mathrm{zp}}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}(i)+\boldsymbol{\eta}(i). \end{aligned} \tag{5}$$

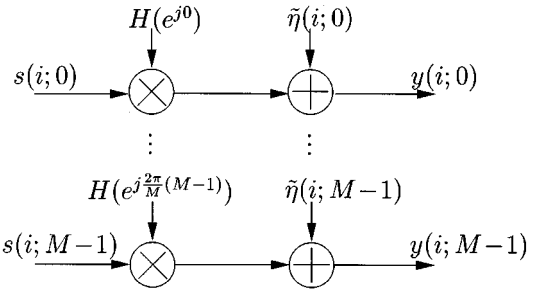


Fig. 2. Equivalent model with parallel flat-fading subchannels.

Because the matrix $\mathbf{H}_0\mathbf{T}_{\rm zp}\mathbf{F}_M^{\mathcal{H}}$ has full rank, time domain zero-forcing (ZF) equalization can be applied to recover the transmitted data as detailed in [20]; an efficient implementation of the ZF equalizer was proposed in [17] which amounts to inverting a $P\times P$ diagonal matrix and performing an P-point FFT afterwards. Furthermore, ZP-OFDM can also adopt an FFT-based low complexity receiver as CP-OFDM [16]. Indeed, if we let $\mathbf{I}_{\rm zp}$ denote the first $P-M$ columns of $\mathbf{I}_M$, the matrix $\mathbf{R}_{\rm zp}:=[\mathbf{I}_M,\mathbf{I}_{\rm zp}]$ maps a $P\times 1$ vector to an $M\times 1$ vector by adding the last $P-M$ elements to its first $P-M$ elements (matrix implementation of the overlap-add operation in block convolution). After performing FFT on $\mathbf{R}_{\rm zp}\mathbf{x}(i)$, we obtain

$$\begin{aligned}\mathbf{y}_{\rm zp}(i)&=\mathbf{F}_M\mathbf{R}_{\rm zp}\mathbf{x}(i)\\&=\mathbf{F}_M\mathbf{R}_{\rm zp}\mathbf{H}_0\mathbf{T}_{\rm zp}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}(i)+\mathbf{F}_M\mathbf{R}_{\rm zp}\boldsymbol{\eta}(i).\end{aligned}\quad(6)$$

By defining $\tilde{\boldsymbol{\eta}}_{\rm zp}(i):=\mathbf{F}_M\mathbf{R}_{\rm zp}\boldsymbol{\eta}(i)$ and verifying that

$$\mathbf{R}_{\rm zp}\mathbf{H}_0\mathbf{T}_{\rm zp}=[\mathbf{I}_M,\mathbf{I}_{\rm zp}]\mathbf{H}_0\left[\mathbf{I}_M^{\rm T},\mathbf{0}^{\rm T}\right]^{\rm T}=\tilde{\mathbf{H}}=\mathbf{T}_{\rm cp}\mathbf{H}_0\mathbf{R}_{\rm cp}$$

we can simplify (6) to

$$\mathbf{y}_{\rm zp}(i)=\mathbf{F}_M\tilde{\mathbf{H}}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}(i)+\tilde{\boldsymbol{\eta}}_{\rm zp}(i)=\mathbf{D}_H\mathbf{s}(i)+\tilde{\boldsymbol{\eta}}_{\rm zp}(i).\quad(7)$$

Equations (4) and (7) reveal that CP-OFDM and ZP-OFDM yield the same equivalent channel model shown in Fig. 2, and both convert the FIR convolutive channel to parallel flat fading subchannels. Through the transmit–receive mappers $(\mathbf{T}_{\rm cp},\mathbf{R}_{\rm cp})$ and $(\mathbf{T}_{\rm zp},\mathbf{R}_{\rm zp})$ introduced herein, it is transparent that both approaches turn linear convolution into circular convolution that can be efficiently implemented via FFTs.

The subspace-based channel estimation methods proposed for ZP-OFDM in [20] and for CP-OFDM in [15] utilize structure information of *pre*-FFT-processed data $\mathbf{x}(i)$. In contrast, our channel estimation methods will rely on *post*-FFT-processed data $\mathbf{y}_{\rm cp}(i)$ of (4) or $\mathbf{y}_{\rm zp}(i)$ of (7); thus, they apply to both CP-OFDM and ZP-OFDM. For uniformity, we subsequently drop the subscripts cp/zp and express (4) and (7) in scalar form as

$$y(i;m)=H(\rho_m)s(i;m)+\tilde{\eta}(i;m),\quad m\in[0,M-1]\quad(8)$$

where we assign a double argument $(i;m)$ to quantities with argument $n=iM+m$, and introduce $\{\rho_m:=e^{j(2\pi/M)m}\}_{m=0}^{M-1}$ for notational brevity. Per subcarrier m, the channel is a complex constant whose blind recovery from the data $y(i;m)$ is reminiscent of phase and carrier estimation in nondata-aided synchronization problems. In fact, our blind channel estimators for PSK modulated OFDM transmissions entail second and higher order powers of $y(i;m)$, and are inspired by well known squaring and higher order synchronization loops, see, e.g., [18, ch. 6]

III. Blind Channel Identifiability and Estimation

We will start with channel identifiability issues from the noiseless version of (8): $y(i;m)=H(\rho_m)s(i;m)$, where we omitted noise since we are concerned with basic feasibility questions first. Our results will be developed under the following conditions that hold true in our OFDM setup:

a1a) Symbols are drawn from a finite alphabet set of size Q; i.e., $s(i;m)\in\{\zeta_q\}_{q=1}^Q$;

a1b) Symbols are equiprobable, i.e., $\Pr(s(i;m)=\zeta_q)=1/Q$, for $q=1,\ldots,Q$.

a2) Noise $\tilde{\eta}(i;m)$ is zero-mean complex circular Gaussian and independent of $s(i;m)$.

Assumption a1a) implies that the following equation holds true: $\prod_{q=1}^Q[s(i;m)-\zeta_q]=0$. Expanding the product yields a Qth-order polynomial in $s(i;m)$

$$s^Q(i;m)+\alpha_1 s^{Q-1}(i;m)+\cdots+\alpha_Q=0\quad(9)$$

where $\alpha_1,\ldots,\alpha_Q$ are determined by the constellation points $\{\zeta_q\}_{q=1}^Q$. The coefficients $\{\alpha_n\}_{n=1}^Q$ cannot all be zeros, otherwise $s^Q(i;m)\equiv 0$ as evidenced by (9). Therefore, we can always find nonzero elements from the coefficient set $\{\alpha_n\}_{n=1}^Q$; let us define J to be the smallest index of such nonzero coefficients, i.e.,

Definition 1: We define J such that $\alpha_J\neq 0$; $\alpha_n=0,\forall\, n<J$.

It then follows easily that $J\leq Q$. Let us check the value of J for widely used signal constellations. For a PSK constellation of size Q: $\{\zeta_q=\exp((j2\pi q)/(Q+j\pi/Q))\}_{q=1}^Q$, we have $J=Q$ and $\alpha_J=1$ since (9) reduces to $s^Q(i;m)+1=0$; therefore, $J=2$ for BPSK and $J=4$ for quadrature phase shift keying (QPSK) signals. For QAM with Q signaling points, e.g., $Q=16,32,64,128,256$, we can easily verify that $J=4$. For example, let us consider 16QAM with signaling points $\zeta_q\in\{\pm a\pm bj\}$ where $a,b\in\{1,3\}$; expansion of the product $\prod_{q=1}^Q[s(i;m)-\zeta_q]$ yields $\alpha_4=272\neq 0$. The facts that $J\leq Q$ and that for large-signal constellations (e.g., QAM) $J\ll Q$ will play a key role in the rest of our paper.

Under a1a) and a1b), we prove in [25, Appendix II] that $\mathrm{E}\{s^J(i;m)\}=-(J/Q)\alpha_J\neq 0$ for any signal constellation. We highlight here that PSK constellations have the nice property that $s^J(i;m)=-(J/Q)\alpha_J=-1$ deterministically for any $s(i;m)$. Thus, a1b) will not be necessary for PSK transmissions in our subsequent derivations.

Starting from $\mathrm{E}\{y^J(i;m)\}=H^J(\rho_m)\mathrm{E}\{s^J(i;m)\}$, we deduce that

$$H^J(\rho_m)=\frac{-Q}{J\alpha_J}\mathrm{E}\left\{y^J(i;m)\right\},\quad\forall\, m\in[0,M-1].\quad(10)$$

For each m, the right-hand side of (10) is available from the FFT processed data either via statistical averaging (across i) or deterministically for PSK. Hence, $H^J(\rho_m)$ can be determined and the question is how one uniquely recovers $H(\rho_m)$ from $H^J(\rho_m)$. To express $H^J(\rho_m)$ in terms of

$h(l)$, we first define $\boldsymbol{\beta}_J^{\mathrm{T}} := [\beta_0, \ldots, \beta_{JL}] = h(l) \star_J h(l)$ the J-fold convolution of $h(l)$ with itself that we denote as $\star_J$. Since time-domain convolution corresponds to frequency-domain multiplication, it follows that: $H^J(\rho_m) := H^J(z)|_{z=\rho_m} := \left(\beta_0 + \cdots + \beta_{JL} z^{-JL}\right)|_{z=\rho_m}$. To determine the coefficients $\beta_0, \ldots, \beta_{JL}$ uniquely from $H^J(\rho_m)$, we need $JL+1$ distinct such equations that become available when ρ_m takes $JL+1$ distinct values. Therefore, selecting block length $M \geq JL+1$ allows one to find $\beta_0, \ldots, \beta_{JL}$ uniquely, which enables determination of $H^J(z) = \sum_{i=0}^{JL} \beta_i z^{-i}$ on the entire complex plane. Observing that $H^J(z)$ contains all the roots of $H(z)$ with multiplicity J, we can extract from the roots of $H^J(z)$ those L roots that form $H(z)$. This implies that $H(z)$ is uniquely identifiable up to a scalar ambiguity, as we summarize in the following.

