A Comparison of Pilot-Aided Channel Estimation Methods for OFDM Systems

Michele Morelli and Umberto Mengali, *Fellow, IEEE*

***Abstract*—This paper deals with the estimation of the channel impulse response (CIR) in orthogonal frequency division multiplexed (OFDM) systems. In particular, we focus on two pilot-aided schemes: the maximum likelihood estimator (MLE) and the Bayesian minimum mean square error estimator (MMSEE). The advantage of the former is that it is simpler to implement as it needs no information on the channel statistics. On the other hand, the MMSEE is expected to have better performance as it exploits prior information about the channel.**

Theoretical analysis and computer simulations are used in the comparisons. At SNR values of practical interest, the two schemes are found to exhibit nearly equal performance, provided that the number of pilot tones is sufficiently greater than the CIRs length. Otherwise, the MMSEE is superior. In any case, the MMSEE is more complex to implement.

***Index Terms*—Channel estimation, maximum likelihood, minimum mean square error, OFDM.**

I. Introduction

ORTHOGONAL frequency division multiplexing (OFDM) has received considerable interest in the last few years for its advantages in high-bit-rate transmissions over frequency-selective fading channels. In OFDM systems, the input high-rate data stream is divided into many low-rate streams [1], [2] that are transmitted in parallel, thereby increasing the symbol duration and reducing the intersymbol interference. These features have motivated the adoption of OFDM as a standard for digital audio broadcasting (DAB) [3], digital video broadcasting (DVB) [4], and broadband indoor wireless systems [5].

Coherent OFDM detection requires channel estimation and tracking. To this purpose, known symbols (pilots) are often multiplexed into the data and channel estimation is performed by interpolation. Channel estimation can be avoided by using differential detection, at the cost, however, of a 3-dB loss in signal-to-noise ratio (SNR).

Several pilot-aided channel-estimation schemes for OFDM applications have been investigated, and [6]–[10] provide a good sample of the results obtained in this area. In particular, the method proposed in [6] provides channel estimates based on piecewise-constant and piecewise-linear interpolations between pilots. It is simple to implement, but it needs a large number of pilots to get satisfactory performance. In [7], a low-rank approximation to the frequency-domain linear minimum mean squared error estimator (MMSEE) is proposed, making use of singular value decomposition techniques. The drawback of this approximation is that it requires knowledge of the channel frequency correlation and the operating SNR. In practice, the system can be designed for fixed values of SNR and channel correlation at the expense of performance losses. The MMSEE studied in [8] exploits channel correlations in time and frequency domains. Like the scheme in [7], it needs knowledge of the channel statistics and the operating SNR. Although it can work in a mismatched mode, its performance degrades if the assumed Doppler frequencies and delay spreads are smaller than the true ones. In [9], channel estimation is performed by two-dimensional interpolation between pilots. Like the method in [8], it is rather robust to Doppler, even though it exhibits performance degradations with lower Doppler frequencies. Finally, [10] investigates the maximum likelihood estimator (MLE). No information on the channel statistics or the operating SNR is required in this scheme.

Some remarks about MLE and MMSEE are of interest. These estimators are based on different assumptions about the channel impulse response (CIR). In the former, the CIR is viewed as a *deterministic but unknown* vector, whereas in the latter, it is regarded as a *random* vector whose particular realization we want to estimate. Correspondingly, the mean squared error (MSE) in the MLE is understood as an average over the observed data, whereas in the MMSEE, the average is taken not only over the data but over the CIR probability density function as well. It follows that the MMSEE has the minimum MSE "on the average," i.e., with respect to all the CIR realizations.

As we will see, the MLE achieves the Cramér-Rao lower bound (CRLB), and therefore, it is the minimum-variance unbiased estimator. No further improvement in MSE is possible as long as the CIR is viewed as a deterministic quantity and the estimator is unbiased. On the other hand, the MMSEE has prior information on the CIR and can exploit this information to do better than the MLE.

The above considerations prompt some important questions. The first is whether it is conceptually possible for the MMSEE to perform below the CRLB. The answer is affirmative since the CRLB is a bound only in the framework of the *classical approach* to estimation (where CIR is a deterministic quantity). When dealing with MMSEE, on the other hand, we adopt a *Bayesian approach*, and the corresponding estimation accuracy depends on prior information. In principle, performance can be as good as desired, provided that sufficient prior information is available. To emphasize this fact, in the following, we denote as MSE and *Bayesian* MSE (BMSE) the mean square errors in the MLE and MMSEE, respectively.