Proposition 1: If the length of a symbol block M (which is a system design parameter) is selected to satisfy $M \geq JL+1$, then identifiability of the L th-order channel $H(z)$ is guaranteed from the demodulated data $y(i;m) = H(\rho_m)s(i;m)$ for any signal constellation under assumptions a1a) and a1b) with J specified as in Definition 1. Especially for PSK constellations, a1b) is not needed and thus channel identifiability is guaranteed even based on a single OFDM block.

As stated in Proposition 1, blind channel identifiability is guaranteed regardless of the channel zero locations. This nice property is not possessed by the subspace-based method derived in [15], [14] for CP-OFDM, where the proposed subspace-based channel estimator does not guarantee channel identifiability for those channels with nulls on some subcarriers. Without assumption a1b) for arbitrary constellations, a more general proof of Proposition 1 is provided in [25].

Having established channel identifiability, we now develop channel estimation algorithms in the presence of additive noise. The circularity of $\tilde{\eta}(i;m)$ in a2) implies that $\mathrm{E}\{\tilde{\eta}^n(i;m)\} = 0$, $\forall\, n > 0$, regardless of the noise color. Note also that $\tilde{\eta}(i;m)$ is complex circular Gaussian even with BPSK because it models complex noise after FFT processing. With J as in Definition 1, a2) implies that

$$\begin{aligned} \mathrm{E}\left\{y^J(i;m)\right\} &= \mathrm{E}\left\{[H(\rho_m)s(i;m) + \tilde{\eta}(i;m)]^J\right\} \\ &= H^J(\rho_m)\mathrm{E}\left\{s^J(i;m)\right\} \end{aligned} \tag{11}$$

and hence (10) still holds true. In practice, $\mathrm{E}\left\{y^J(i;m)\right\}$ is replaced by consistent sample averages and thus $H^J(\rho_m)$ is estimated as

$$\hat{H}^J(\rho_m) = \frac{-Q}{J\alpha_J}\left(\frac{1}{I}\sum_{i=0}^{I-1} y^J(i;m)\right), \quad m \in [0, M-1] \tag{12}$$

where I is the total number of blocks averaged. We next collect $\hat{H}^J(\rho_m)$ from (12) in an $M \times 1$ vector:[1] $\hat{\hat{\mathbf{H}}}_J := [\hat{H}^J(\rho_0), \ldots, \hat{H}^J(\rho_{M-1})]^{\mathrm{T}}$. Define the matrix $\mathbf{V}_J$ to be a scaled version of the first $JL+1$ columns of FFT matrix $\mathbf{F}_M$ as follows: $\mathbf{V}_J := \sqrt{M}\mathbf{F}_M(:, 1{:}JL+1)$, where Matlab's notation $\mathbf{A}(:, i{:}k)$ is used to denote a submatrix formed by the ith to kth columns of matrix $\mathbf{A}$. Because $\hat{\hat{\mathbf{H}}}_J = \mathbf{V}_J\boldsymbol{\beta}_J$ describes the $\mathcal{Z}$-transform of $\boldsymbol{\beta}_J$ evaluated at $\{z = \rho_m\}_{m=0}^{M-1}$, the vector $\boldsymbol{\beta}_J$ can be estimated by a simple matrix inversion († denotes matrix pseudo-inverse) as

$$\hat{\boldsymbol{\beta}}_J = \mathbf{V}_J^{\dagger}\hat{\hat{\mathbf{h}}}_J = \frac{1}{M}\mathbf{V}_J^{\mathcal{H}}\hat{\hat{\mathbf{h}}}_J \tag{13}$$

where $\mathbf{V}_J^{\dagger} = (1/M)\mathbf{V}_J^{\mathcal{H}}$ can be easily checked. Relying on the FFT-processed data $y(i;m)$ in (8) and capitalizing on the finite alphabet of $s(i;m)$, (12) and (13) provide sufficient information to recover the channel vector $\mathbf{h} := [h(0), \ldots, h(L)]^{\mathrm{T}}$.

Starting from time domain vector (13) only, and based on the fact that $\boldsymbol{\beta}_J$ has all L roots of $\mathbf{h}$ with multiplicity J, we can search over JL noisy roots of $\hat{\boldsymbol{\beta}}_J$ and find that combination of L roots that matches $\mathbf{h}$ best (this corresponds to the Root Selection (RS) algorithm in [25]). However, complexity of the RS increases fast as J, L increase, and, this nonlinear algorithm will exhibit sensitivity when roots are to be found in the presence of noise. It is possible however, to derive a low-complexity linear equation (LE)-based channel estimator by observing that $\boldsymbol{\beta}_J$ is the J-fold linear convolution of $\mathbf{h}$ [25]. Although very simple to implement, the LE method is prone to noise-induced error propagation when $|h(0)|$, or $|h(L)|$, is small. Therefore, the LE method is only well motivated for wired-line applications where the strong path is synchronized to be the first path. Next, we will describe channel estimators starting from the frequency domain estimates in (12).

A. Minimum-Distance Algorithms

Given $\hat{H}^J(\rho_m)$ estimates from (12), we develop here two Minimum Distance (MD) algorithms. For each $m \in [0, M-1]$, we have $\hat{H}(\rho_m) = \lambda_m[\hat{H}^J(\rho_m)]^{1/J}$, where $\lambda_m \in \{e^{j(2\pi/J)n}\}_{n=0}^{J-1}$ is the corresponding scalar ambiguity in taking the Jth root. To resolve these ambiguities, we exhaustively search over all J^M possible vectors $\hat{\hat{\mathbf{h}}}_1 := [\lambda_0[\hat{H}^J(\rho_0)]^{1/J}, \ldots, \lambda_{M-1}[\hat{H}^J(\rho_{M-1})]^{1/J}]^{\mathrm{T}}$. For each $\hat{\hat{\mathbf{h}}}_1$, we compute the corresponding time domain vector through Least Squares fitting: $\hat{\mathbf{h}}_1 = \mathbf{V}_1^{\dagger}\hat{\hat{\mathbf{h}}}_1 = (1/M)\mathbf{V}_1^{\mathcal{H}}\hat{\hat{\mathbf{h}}}_1$, where $\mathbf{V}_1$ is a scaled version of the first $L+1$ columns of FFT matrix $\mathbf{F}_M$: $\mathbf{V}_1 := \sqrt{M}\mathbf{F}_M(:, 1{:}L+1)$. The matrix $\mathbf{V}_1$ relates channel frequency response values with time domain channel coefficients, and $\mathbf{V}_1^{\dagger} = (1/M)\mathbf{V}_1^{\mathcal{H}}$ can be easily verified. Channel estimates are then found by minimizing the Euclidean distance

$$\hat{\mathbf{h}} = \arg\min_{\hat{\mathbf{h}}_1} \|\hat{\boldsymbol{\beta}}_J - \hat{\mathbf{h}}_1 \star_J \hat{\mathbf{h}}_1\|. \tag{14}$$

Instead of (14), a joint maximum likelihood (ML) channel and symbol estimation could be pursued based on the noisy data in (8), by searching for the best fit over all possible symbol and channel combinations. Because both ML and MD resort to exhaustive search, it is of interest to compare them. If we collect I blocks (each of length M), the complexity of such an ML estimator for a Q-ary constellation is proportional to Q^{IM}, while the complexity of the MD in (14) reduces to only J^M by exploiting the known signal constellation to eliminate the unknown symbols $s(i;m)$. For PSK constellations, where $J = Q$, the ML can have the same complexity with MD (J^M) only if it operates on a single block instead of all available I blocks.

[1] In our following derivations, a vector with "^" indicates that its entries are collected from frequency domain channel response.

In this case however, the MD will perform considerably better than ML because it decreases additive noise through block averaging [cf. (12)], when compared to the ML that relies on the scalar noisy samples of (8). Although less complex than ML, the MD approach is still too complex to be implemented because often $M \gg L$ in practical systems. Nevertheless, MD is always useful as a bound to benchmark performance of simpler channel estimators.

Because the channel order $L < M$, the MD method does not require the entire vector $\hat{\tilde{\mathbf{h}}}_1$ to obtain the time-domain channel estimate $\hat{\tilde{\mathbf{h}}}_1$. This observation motivates the following modified Minimum Distance (MMD) algorithm:

1) Select $\bar{N}$ elements from $\hat{\tilde{\mathbf{h}}}_1$ and form a new $\bar{N} \times 1$ vector $\bar{\tilde{\mathbf{h}}}$. For $J^{\bar{N}}$ possible $\bar{\tilde{\mathbf{h}}}$, we obtain the time domain $\hat{\mathbf{h}}_1 = \bar{\mathbf{V}}^{\dagger}\bar{\tilde{\mathbf{h}}}$, where $\bar{\mathbf{V}}$ is constructed from $\mathbf{V}_1$ by keeping only the corresponding $\bar{N}$ rows. The only requirement on $\bar{N}$ is to satisfy $\bar{N} \geq L+1$, which guarantees that $\bar{\mathbf{V}}^{\dagger}$ exists. We can choose those $\bar{N}$ elements in $\hat{\tilde{\mathbf{h}}}_1$ with largest absolute values, or simply focus on $\bar{N}$ equispaced subcarriers.
2) Obtain channel estimates as in (14).

The complexity is now reduced from J^M to $J^{\bar{N}} \geq J^{L+1}$, rendering the MMD algorithm affordable for low order channels. Presence of (say N_p) pilot tones will reduce the complexity of MMD further. Indeed, as we will describe later on, we can determine the phases λ_m on those pilot carriers [cf. (15)]. Including those frequency response in $\bar{\tilde{\mathbf{h}}}$, we only need to search over $J^{\bar{N}-N_p}$ possible choices. The complexity with the help of pilot tones is $1/J^{N_p}$ of that without pilot tones, which makes the MMD applicable to longer channels with the same complexity.

B. Phase Directed (PD) Algorithm

As described in the MD approach, for each m we obtain $\hat{H}(\rho_m) = \lambda_m[\hat{H}^J(\rho_m)]^{1/J}$ up to a scalar ambiguity $\lambda_m \in \{e^{j(2\pi/J)n}\}_{n=0}^{J-1}$. Suppose that initial estimates $\hat{H}_0(\rho_m)$ are available through the MMD algorithm, or, via the RS or LE algorithms in [25]. For each $m \in [0, M-1]$, we can resolve the phase ambiguity by searching over J candidate phase values

$$\hat{\lambda}_m = \arg\min_{\lambda_m} \left| \hat{H}_0(\rho_m) - \lambda_m \left[\hat{H}^J(\rho_m)\right]^{1/J} \right|^2. \quad (15)$$

Therefore, we can improve channel estimation accuracy through what we term Phase Directed (PD) steps that we describe next:

Step 1) Set $i_1 = 0$, find an initial estimate $\hat{\mathbf{h}}_0$ using low complexity methods, and calculate the frequency response $\hat{H}_0(\rho_m)$ for $m = 0, 1, \ldots, M-1$.

Step 2) In each successive iteration, set $i_1 := i_1 + 1$, and

a) Resolve phase ambiguities by replacing $\hat{H}_0(\rho_m)$ with $\hat{H}_{i_1-1}(\rho_m)$ in (15), and then form the vector $\hat{\hat{\mathbf{h}}}_1 := [\hat{\lambda}_0[\hat{H}^J(\rho_0)]^{1/J}, \ldots, \hat{\lambda}_{M-1}[\hat{H}^J(\rho_{M-1})]^{1/J}]^{\mathrm{T}}$;

b) Update channel estimates $\hat{\mathbf{h}}_{i_1} = \mathbf{V}_1^{\dagger}\hat{\hat{\mathbf{h}}}_1 = (1/M)\mathbf{V}_1^{\mathcal{H}}\hat{\hat{\mathbf{h}}}_1$, and their frequency response using $\hat{\hat{\mathbf{h}}}_{i_1} = \mathbf{V}_1\hat{\mathbf{h}}_{i_1}$.

Step 3) Repeat S2 several (say I_1) times, or, continue until $\hat{\mathbf{h}}_{i_1} \approx \hat{\mathbf{h}}_{i_1-1}$ in the Euclidean norm sense.

Note that each iteration of Step 2) entails one M-point inverse FFT and one M-point FFT; indeed, $\mathbf{V}_1^{\mathcal{H}}\hat{\hat{\mathbf{h}}}_1$ amounts to performing an M-point inverse FFT on $\hat{\hat{\mathbf{h}}}_1$ and truncating the output by keeping only the first $L+1$ entries; $\mathbf{V}_1\hat{\mathbf{h}}_{i_1}$ amounts to performing an M-point FFT on an $M \times 1$ vector formed after zero-padding $\hat{\mathbf{h}}_{i_1}$. Thus, the PD iterations have low computational complexity. It is also worth recalling that $J = 2$ for BPSK and hence $\lambda_m \in \{\pm 1\}$. For the widely used QPSK, 16QAM and 64QAM we have $J = 4$ so that $\lambda_m \in \{\pm 1, \pm j\}$. These few phase values can be resolved easily via (15) with high accuracy even with coarse initial channel estimates. Our experience with simulations supports that the PD algorithm achieves the best performance with only two or three iterations. Hence, PD is capable of performing close to the complex yet accurate MD algorithm with much lower complexity, as will be illustrated in our simulations.

An alternative means of capitalizing on the finite-alphabet is the decision-directed (DD) algorithm [13], [9], [25]. Based on initial channel estimates $\{\hat{H}_0(\rho_m)\}_{m=0}^{M-1}$, the DD algorithm consists of equalizing (8) using $\hat{s}(i;m) = y(i;m)/\hat{H}_0(\rho_m)$, and then projecting the symbol estimates onto the finite alphabet. New channel estimates are then obtained by treating the resulting symbol estimates as known symbols, and the process of alternating between channel and symbol estimation steps is applied repeatedly until possible convergence.

Serving a similar purpose, the PD algorithm outperforms the DD iteration for two reasons.

1) The DD algorithm diverges easily at *low SNR* because symbol by symbol detection has poor performance. The smallest Bit Error Rate (BER) will be that corresponding to known channels. To lower this BER level, Forward Error Correction was considered in [13], where the Viterbi decoder output was re-encoded to offer symbol estimates at the expense of large decoding delay and computational complexity. In contrast, the PD algorithm reduces noise effects (and thus channel estimation error) to arbitrarily low levels by averaging over more and more blocks. Therefore, channel estimation accuracy improves consistently.
2) DD performance degrades as the *constellation size* Q increases because the minimum Euclidean distance among constellation points decreases and thus symbol decisions become less reliable at the same SNR. In the PD algorithm however, $J = 4 \ll Q$ for all QAM signals, and the same number of possible phase values $\{\pm 1, \pm j\}$ is present regardless of the constellation size.

The main difference between PD and DD is that DD alternates between channel estimation and symbol detection while PD avoids symbol estimation by decoupling channel estimation from symbol recovery. Therefore, PD is immune to the well-known error propagation phenomenon that is present in DD iterations and is responsible for catastrophic "run-away" effects at low SNR.

The price paid by PD is a certain amount of "mismatch noise" that arises due to finite sample effects when ensemble quantities in (10) are replaced by sample averages in (12). This mismatch is not present with PSK signals because $s^J(i;m) = -1$

(or 1 for BPSK) deterministically; i.e., ensemble averages coincide with sample averages. To elaborate further on this mismatch noise consider QAM signaling and noise-free reception, where the DD algorithm converges to the ML solution and perfect channel estimates are obtained. PD on the other hand, will be affected by the finite sample size and its channel estimation error curves will level-off (floor effect). With noise-free data, however, PD will yield accurate enough channel estimates, provided that sufficient blocks are used in the sample averaging. On the one hand, the error floor can be lowered to a prescribed level by collecting more and more blocks. On the other hand, it is even possible to eliminate the error floor by applying only one DD iteration after the PD iterations. This remedy though is only effective at high SNR values and for large-signal constellations, as will be illustrated in our simulations.

In addition to having low complexity and high channel estimation accuracy, the PD algorithm lends itself naturally to tracking slow channel variations. Among many possible window choices, we here only present the simplest rectangular shape where we form sliding-window (of size W) channel estimates by averaging only the W most recently received blocks $\{\mathbf{y}(i)\}_{i=I-W+1}^{I}$ in (12). Note that the window shape and size can be chosen to balance the trade-off between reduced-variance estimation and enhanced tracking capability. The resulting windowed channel update is

$$\hat{H}_{i+1}^{J}(\rho_m)=\hat{H}_{i}^{J}(\rho_m)+\frac{y^J(i+1;m)-y^J(i-W+1;m)}{\mathrm{E}\{s^J(i+1;m)\}}. \tag{16}$$

We can apply the PD algorithm to the new estimates $\{\hat{H}_{i+1}^{J}(\rho_m)\}_{m=0}^{M-1}$ and use the current $\hat{H}_{i}^{J}(\rho_m)$ as an initial estimate $\hat{H}_0(\rho_m)$ in (15). Combining channel estimates in (16) with the phase-resolving step of (15) equips PD with a low-complexity tracking capability.

PD advantages over DD have important practical ramifications. In a nutshell: either starting from a low-complexity blind channel estimator, or, by exploiting training sequences present in current OFDM standards as we will discuss in Section IV-E, the PD algorithm yields the best achievable MD performance with reduced complexity and is capable of tracking slow channel variations.

C. Distinct Features

We present next some unique characteristics of our blind channel estimators.

1) Convergence Issues and Noise Effects: As we pointed out earlier, our channel estimators do not rely on statistical properties of $s(i;m)$ for PSK transmissions, and thus symbols $s(i;m)$ do not have to be equiprobable as required by assumption a1b). It then follows that even *one* block is sufficient to yield reliable channel estimates at high SNR with guaranteed channel identifiability, which is never the case for other blind channel estimators, e.g., the cyclostationarity-based [12] or the subspace-based [15], [20] approaches. Recall that subspace-based methods need to collect enough blocks to render the data covariance matrix full rank: at least $2M$ blocks are needed for CP-OFDM [15] and at least M blocks for ZP-OFDM [20].

For constellations other than PSK, block averaging in (12) approximates the ensemble average $\mathrm{E}\{s^J(i;m)\}$ by the sample average $(1/I)\sum_{i=0}^{I-1}s^J(i;m)$. In this case, we need to collect a sufficient number of blocks to decrease the finite sample effects that are prevalent especially at high SNR, where additive noise can be omitted. Recall however, that $s^J(i;m)$ assumes a small number of values, e.g., four distinct values for 16QAM and 16 distinct values for 64QAM signaling. This alleviates the requirement on the number of blocks I.

Requiring short data records equips our channel estimators with the ability to track even rapid channel variations. At high SNR, since one block suffices for PSK-based channel estimation, the channel is allowed to vary even from block to block.