Manuscript received March 8, 2000; revised August 30, 2001. The associate editor coordinating the review of this paper and approving it for publication was Dr. Athina Petropulu.

The authors are with the Department of Information Engineering, University of Pisa, Pisa, Italy.

Publisher Item Identifier S 1053-587X(01)10498-8.

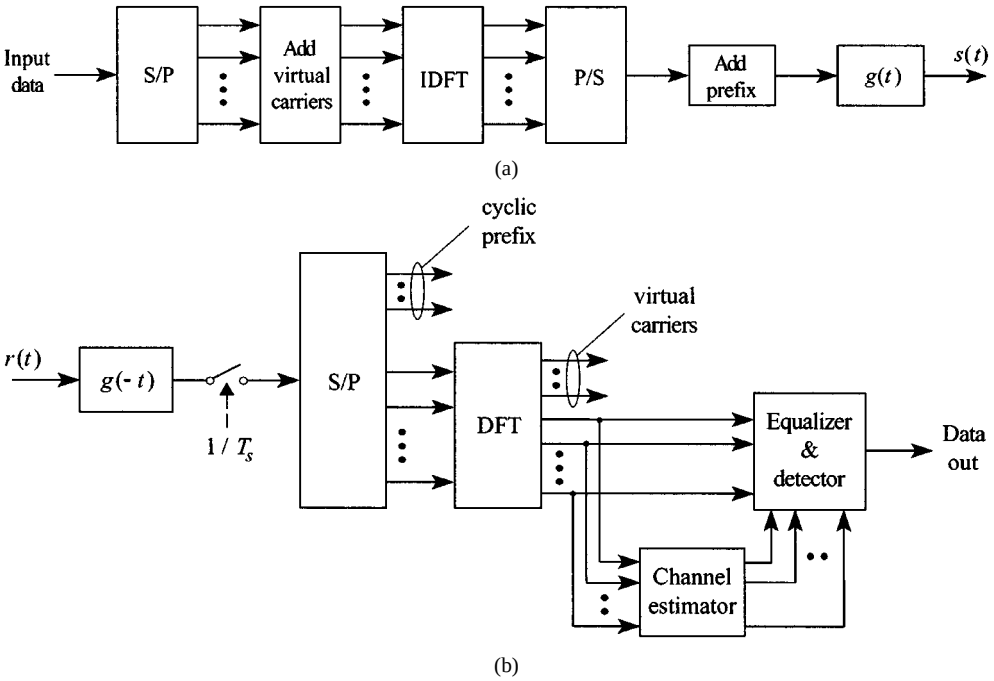

Fig. 1. (a) Block diagram of the OFDM transmitter. (b) Block diagram of the OFDM receiver.

Having established that MMSEE can do better than the CRLB, we wonder how much better it can do in practice and under which operating conditions. We also wonder about the price to pay in terms of computational complexity.

The purpose of this paper is to answer these questions. The discussion is organized as follows. The next section describes the signal model and introduces some basic notations. Section III revisits MLE and MMSEE and compares their complexity. Performance is assessed and is compared in Section IV. Section V discusses analytical and simulation results, and Section VI offers some conclusions.

II. Signal Model

We consider an OFDM system employing N subcarriers for the transmission of P parallel data symbols. Notice that $N-P$ subcarriers (virtual carriers) at the edges of the spectrum are not used to avoid aliasing problems at the receiver [11]. The block diagram of the OFDM transmitter is shown in Fig. 1(a). The stream of data $\{c_i\}$ (belonging to a PSK or QAM constellation) is serial to parallel (S/P) converted and partitioned into adjacent blocks of length $P=2N_\alpha+1$. After insertion of $N-P$ zeros, the lth OFDM block

$$\boldsymbol{c}^{(l)}=\left[c_0^{(l)},c_1^{(l)},\ldots,c_{N_\alpha}^{(l)},0,0,\ldots,0,c_{-N_\alpha}^{(l)},\ldots,c_{-2}^{(l)},c_{-1}^{(l)}\right] \tag{1}$$

is fed to an N-point inverse discrete Fourier transform (IDFT) unit that produces the N-dimensional vector $\boldsymbol{b}^{(l)}$ of time-domain samples. In order to eliminate any interference between adjacent OFDM symbols, an N_G-point cyclic prefix (longer than the overall channel impulse response) is appended to $\boldsymbol{b}^{(l)}$. The resulting extended vector drives a linear modulator with impulse response $g(t)$ and signaling interval $T_s=T/(N+N_G)$, where T is the OFDM symbol duration. In the following, we assume that $g(t)$ is a root-raised-cosine function with rolloff α.