To appreciate how our blind channel estimators handle additive noise, it is of interest to compare them with benchmark methods that are based on training. To investigate the additive noise effects only and avoid the additional approximation error between the ensemble average $\mathrm{E}\{s^J(i;m)\}$ and its sample average approximation, we here only focus on PSK constellations. For one training block $\{s^J(i;m)\}_{m=0}^{M-1}$, we receive $y(i;m)=H(\rho_m)s(i;m)+\tilde{\eta}(i;m)$. The effective signal to noise ratio (SNR) for channel estimation is thus $|H(\rho_m)|^2\rho_s^2/\rho_{\tilde{\eta}}^2$, where $\rho_s^2(\rho_{\tilde{\eta}}^2)$ stands for signal (noise) power. For each block, our blind methods will first take the Jth power of the received FFT-processed data to obtain

$$\begin{aligned}y^J(i;m)=&H^j(\rho_m)s^J(i;m)\\&+J\tilde{\eta}(i;m)s^{J-1}(i;m)H^{J-1}(\rho_m)+\mathrm{O}(\tilde{\eta}^2(i;m)).\end{aligned} \tag{17}$$

Omitting the noise term $\mathrm{O}(\tilde{\eta})^2(i;m))$ at high SNR, the SNR pertinent to channel estimation is now $|H(\rho_m)|^2\rho_s^2/(J^2\rho_{\tilde{\eta}}^2)$, showing $10\log_{10}(J^2)$ dB loss relative to the training-based method if both of them use only a single block. However, this noise enhancement can be remedied by averaging across many (say I) blocks as per (12), which decreases the noise power and thus increases the SNR by $10\log_{10}(I)$ dB. Therefore, our method based on $I=4$ BPSK symbols and $I=16$ for QPSK approximates the channel estimation accuracy reachable by one training block. Relative to a practical system that relies on N_t training blocks ($N_t=2$ in HIPERLAN/2 [4]), our blind channel estimation method will catch up after the receiver accumulates about $I=J^2N_t$ blocks. Note that the training method can only use finite training blocks and the number of blocks N_t is usually fixed in practical systems, while our blind method can exploit the entire received data record (say I blocks). With only a few (about $I=N_tJ^2$) blocks, our blind method catches up with the training-based method (with N_t known training blocks), and outperforms it as more blocks become available ($I\gg N_tJ^2$).

Furthermore, the channel estimators herein do not require knowledge of the AGN color. In contrast, noise in [15] is assumed white, or, the noise covariance matrix must be estimated in order to pre-whiten the received data covariance matrix as suggested in [20]. Note that colored noise has to be accounted for DSL applications because strongly correlated near–end cross talk (NEXT) constitutes a major impairment [1]. In such

cases, the subspace approaches [20], [15] or second order statistics-based methods should calculate the noise color first which certainly increases complexity.

2) Scalar Ambiguity Determination: All blind estimators of a channel $\mathbf{h}$ come with an inherent (generally complex) scalar ambiguity, i.e., blind channel estimates $\hat{\mathbf{h}}$ satisfy: $\mathbf{h} = \alpha\hat{\mathbf{h}}$, where $\alpha \in \mathbb{C}$. Usually, blind estimators based on second-order statistics can also determine the amplitude of α (denoted by $|\alpha|$) and leave the phase of α (denoted by $\angle\alpha$) belonging to a infinite set: $\angle\alpha \in [0, 2\pi]$. To resolve this scalar (or phase) ambiguity, a few pilot tones are used in practice to determine α by comparing the estimated with the known frequency response on those pilot carriers. Because both channel estimates as well as pilot tones are contaminated by noise, such a matching may not be accurate. It can even be misleading if the channel response on those pilot carriers is not estimated accurately. However, because $H^J(z) = \hat{H}^J(z)$ in our approach, the scalar ambiguity in $\mathbf{h} = \alpha\hat{\mathbf{h}}$ must satisfy: $\alpha^J = 1$, which reveals that α has unit amplitude: $|\alpha| = 1$, and its phase belongs to a finite set: $e^{j\angle\alpha} \in \{e^{j(2\pi/J)n}\}_{n=0}^{J-1}$. As $J = 2$ or 4 for most practical signal constellations, the set is simply $\{\pm 1\}$ or $\{\pm 1, \pm j\}$. Thus, the task of recovering α is simplified considerably. Only one pilot symbol suffices to resolve this phase uncertainty by picking the correct value from a few choices.

Note that our simplified scalar determination can also be utilized by other blind channel estimation methods. For example, the subspace methods of [15] and [20] can first estimate the channel as $\hat{\mathbf{h}}_s$ based on structure information, which has a scalar ambiguity: $\mathbf{h} = \alpha_s\hat{\mathbf{h}}_s$. They can then exploit the finite alphabet property of the symbols and determine α_s^J by matching $\alpha_s^J\hat{\mathbf{h}}_s \star_J \hat{\mathbf{h}}_s$ with $\hat{\beta}_J$. Solving α_s as: $\alpha_s = \lambda_s(\alpha^J)^{1/J}$ yields J phase possibilities: $\lambda_s \in \{e^{j(2\pi/J)n}\}_{n=0}^{J-1}$, which suggest that in order to determine the scalar ambiguity α_s one needs to search λ_s only over J finite values. The complexity of this scalar LS fitting is negligible when compared to the SVD complexity for matrices of dimension $(2P - L) \times (2P - L)$ [15], or, $P \times P$ [20]. Therefore, by reducing the scalar ambiguity determination to a phase uncertainty search over a finite set, one can resolve the scalar ambiguity accurately using a few pilots even for other blind channel estimation methods [15], [20].

IV. Practical Considerations

In this section, we equip our channel estimation method with several practical features.

A. Frequency Guard Intervals

In all standardized OFDM systems, null carriers (zeros in $\mathbf{s}(i)$) are used to provide frequency guard against interference from adjacent OFDM systems. Similarly, certain subcarriers are discarded to avoid channel fades in DMT [19]. The resulting $\mathbf{s}(i)$ contains only K useful symbols and $M - K$ zeros which are distributed inside the OFDM symbol vector. Our method is directly applicable to this setting simply by considering the channel frequency response on the nonzero K subcarriers. As per Proposition 1, channel identifiability is still guaranteed when $K \geq JL + 1$.

The need for frequency guard-zeros is often ignored in the literature as was pointed out in [14]. For the subspace-based methods, frequency guard zeros call for increased dimensionality of the noise subspace because the correlation matrix of $\mathbf{R}_{ss} = \mathrm{E}\{\mathbf{s}(i)\mathbf{s}^{\mathcal{H}}(i)\}$ no longer has rank $2M$ but $2K$ [14]. However, when the channel happens to have nulls on subcarriers (whether used or not), the modified subspace-based method for CP-OFDM will still be unable to guarantee uniqueness in estimating channels with zeros located on subcarriers [14], [15].

B. Constellation and Power Loading

As we see in (12), frequency response estimates on different subcarriers are independent of each other. Therefore, information symbols transmitted on different subcarriers can be drawn from different signal constellations and each can be possibly loaded with different power [1]. Our algorithms are directly applicable to such cases by defining J to be the least common multiple of $J_0, J_1, \ldots, J_{M-1}$, where J_m is defined in Definition 1 for the specific signal constellation transmitted on the mth subcarrier. Recall that if we only load BPSK, QPSK, 16QAM, 64QAM signals, the overall J is still 4 and the complexity remains the same.

C. Variable Frame Designs

In addition to variable loading, the constellation may also vary along the data burst on each subcarrier. For example, one may wish to load constellations A and B onto the same carrier in a single data frame with frame length $I = I_a + I_b$, where I_a and I_b are the respective block lengths. For such a setting, our methods only need to replace (12) by

$$\hat{H}^J(\rho_m) = \frac{1}{\mathrm{E}\{s^J(i_a;m)\}}\left(\frac{1}{I_a}\sum_{i_a=0}^{I_a-1} y^J(i_a;m)\right) + \frac{1}{\mathrm{E}\{s^J(i_b;m)\}}\left(\frac{1}{I_b}\sum_{i_b=0}^{I_b-1} y^J(i_b;m)\right). \quad (18)$$

Therefore, different constellations can be loaded arbitrarily onto different subcarriers and even over different segments on each subcarrier. In other words, our blind channel estimation methods offer flexibility for variable frame designs. For instance, one could opt for BPSK signaling at the beginning of the frame to allow for quick channel acquisition with low complexity. Larger constellations can then be transmitted to increase bandwidth efficiency. As soon as sufficient blocks are collected, the channel can be refined through our low-cost PD algorithm operating on (18), or its online counterparts of (16).

D. Coded Transmissions

So far we have assumed that the transmitted symbols $s(i;m)$ are uncoded and equiprobable. In practice however, error control codes are usually employed to combat channel nulls (or deep fades). With coded symbols, assumption a1b) may not hold true. Because with PSK constellations we do not require a1b), our method applies if the coded symbols are transmitted with PSK modulation. For other constellations (e.g., QAM), our method is applicable only when a1b) holds true. This is

the case with practical OFDM systems utilizing BICM [2], where coded bits are interleaved before constellation mapping. Interestingly, interleaving has been adopted by many standards, see e.g., IEEE802.11a [6] and HIPERLAN/2 [4] (see also [23, Fig. 5] for a transceiver diagram). Hence, our channel estimators can be applied even for constellations other than PSK in practical OFDM systems.

E. Application to Semi-Blind Scenarios

In practical systems, training sequences are usually provided for synchronization or quick channel acquisition. To capitalize further on the training sequence, semi-blind implementation of the subspace-based method was proposed in [14] to facilitate convergence and also enable tracking of slow channel variations. Similarly, we can exploit the training sequence to initialize our PD algorithm either for improving channel estimation accuracy, or, for tracking channel variations. Specifically, to improve channel accuracy along the received blocks, we initialize PD with training-based channel estimates and update $\{\hat{H}(\rho_m)\}_{m=0}^{M-1}$ using:

$$\hat{H}_I^J(\rho_m) = \frac{I-1}{I}\hat{H}_{I-1}^J(\rho_m) + \frac{1}{I}\frac{y^J(I;m)}{\mathrm{E}\{s^J(I;m)\}} \tag{19}$$

where I denotes the number of blocks. On the other hand, we can initialize the tracking algorithm of Section III-B [cf. (16)] with training-based channel estimates. Such a semi-blind channel estimation method has only linear complexity and is thus very simple when compared to the cubic complexity of matrix SVD decompositions required by subspace-based methods.

Two important benefits become available with semi-blind scenarios. First, M does not need to be greater than $JL+1$ as required by Proposition 1. Instead, $M \geq L+1$ is enough to guarantee channel identifiability. Second, pilot tones are not needed to resolve the scalar ambiguity, which improves bandwidth efficiency usually by 5 – 15% [13]. In a nutshell, the PD algorithm is attractive in practical systems involving training sequences, thanks to its low complexity and high performance.

V. Comparisons and Simulated Performance

In this section, we illustrate the merits of our blind channel estimators through simulations. We also test semi-blind implementations of our method on a realistic HIPERLAN/2 system with training sequences. The figure of merit for channel estimation is the normalized least-squares channel error (NLSCE) defined in the frequency domain as

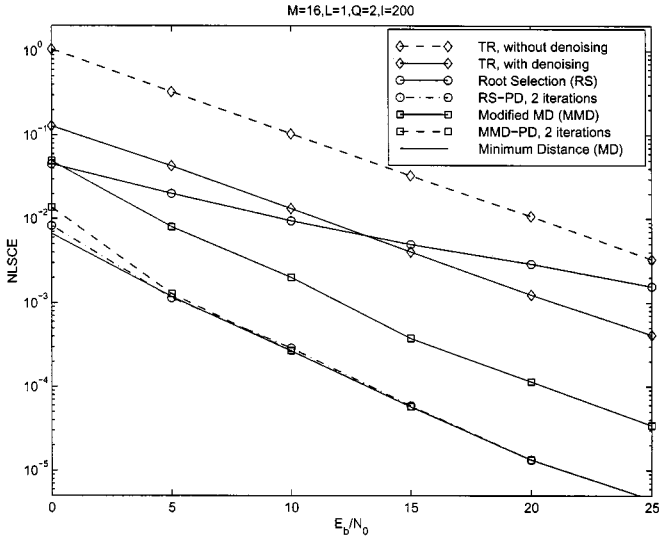
$$\frac{\sum_{m=0}^{M-1}[H(\rho_m)-\hat{H}(\rho_m)]^2}{\sum_{m=0}^{M-1}[H(\rho_m)]^2} = \frac{||\tilde{h}-\hat{\tilde{h}}||^2}{||\tilde{h}||^2} = \frac{||\Delta\tilde{h}||^2}{||\tilde{h}||^2} \tag{20}$$

where $\tilde{\mathbf{h}} := [H(\rho_0),\ldots,H(\rho_{M-1})]^{\mathrm{T}} = \mathbf{V}_1\mathbf{h}$ and $\Delta\tilde{\mathbf{h}} := \tilde{\mathbf{h}} - \hat{\tilde{\mathbf{h}}}$.

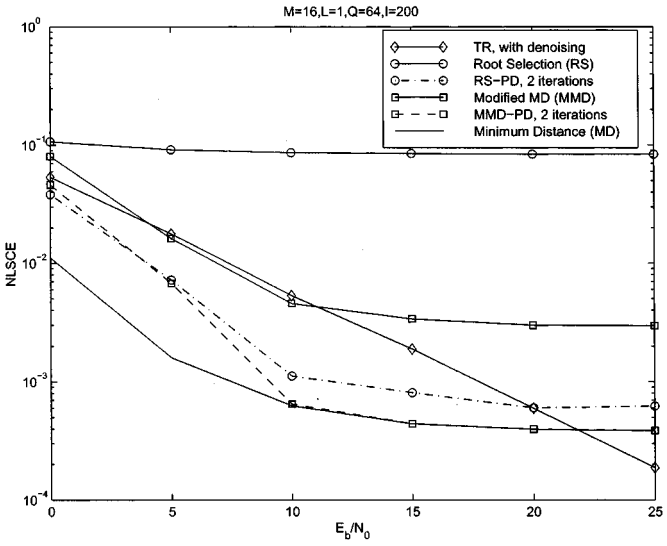
The reason behind adopting (20) rather than the time-domain NLSCE $||\mathbf{h}-\hat{\mathbf{h}}||^2/||\mathbf{h}||^2 := ||\Delta\mathbf{h}||^2/||\mathbf{h}||^2$ is that our final step is channel equalization in the frequency domain using $\hat{s}(i;m) = y(i;m)/\hat{H}(\rho_m)$. Therefore, (20) directly measures the channel estimation accuracy and its influence on the overall bit error rate (BER). To establish the relationship between frequency- and time-domain NLSCE, we start from: $||\Delta\tilde{\mathbf{h}}||^2/||\tilde{\mathbf{h}}||^2 = \left(\Delta\mathbf{h}^{\mathcal{H}}\mathbf{V}_1^{\mathcal{H}}\mathbf{V}_1\Delta\mathbf{h}\right)/\left(\mathbf{h}^{\mathcal{H}}\mathbf{V}_1^{\mathcal{H}}\mathbf{V}_1\mathbf{h}\right)$. Recalling that $\mathbf{V}_1$ is proportional to the first $L+1$ columns of the $M\times M$ FFT matrix when all subcarriers $\{\rho_m\}_{m=0}^{M-1}$ are used, we verify that $\mathbf{V}_1^{\mathcal{H}}\mathbf{V}_1 = M\mathbf{I}$ and $\mathbf{V}_1$ has unity condition number: $\mathrm{cond}(\mathbf{V}_1)=1$. Therefore, we obtain the same NLSCE in both frequency and time domains: $||\Delta\tilde{\mathbf{h}}||^2/||\tilde{\mathbf{h}}||^2 = ||\Delta\tilde{\mathbf{h}}||^2/||\mathbf{h}||^2$. However, due to the presence of frequency guard intervals or inactive subcarriers in real systems, we will only concentrate on those active subcarriers. As a result, $\tilde{\mathbf{h}}$ includes only active subcarriers in (20) and $\mathbf{V}_1$ reduces to $\mathbf{V}_a$ by keeping only those rows corresponding to the active subcarriers. If active subcarriers are not equispaced, the orthogonality among columns of $\mathbf{V}_a$ is destroyed and $\mathrm{cond}(\mathbf{V}_a) > 1$. For the subcarrier allocation in the HIPERLAN/2 system, it turns out that $\mathrm{cond}(\mathbf{V}_a) = 31.5$. It is thus worth clarifying that the estimation accuracy is different when translating channel estimates from one domain to the other, and that the frequency-domain NLSCE is more meaningful because our final goal is to equalize the channel in the frequency domain. On the other hand, if channel equalization were performed in the time domain (as is done in [20]), the time-domain NLSCE would be the natural choice. Next, we test our blind channel estimators in various scenarios.

Test Case 1: We set $M=16$, $L=1$ (two-ray channels), $I=200$, and average over 500 randomly generated channels each with independent Rayleigh distributed taps. We test four widely used signal constellations: BPSK, QPSK, 16QAM, 64QAM. However, due to space limitations and the similar behavior exhibited by BPSK and QPSK, 16QAM and 64QAM, we will only show representative results for BPSK and 64QAM. To allow for a fair comparison among different signal constellations, we normalize signal power to the same bit energy E_b [18]. Accounting for the relationship between symbol energy and bit energy for constellations of size Q: $E_s = E_b\log_2 Q$, the curves depicting channel NLSCE with respect to (w.r.t.) bit-energy-to-noise-ratio E_b/N_0 imply also those w.r.t. E_s/N_0.

To benchmark our blind channel estimation methods, we also examined a standard training (TR)-based method that relies on two ($N_t = 2$) known symbol blocks: $\hat{H}(\rho_m) = (1/2)[(y(0;m)/s(0;m)) + (y(l;m)/s(1;m))]$, $\forall_m \in [0, M-1]$. Because the channel is parameterized by $L+1$ coefficients, we can improve channel estimation accuracy through denoising [22]. Specifically, we obtain training-based time domain channel estimates through LS fitting $\hat{\mathbf{h}}_{\mathrm{tr}} = \mathbf{V}_1^{\dagger}\bar{\mathbf{h}}_{\mathrm{tr}} = 1/M\mathbf{V}_1^{\mathcal{H}}\bar{\mathbf{h}}_{\mathrm{tr}}$, and the denoised frequency domain channel estimates as: $\hat{\tilde{\mathbf{h}}}_{\mathrm{tr}} = \mathbf{V}_1\hat{\mathbf{h}}_{\mathrm{tr}}$. Denoising is well motivated since we have oversampled the channel frequency response (M samples are used instead of $L+1$); thus, we can reduce noise effects through lowpass filtering and obtain a significant gain of $10\log_{10}[M/(L+1)]$ dB [13]. This gain is 9 dB in our experimental setting as confirmed by observing the 9-dB gap between curves with and without denoising in Fig. 3(a). In the HIPERLAN context which will be discussed in test case 2, this gain is $10\log_{10}[M/(L+1)] = 10\log_{10}(64/16) = 6$ dB.