The OFDM receiver is sketched in Fig. 1(b). After matched filtering, the signal is sampled at rate $1/T_s$ and serial to parallel converted. Next, the cyclic prefix is removed, and the received samples are passed to an N-point discrete Fourier transform (DFT) unit. As in [7], we assume that the channel variations are negligible over one block of data, and we indicate with $\boldsymbol{h}^{(l)}=[h^{(l)}(0),h^{(l)}(1),\ldots,h^{(l)}(L-1)]^T$ (the superscript $(\cdot)^T$ indicates vector transpose) the T_s-spaced samples of the overall CIR. Denoting

$$H^{(l)}(n)=\sum_{k=0}^{L-1}h^{(l)}(k)e^{-j2\pi nk/N} \tag{2}$$

the DFT of $\boldsymbol{h}^{(l)}$ and dropping the block identifier l for simplicity, the output of the DFT unit is found to be [12]

$$X(n)=c_nH(n)+w(n)\quad |n|\le N_\alpha \tag{3}$$

where $N_\alpha=\text{int}[N(1-\alpha)/2]$, and int$(\cdot)$ means *integer part* of the enclosed quantity. In (3), the c_n are the useful data symbols, and $w(n)$ is the channel noise, which is modeled as a white Gaussian process with zero mean and variance $\sigma^2=E\{|w(n)|^2\}$.

Coherent detection requires knowledge of the sampled channel frequency response $H(n)$. In this study, we assume that some known symbols (pilots) are multiplexed into the data stream, and channel estimation is performed by interpolation between pilots. A total of N_p pilots $\{a_n;0\le n\le N_p-1\}$ are inserted in the OFDM block at known locations $\{i_n;0\le n\le N_p-1\}$. Denoting by $\boldsymbol{X}=[X(i_0),X(i_1),\ldots,X(i_{N_p-1})]^T$ the N_p-dimensional vector containing the DFT output at the pilot locations, from (2) and (3), we have

$$\boldsymbol{X}=\boldsymbol{ABh}+\boldsymbol{w} \tag{4}$$

where $\boldsymbol{A}$ is a diagonal matrix

$$\boldsymbol{A} = \text{diag}\{a_0, a_1, \ldots, a_{N_p-1}\} \tag{5}$$

and $\boldsymbol{B}$ is an $N_p \times L$ matrix with entries

$$[\boldsymbol{B}]_{n,k} = e^{-j2\pi k i_n/N} \quad 0 \le n \le N-1, \quad 0 \le k \le L-1. \tag{6}$$

Vector $\boldsymbol{w}$ has a Gaussian distribution with zero mean and covariance matrix

$$\boldsymbol{C}_w = \sigma^2 \boldsymbol{I}_{N_p} \tag{7}$$

where $\boldsymbol{I}_{N_p}$ is the identity matrix of order N_p. Pilot symbols are taken from a PSK constellation, i.e., $|a_n| = 1$. Then, premultiplying both sides of (4) by $\boldsymbol{A}^H$ (the superscript $(\cdot)^H$ indicates Hermitian transpose) produces

$$\boldsymbol{Z} = \boldsymbol{B}\boldsymbol{h} + \tilde{\boldsymbol{w}} \tag{8}$$

where $\boldsymbol{Z} = \boldsymbol{A}^H \boldsymbol{X}$ has entries

$$Z(n) = a_n^* X(i_n) \quad 0 \le n \le N_p - 1 \tag{9}$$

and $\tilde{\boldsymbol{w}} = \boldsymbol{A}^H \boldsymbol{w}$ is statistically equivalent to $\boldsymbol{w}$. The goal is to derive estimates of the channel frequency response $H(n)$ from the observation of $\boldsymbol{Z}$.