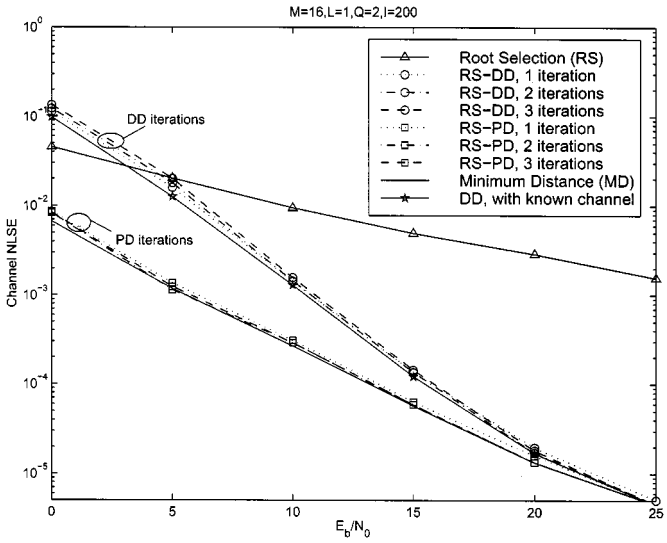


(a)

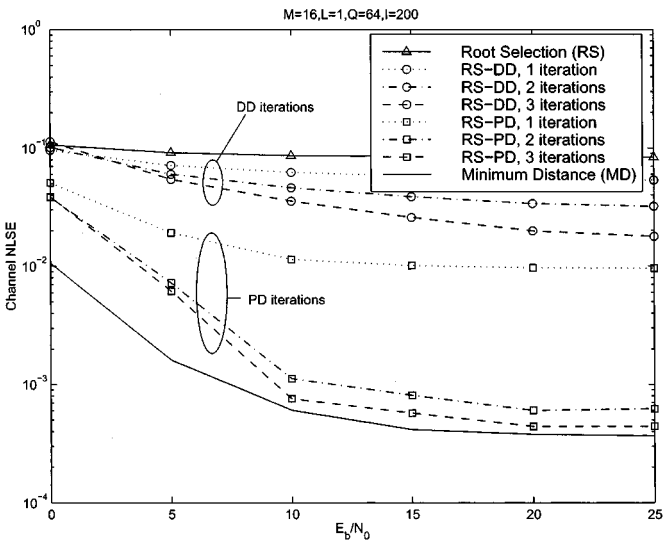


(b)

Fig. 3. Comparisons of different blind channel estimators:(a) BPSK and (b) 64QAM.



(a)



(b)

Fig. 4. Testing DD and PD initialized with the RS channel estimator: (a) BPSK and (b) 64QAM.

Note that going back to the impulse response (time-domain) vector with $L+1$ free parameters, we have already introduced this denoising step in all of our algorithms.

We demonstrate the nice features of our proposed method from the following four different perspectives.

1a) Channel estimation accuracy: We present in Fig. 3 the RS, MD, MMD (with $\bar{N} = L+1$ equispaced subcarriers) channel estimators as well as: i) the RS of [25] followed by 2 PD iterations; ii) the MMD followed by 2 PD iterations; and iii) the training-based (TR) method. First, we see that MD benchmarks all other blind channel estimators and outperforms the TR method. The RS method yields coarse channel estimates because as we mentioned in [25] root finding is a noise-sensitive operation especially with large constellations. The MMD offers rather good channel estimates, it outperforms the TR with BPSK and is comparable to TR with 64QAM. Only 2 PD $(I_1 = 2)$ refinements suffice to make the RS and the MMD algorithms approach the best achievable MD performance.

Error floor appears when our method is applied to large signal constellations at high SNR. This noise-induced floor comes from the mismatch between the sample average (12) and the theoretical ensemble average (10). As we mentioned, PSK signals are an exception to this rule. We will comment further on the floor effects after we compare PD with DD.

1b) PD versus DD comparisons: First, we feed both the DD and the PD with rather coarse RS estimates. We see in Fig. 4(a) that at low SNR, DD diverges due to a poor symbol estimation step. However, PD improves significantly by averaging across blocks. Fig. 4(a) also depicts the performance of DD when initialized by known channel values. Its proximity to RS-DD suggests that the BER in both cases is almost the same. Its gap from the RS-PD (or the MD) testifies to the decoupling property of our channel estimation from symbol recovery. At a certain SNR, the achievable BER remains constant, while the channel estimation accuracy can be improved consistently as we collect more and more blocks [cf. (12)]. In Fig. 4(b), we observe

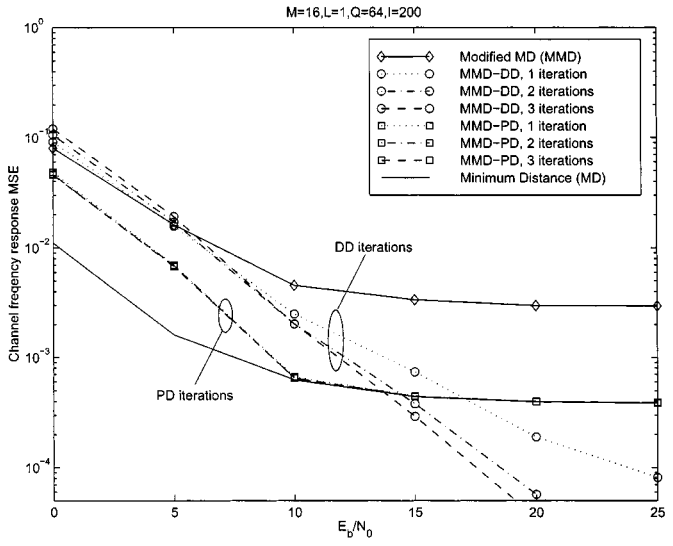


Fig. 5. MMD-DD versus MMD-PD (64QAM).

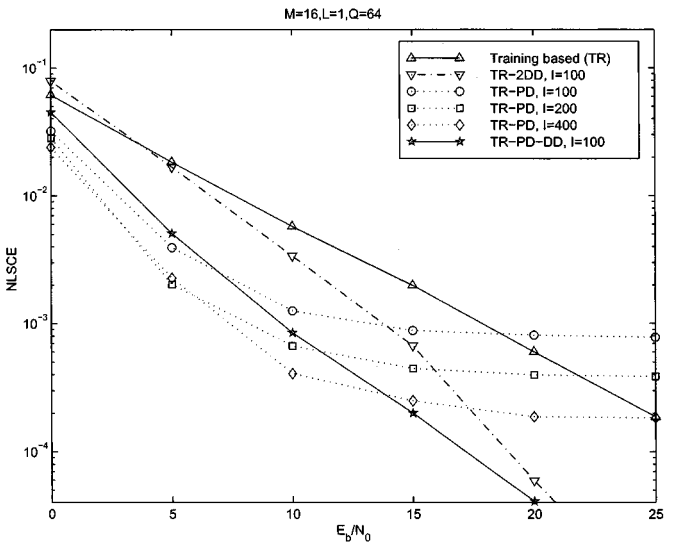


Fig. 6. Error floor effects (64QAM).

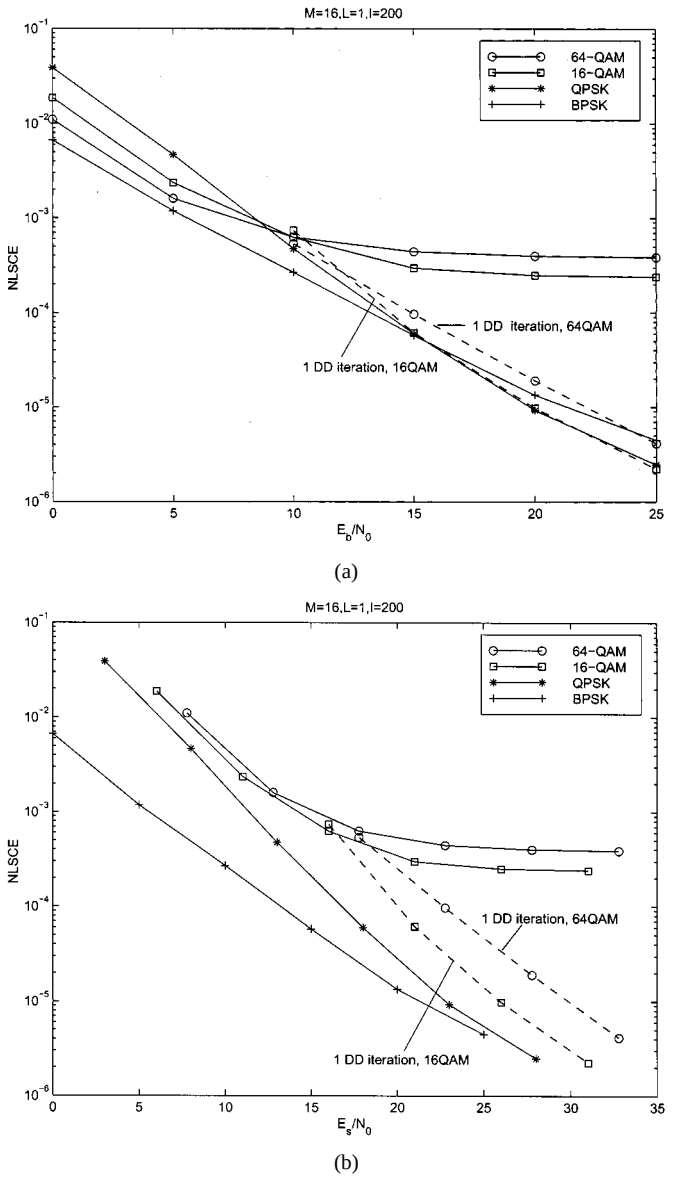


Fig. 7. Robustness of proposed channel estimators to variable constellation size. (a) NLSCE versus E_b/N_0. (b) NLSCE versus E_s/N_0.

that PD converges much faster than DD for 64QAM and yields much more reliable channel estimates. This happens because DD quantizes each symbol estimate to $Q = 64$ values, while PD needs only to pick one out of J ($J = 4 \ll Q$ in this case) phase values for each $H(\rho_m)$, which has been already denoised through the averaging of (12).