III. Pilot-Aided Channel Estimation

Let $\boldsymbol{H} = \{H(n); |n| \le N_\alpha\}$ be the vector containing the channel frequency response. From (2), it is seen that

$$\boldsymbol{H} = \boldsymbol{G}\boldsymbol{h} \tag{10}$$

where $\boldsymbol{G}$ is a matrix with entries

$$[\boldsymbol{G}]_{n,k} = e^{-j2\pi nk/N} \quad |n| \le N_\alpha, \quad 0 \le k \le L-1. \tag{11}$$

From the invariance property of MLE [13, p. 185] and MMSEE [13, p. 349], it follows that if $\hat{\boldsymbol{h}}$ is the estimate of $\boldsymbol{h}$ (either ML or MMSE), then the corresponding estimate of $\boldsymbol{H}$ can be computed as

$$\hat{\boldsymbol{H}} = \boldsymbol{G}\hat{\boldsymbol{h}}. \tag{12}$$

A. Maximum Likelihood Estimator

The MLE is based on the assumption that $\boldsymbol{h}$ is a deterministic but unknown vector. The estimate of $\boldsymbol{h}$ is derived from the linear model (8) and is given by [13, p. 186]

$$\hat{\boldsymbol{h}}_{\text{MLE}} = \boldsymbol{D}^{-1}\boldsymbol{B}^H \boldsymbol{Z} \tag{13}$$

where $\boldsymbol{D}$ is a square matrix

$$\boldsymbol{D} = \boldsymbol{B}^H \boldsymbol{B} \tag{14}$$

whose entries are computed from (6)

$$[\boldsymbol{D}]_{n,k} = \sum_{m=0}^{N_p-1} e^{j2\pi(n-k)i_m/N} \quad 0 \le n,k \le L-1. \tag{15}$$

Note that (13) can be written in scalar form as

$$\hat{h}_{\text{MLE}}(k) = \sum_{m=0}^{N_p-1} Z(m) \sum_{n=0}^{L-1} [\boldsymbol{D}^{-1}]_{k,n} e^{j2\pi n i_m/N}. \tag{16}$$

Hence, substituting into (12) yields

$$\hat{H}_{\text{MLE}}(n) = \sum_{m=0}^{N_p-1} Z(m) p_{\text{MLE}}(n,m). \tag{17}$$

with

$$p_{\text{MLE}}(n,m) = \sum_{k=0}^{L-1} [\boldsymbol{D}^{-1}\boldsymbol{B}^H]_{k,m} e^{-j2\pi nk/N}. \tag{18}$$

B. Minimum Mean Square Error Estimator

The MMSEE is based on the *Bayesan approach* to statistical estimation, in which $\boldsymbol{h}$ is modeled as a random vector. In deriving the MMSEE, we assume that $\boldsymbol{h}$ is zero-mean and Gaussian and is uncorrelated with $\boldsymbol{w}$. Under these assumptions, it is found that[13, p. 364]

$$\hat{\boldsymbol{h}}_{\text{MMSEE}} = \boldsymbol{V}^{-1}\boldsymbol{B}^H \boldsymbol{Z} \tag{19}$$

where

$$\boldsymbol{V} = \sigma^2 \boldsymbol{C}_{\boldsymbol{h}}^{-1} + \boldsymbol{D} \tag{20}$$

and

$$\boldsymbol{C}_{\boldsymbol{h}} = E\{\boldsymbol{h}\boldsymbol{h}^H\} \tag{21}$$

is the covariance matrix of $\boldsymbol{h}$. Substituting (19) into (12) yields the MMSEE of the channel frequency response

$$\hat{H}_{\text{MMSEE}}(n) = \sum_{m=0}^{N_p-1} Z(m) p_{\text{MMSEE}}(n,m) \tag{22}$$

where

$$p_{\text{MMSEE}}(n,m) = \sum_{k=0}^{L-1} [\boldsymbol{V}^{-1}\boldsymbol{B}^H]_{k,m} e^{-j2\pi nk/N}. \tag{23}$$

C. Remarks

i) Equation (13) indicates that MLE requires the invertibility of $\boldsymbol{D}$. Such a condition is met if and only if $\boldsymbol{B}$ is full rank and $N_p \ge L$. This means that the number of pilots must be not smaller than the number of channel taps. On the other hand, from (19), we see that MMSEE requires the invertibility of $\boldsymbol{V}$. For this to hold, however, $\boldsymbol{B}$ need not be full rank. Thus, the MMSEE can exist even if $N_p < L$.

ii) If $\boldsymbol{B}$ is a square matrix $(N_p = L)$, from (13) and (14), it is seen that $\hat{\boldsymbol{h}}_{\text{MLE}}$ reduces to the channel estimator proposed by Negi and Cioffi [10]

$$\hat{\boldsymbol{h}}_{\text{MLE}} = \boldsymbol{B}^{-1}\boldsymbol{Z}. \tag{24}$$

iii) From (19), it is clear that MMSEE is linear in the observed vector $\boldsymbol{Z}$. This is a consequence of the Gaussian

assumption on $\boldsymbol{h}$. It follows that MMSEE coincides with the maximum *a posteriori* (MAP) estimator [13, p. 485].

iv) Comparing (19) and (20) with (13) and (14), it is seen that MMSEE reduces to MLE at high SNRs (i.e., when $\sigma^2 \to 0$) or if no prior information on $\boldsymbol{h}$ is given (i.e., when $\boldsymbol{C}_h^{-1} \to \boldsymbol{0}$, $\boldsymbol{0}$ being the null matrix).

v) Comparing (17) and (22), it appears that MLE and MMSEE have the same form

$$\hat{H}(n) = \sum_{m=0}^{N_p-1} Z(m)p(n,m) \tag{25}$$

which amounts to an interpolation between the samples $Z(m)$. Equations (14) and (18) indicate that the coefficients $p_{\mathrm{MLE}}(n,m)$ in the MLE depend only on the pilot symbols' locations, which are known. Hence, the $p_{\mathrm{MLE}}(n,m)$ can be precomputed and stored. On the contrary, from (20) and (23), it is seen that the coefficients $p_{\mathrm{MMSEE}}(n,m)$ in the MMSEE also depend on the channel covariance matrix $\boldsymbol{C}_h$ and the noise variance σ^2. Thus, suitable schemes must be found to estimate these parameters. Furthermore, even assuming that estimates of $\boldsymbol{C}_h$ and σ^2 are available, the computation of $P_{\mathrm{MMSEE}}(n,m)$ still requires inverting $\boldsymbol{C_h}$ and $\boldsymbol{V}$. In conclusion, the MMSEE is much more complex than the MLE.

vi) MLE and MMSEE can be implemented in two different ways.

a) Compute $\hat{H}(n)$ from (25).

b) Compute first $\hat{\boldsymbol{h}}$ from (13) or (19) and then derive $\hat{H}(n)$ as the DFT of $\hat{\boldsymbol{h}}$.

Procedure a) requires storing the $(2N_\alpha+1)N_p$ coefficients $p(n,m)$ and performing $4(2N_\alpha+1)N_p$ real multiplications and $(4N_p-1)(2N_\alpha+1)$ real additions. In writing these figures, we have borne in mind that a complex multiplication corresponds to four real multiplications and two real additions, whereas a complex addition is equivalent to two real additions. Procedure b) requires storing $\boldsymbol{D^{-1}B}^H$ for the MLE and $\boldsymbol{V^{-1}B}^H$ for the MMSEE. Furthermore, use of the fast Fourier transform (FFT) involves $4LN_p + 2\rho(2N_\alpha+1)\log_2 N$ real multiplications and $4(N_p-2)L + 3\rho(2N_\alpha+1)\log_2 N$ real additions. In these figures, ρ is the pruning factor [14] and represents the computational saving obtained in the FFT calculation by eliminating the operations on zeros. It turns out that ρ is given by

$$\rho = 1 - \frac{\log_2(N/L) - 2(1-L/N)}{\log_2 N}. \tag{26}$$

A comparison between the computational load in the two procedures is tricky because many parameters are involved, and different answers may be given, depending on the operating conditions. In any case, procedure a) requires more memory storage than b) since $L \ll 2N_\alpha + 1$ in many practical situations.

IV. Performance Analysis

The performance of MLE and MMSEE can be expressed either in terms of *individual* mean square error

$$\gamma(n) = E\{|\hat{H}(n) - H(n)|^2\} \tag{27}$$

or *total* mean square error

$$\Gamma = \frac{1}{2N_\alpha+1} \sum_{n=-N_\alpha}^{N_\alpha} \gamma(n). \tag{28}$$

A. Performance of MLE

With the MLE, the vector $\boldsymbol{h}$ is a constant, and the expectation in (27) is taken over the thermal noise. In Appendix A, it is shown that MLE is unbiased, i.e.,

$$E\{\hat{H}_{\mathrm{MLE}}(n)\} = H(n) \tag{29}$$

and the individual MSE is given by

$$\gamma_{\mathrm{MLE}}(n) = \sigma^2 \sum_{k=0}^{L-1}\sum_{m=0}^{L-1} [\boldsymbol{D}^{-1}]_{k,m} e^{j2\pi n(m-k)/N}. \tag{30}$$

Substituting (30) into (28) yields

$$\Gamma_{\mathrm{MLE}} = \sigma^2 \sum_{k=0}^{L-1}\sum_{m=0}^{L-1} [\boldsymbol{D}^{-1}]_{k,m} q(m-k) \tag{31}$$

where

$$q(m) = \frac{\sin[\pi(2N_\alpha+1)m/N]}{(2N_\alpha+1)\sin[\pi m/N]}. \tag{32}$$

In Appendix A, it is also shown that (31) coincides with the CRLB.