Second, we initialize DD and PD with better MMD estimates. For BPSK, we obtain corresponding curves similar to Fig. 4(a). However, Fig. 5 shows that PD performs worse than DD at high SNR for 64QAM and error floor appears.

1c) Error floor effects: Starting from (12), we can verify that the aforementioned noise-induced error floor appears in our blind channel estimation method for constellations other than PSK. However, the error floor can be lowered by collecting more blocks. In Fig. 6, we initialize PD and DD with channel estimates obtained from training. We see that the floor level decreases as the number of blocks increases. We can eliminate this error floor by applying one DD iteration to TR-PD (denoted by TR-PD-DD in Fig. 6). In Fig. 6, we see that TR-PD-DD outperforms TR-PD for $E_b/N_0 \geq 7$ dB. At low SNR, it diverges due to the limited symbol estimation accuracy. There is no error floor for PSK transmissions and TR-PD is always better than TR-PD-DD since it even outperforms the DD with known channel initialization, as evidenced by Fig. 4(a). Therefore, applying one DD iteration is only effective at high SNR and for large signal constellations other than PSK. However, in such cases we see from Fig. 6 that PD offers accurate enough channel estimates.

1d) Robustness to constellation size: Our algorithms rely on the finite alphabet property of source symbols and one would expect their performance to be dependent on the constellation size. However, our channel estimators separate channel estimation from symbol detection by averaging out symbols in (12); hence, they are robust to the constellation size. To illustrate this point, we compare the achievable NLSCE through the MD approach for BPSK, QPSK, 16QAM and 64QAM. Fig. 7(a) depicts the comparisons with the same

TABLE I
SUBCARRIER ALLOCATION PER OFDM SYMBOL IN HIPERLAN2

$1\cdots6$	$7\cdots11$	12	$13\cdots25$	26	$27\cdots32$	33	$34\cdots39$	40	$41\cdots53$	54	$55\cdots59$	$60\cdots64$
$0\cdots0$	$X\cdots X$	P_1	$X\cdots X$	P_2	$X\cdots X$	0	$X\cdots X$	P_3	$X\cdots X$	P_4	$X\cdots X$	$0\cdots0$

E_s/N_0, while Fig. 7(b) translates Fig. 7(a) to the same E_s/N_0. First, we see that BPSK signaling leads to the best achievable channel estimation accuracy because $J = 2$ implies less noise enhancement compared to $J = 4$ that is used for all other constellations. With the same E_s/N_0, Fig. 7(a) illustrates that the channel estimation accuracy for larger constellations outperform that for smaller constellations at low SNR, while we know that the corresponding BER's obey the reverse order [18]. At high SNR, 64QAM and 16QAM exhibit error floor and are thus worse than QPSK signals. We note that 16QAM has lower floor than 64QAM, because $s^4(n)$ takes 4 distinct values as compared to 16 values assumed by 64QAM. To test the performance between different constellations without error floor, we also plot the curves for 16 and 64QAM with one DD iteration. We see that our channel estimation methods are comparable for QPSK, 16QAM and 64QAM signals. The small gaps come from the finite sample effects. Hence, we have verified that our channel estimators offer channel estimates that are robust against different constellations.

Test Case 2: Here we test application of the PD algorithm to the HIPERLAN/2 system. First, we briefly introduce the HIPERLAN/2 standard [4], [23] which is a broadband wireless communication system operating at 5 GHz with bandwidth 20 MHz. Typical SNR values are 5–10 dB and terminal speeds are ≤ 3 m/s. OFDM transmission is utilized with IFFT of length $M = 64$ and CP of length $P - M = 16$ [cf. Fig. 1]. Among the $M = 64$ available subcarriers, only 52 are active out of which 48 bear information symbols while 4 of them are serving as pilots. The assignment of subcarriers in each OFDM symbol is summarized in Table I, where 0 stands for no transmission on the corresponding subcarrier, X stands for data, and $P_1 \ldots P_4$ denote the four pilot symbols. Each Medium Access Control (MAC) frame of 500 OFDM symbols consists of several transport channels (e.g., Broadcast channel (BCH), Access feedback channel (ACH)) in addition to downlink and/or uplink data transmissions [5]. The number of OFDM symbols used by downlink and/or uplink are dynamically assigned [5]. We focus on downlink transmissions here and assume that each data burst contains $N_{\rm sym}$ OFDM symbols, where $N_{\rm sym}$ varies in the range of $[0, 500)$. In front of each downlink data burst, two known OFDM symbols are inserted for channel acquisition. Based on these training-based initial channel estimates, we propose here to apply the PD algorithm to improve channel estimation accuracy and track channel variations.

In this HIPERLAN context, we first check the influence of inactive subcarriers. We assume that the channel is AWGN [18] $(h(0) = 1, h(l) = 0, \forall\, l \in [1, 16])$ and adopt QPSK signaling with $I = 200$ blocks. The subspace channel estimator is applied: $\hat{\mathbf{h}} = \arg\min_{\mathbf{h}}(\mathbf{h}^{\mathcal{H}}\mathbf{Q}\mathbf{h})$, $\mathbf{h} \neq 0$, where Q is obtained according to [15]. The scalar ambiguity is removed using 4 pilot tones. We also implement the joint subspace and pilot optimization algorithm of [14]: $\hat{\mathbf{h}} = \arg\min_{\mathbf{h}}(\mathbf{h}^{\mathcal{H}}\mathbf{Q}\mathbf{h} + \beta\sum_{n=1}^{4}|\sum_{l=0}^{L} h(l)p_n^{-l} - \hat{H}(p_n)|^2)$, where $\{p_n\}_{n=1}^4$ are four pilot carriers and β is a weighting factor that is set to be 1 here. In Fig. (8), we see that the subspace method suffers considerably when only 52 subcarriers are used instead of 64. When E_b/N_0 is in the practical range of 5 – 15 dB, subspace-based methods are outperformed by the training-based method with $N_t = 2$ known blocks. For this reason, we will not plot subspace-based methods in our following experiments. On the other hand, the training-based method and our TR-PD (1 iteration, i.e., $I_1 = 1$) are not affected by the null carriers. By averaging $I = 200$ blocks, TR-PD should perform approximately $10\log_{10}(I/(N_t J^2)) = 7.95$ dB better than the training-based method in this setting. In fact, we observe that the gap between these two methods is larger than 8 dB and increases as SNR increases.

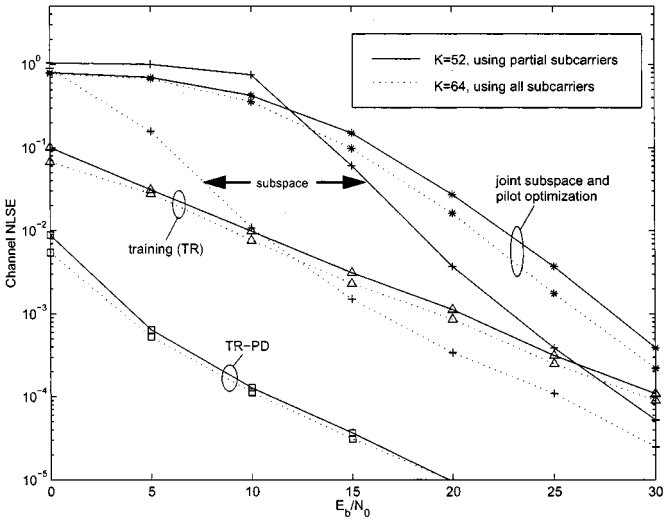
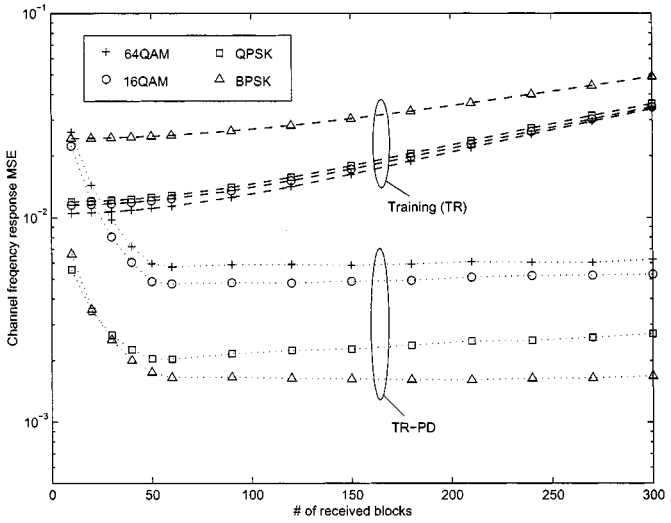
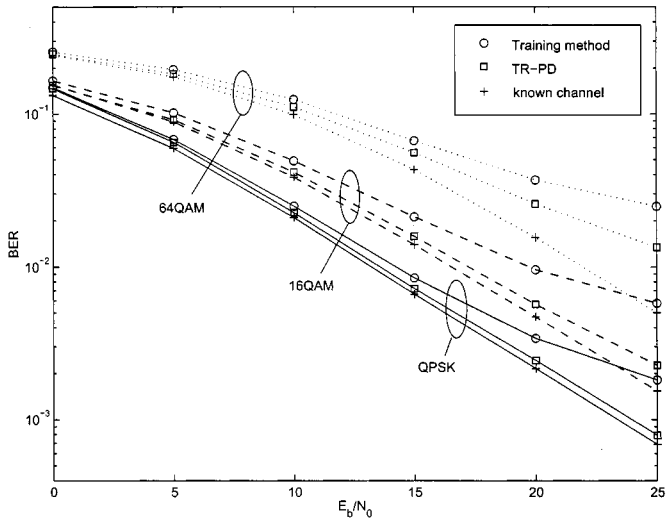


Fig. 8. Comparison with and without null carriers.

Next, we compare performance of the classical training-based method against our semi-blind TR-PD algorithm using the channel model B in [3] with $L + 1 = 16$ Rayleigh distributed taps. We test slowly time-varying channels with Jakes's Doppler spectrum at $\nu = 1.5$ m/s and a typical SNR of $E_b/N_0 = 10$ dB. We assume that each data burst has $N_{\rm sym} = 300$ symbols, out of which the first two symbols are known. We also assume that the known training symbols are drawn from the same constellation as the data symbols, and we here study four widely used constellations: BPSK, QPSK, 16QAM and 64QAM. For our TR-PD method, we use all the received data blocks when the received number of blocks satisfies: $I < 50$. Once $I > 50$, we only use the most recent 50 blocks to perform channel estimation which corresponds to set the sliding window length $W = 50$ in (16). Fig. 9(a) verifies that TR-PD tracks closely slow channel variations while TR-based channel estimates drift away. Fig. 9(a) also reveals that TR-PD yields better channel estimates for BPSK,



(a)



(a)

Fig. 9. Application to the HIPERLAN/2 system. (a) NLSCE versus # of blocks ($E_b/N_0 = 10$ db, $v = 1.5$ m/s). (b) BER versus E_b/N_0($v = 1.5$ m/s).

QPSK than for 16QAM and 64QAM because of finite-sample effects for QAM signals [cf. Section III-C-I]. Furthermore, we can infer from Fig. 9(a) that it is worthwhile to apply TR-PD to improve channel estimation accuracy when the data burst has length $N_{\rm sym} > 20$ for BPSK, QPSK and $N_{\rm sym} > 50$ for 16QAM, 64QAM.

To check the overall performance of channel estimation and equalization, we also tested the (uncoded) BER averaged over the entire data burst which is assumed to have length $N_{\rm sys} = 150$. The TR-PD channel estimates are updated every 25 OFDM symbols with sliding window length $W = 50$. From Fig. 9(b) we see that the training-based method incurs significant SNR loss relative to the benchmark BER obtained with perfect channel knowledge when the channel varies with time. To alleviate this SNR loss, frequent retraining is needed within long data bursts which of course decreases bandwidth efficiency. However, as demonstrated in Fig. 9(b), the TR-PD can track slow channel variations accurately which allows the system to approach the benchmark BER without sacrificing bandwidth efficiency.

VI. Conclusion

A finite-alphabet based channel estimation approach has been developed in this paper and shown to possess attractive features not only for the classical CP-OFDM but also for the modified ZP-OFDM. Being simple and flexible it holds high application potential for existing OFDM standards and future generation multicarrier systems. Those include also wideband CDMA hybrids that rely on multicarrier transmissions to suppress multiuser interference and reduce channel estimation to the OFDM-like single-user problem treated in this paper, see, e.g., [24], [26]. Incorporation of the novel channel estimators to multiuser generalized OFDM systems with frequency-hopping is dealt with in [25]. The channel estimation algorithms derived herein offer choices to trade off performance with complexity, while guaranteeing channel identifiability regardless of channel zero locations. Unlike decision-directed alternatives, they decouple channel from symbol estimation. This decoupling enhances their robustness to constellation sizes and reduces estimation of the scalar ambiguity inherent to all blind channel estimators to a search over finite values. Especially with PSK modulated signals, high estimation accuracy is achieved with minimal received data. By exploiting training data usually specified in OFDM standards, semi-blind and adaptive implementations were shown to improve estimation accuracy and be capable of tracking channel variations with surprisingly low complexity.

References

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, pp. 5–14, May 1990.

[2] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.

[3] "Channel models for HIPERLAN/2 in different indoor scenarios," European Telecommunications Standards Institute, Sophia-Antipolis, Valbonne, France, Norme ETSI, document 3ERI085B, 1998.

[4] *Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) layer*, ETSI TS 101 475.

[5] *Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Data Link Control (DLC) Layer; Part 1: Basic Data Transport Functions*, ETSI TS 101 761-1.

[6] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, IEEE 802.11a, 1999.

[7] "The DWMT: A multicarrier transceiver for ADSL using M-band wavelets," ANSI, 1993.

[8] M. de Courville, P. Duhamel, P. Madec, and J. Palicot, "Blind equalization of OFDM systems based on the minimization of a quadratic criterion," in *Proc. ICC*, vol. 3, Dallas, TX, June 1996, pp. 1318–1321.

[9] P. Frenger and A. Svensson, "Decision-Directed coherent detection in multicarrier systems on Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 490–498, Mar. 1999.

[10] G. B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Lett.*, vol. 4, pp. 184–187, June 1997.

[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Laurel Park, MD: Johns Hopkins Univ. Press, 1996.

[12] R. W. Heath and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Processing*, vol. 47, pp. 848–856, Mar. 1999.

[13] V. Mignone and A. Morello, "CD3-OFDM: A Novel Demodulation Scheme for Fixed and Mobile Receivers," *IEEE Trans. Commun.*, vol. 44, pp. 1144–1151, Sept. 1996.

[14] B. Muquet, M. de Courville, and P. Duhamel, "A subspace based blind and semi-blind channel identification method for OFDM systems," submitted for publication.

[15] B. Muquet, M. de Courville, P. Duhamel, and V. Buenac, "A subspace based blind and semi-blind channel identification method for OFDM systems," in *Proc. IEEE-SP Workshop on Signal Proc. Advances in Wireless Comm.*, Annapolis, MD, May 9–12, 1999, pp. 170–173.

[16] B. Muquet, M. de Courville, P. Duhamel, and G. B. Giannakis, "OFDM with Trailing Zeros Versus OFDM with Cyclic Prefix: Links, Comparisons and Application to the HIPERLAN/2 System," in *Proc. Int. Conf. on Communications*, New Orleans, LA, June 2000, pp. 1049–1053.

[17] B. Muquet, M. de Courville, G. B. Giannakis, Z. Wang, and P. Duhamel, "Reduced Complexity Equalizers for Zero-Padded OFDM Transmissions," in *Proc. Int. Conf. on Acoust. Speech and Signal Processing*, vol. 5, Istanbul, Turkey, June 5–9, 2000, pp. 2973–2976.

[18] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

[19] A. Ruiz, J. M. Cioffi, and S. Kasturia, "Discrete multiple tone modulation with coset coding for the spectrally shaped channel," *IEEE Trans. Commun.*, vol. 40, pp. 1012–1029, June 1992.

[20] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part II: Blind channel estimation, synchronization, and direct estimation," *IEEE Trans. Signal Processing*, vol. 47, pp. 2007–2022, July 1999.

[21] F. Tufvesson and T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems," in *Proc. Vehicular Technology Conf.*, vol. 3, Phoenix, AZ, May 1997, pp. 1639–1643.

[22] J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in OFDM systems," in *Proc. Vehicular Technology Conf.*, vol. 2, Chicago, IL, July 1995, pp. 815–819.

[23] R. D. J. van Nee, G. A. Awater, M. Morikura, H. Takanashi, M. A. Webster, and K. W. Halford, "New high-rate wireless LAN standards," *IEEE Commun. Mag.*, vol. 37, Dec. 1999.

[24] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Processing Mag.*, pp. 29–48, May 2000.

[25] S. Zhou, G. B. Giannakis, and A. Scaglione, "Long codes for generalized FH-OFDMA through unknown multipath channels," *IEEE Trans. Commun.*, vol. 49, pp. 721–733, Apr. 2001.

[26] S. Zhou, G. B. Giannakis, and A. Swami, "Frequency-hopped generalized MC-CDMA for multipath and interference suppression," in *Proc. MILCOM Conf.*, Los Angeles, CA, Oct. 22–25, 2000, pp. 937–941.

Shengli Zhou (S'99) was born in Anhui, China, in 1974. He received the B.S. and M.Sc. degrees in electrical engineering and information science from the University of Science and Technology of China (USTC), in 1995 and 1998, respectively. He is now working toward the Ph.D. degree in the Department of Electrical and Computer Engineering at the University of Minnesota, Minneapolis.

His broad interests lie in the areas of communications and signal processing, including transceiver optimization, blind channel estimation and equalization algorithms, wireless, multicarrier, space–time coded and spread-spectrum communication systems.

Georgios B. Giannakis (F'96) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, in 1981 and the M.Sc. degrees in electrical engineering and mathematics and the Ph.D. degree in electrical engineering from the University of Southern California (USC), in 1983, 1986, and 1986.

After lecturing for one year at USC, he joined the University of Virginia in 1987, where he became a Professor of Electrical Engineering in 1997. Since 1999, he has been a Professor with the Department of Electrical and Computer Engineering at the University of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification—subjects on which he has published more than 125 journal papers, 250 conference papers and two edited books. Current research topics focus on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, redundant precoding and space-time coding for block transmissions, multicarrier, and wide-band wireless communication systems.

Dr. Giannakis has received three best paper awards from the IEEE Signal Processing (SP) Society (1992, 1998, 2000) and the Society's Technical Achievement Award in 2000. He co-organized three IEEE-SP Workshops (HOS in 1993, SSAP in 1996 and SPAWC in 1997) and guest (co-) edited four special issues. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, a secretary of the SP Conference Board, a member of the SP Publications Board and a member and vice-chair of the Statistical Signal and Array Processing Committee. He is a member of the Editorial Board for the PROCEEDINGS OF THE IEEE, he chairs the SP for Communications Technical Committee, and he serves as the Editor-in-Chief for the IEEE SIGNAL PROCESSING LETTERS. He is a member of the IEEE Fellows Election Committee, the IEEE-SP Society's Board of Governors, and a frequent consultant for the telecommunications industry.