Although the MLE is derived for a deterministic $\boldsymbol{h}$, it can also be applied with a random CIR. In other words, the deterministic assumption is made to derive the MLE, which is then used with a random channel. In this context, it makes sense to consider the average of MSE taken over the CIR realizations. However, since $\gamma_{\mathrm{MLE}}(n)$ does not depend on $\boldsymbol{h}$, it is clear that individual and total errors are not affected by the averaging operation. Thus, they coincide with the Bayesian mean square errors.

B. Performance of the MMSEE

With the MMSEE, the expectation in (27) is taken over the noise and the CIR probability density function. The covariance matrix

$$\boldsymbol{C}_{\hat{h}} = E\left\{\hat{\boldsymbol{h}}_{\mathrm{MMSEE}}\hat{\boldsymbol{h}}^H_{\mathrm{MMSEE}}\right\} \tag{33}$$

can be written as [13, p. 391]

$$\boldsymbol{C}_{\hat{h}} = \sigma^2\boldsymbol{V}^{-1} \tag{34}$$

where $\boldsymbol{V}$ is defined in (20). Then, substituting (12) into (27) and (28) and bearing in mind (34) yields the individual and total BMSE

$$\gamma_{\mathrm{MMSEE}}(n) = \sigma^2 \sum_{k=0}^{L-1}\sum_{m=0}^{L-1} [\boldsymbol{V}^{-1}]_{k,m} e^{j2\pi n(m-k)/N} \tag{35}$$

$$\Gamma_{\mathrm{MMSEE}} = \sigma^2 \sum_{k=0}^{L-1}\sum_{m=0}^{L-1} [\boldsymbol{V}^{-1}]_{k,m} q(m-k) \tag{36}$$

where $q(m)$ is defined in (32).

C. Performance Comparison

From (31) and (32) and (35) and (36), it is seen that the estimation accuracy of MLE and MMSEE depends on

i) the noise variance σ^2;
ii) the number of channel taps L;
iii) the DFT size N;
iv) the number of modulated subcarriers $2N_\alpha + 1$;
v) the number N_p and locations of the pilots (through $\boldsymbol{D}$);
vi) the prior channel covariance matrix $\boldsymbol{C}_h$ (only for MMSEE).

Given L, N, N_α, and N_p, it would be interesting to determine the optimal pilots' locations that minimize the total BMSE. The solution to this problem is given in [10] for MLE when there are no suppressed carriers. It is found that the best performance is achieved by resorting to uniformly spaced pilots with separation interval $\Delta = N/N_p$, i.e.,

$$i_m - i_{m-1} = N/N_p \quad 1 \le m \le N_p - 1. \tag{37}$$

Correspondingly, $\boldsymbol{D}$ reduces to

$$\boldsymbol{D} = N_p \times \boldsymbol{I}_L \tag{38}$$

and $\gamma_{\mathrm{MLE}}(n)$ becomes

$$\gamma_{\mathrm{MLE}}(n) = \frac{\sigma^2 L}{N_p}. \tag{39}$$

Assuming that the entries of $\boldsymbol{h}$ are zero-mean independent Gaussian random variables with variance $\sigma_k^2, k = 0, 1, \ldots, L-1$, it can be shown that

$$\gamma_{\mathrm{MMSEE}}(n) = \lambda \times \gamma_{\mathrm{MLE}}(n) \tag{40}$$

where λ is defined as

$$\lambda = \frac{1}{L} \sum_{k=1}^{L} \frac{1}{1 + \sigma^2/(\sigma_k^2 N_p)} \tag{41}$$

and is less than unity. Notice that $\gamma_{\mathrm{MLE}}(n)$ and $\gamma_{\mathrm{MMSEE}}(n)$ are independent of the subcarrier index and coincide with the total estimation errors Γ_{MLE} and Γ_{MMSEE}. As is intuitively expected, (39) and (40) say that the estimation accuracy degrades as the number of channel taps increases and the number of pilots decreases.

Equation (37) indicates the optimal pilots' locations when there are no virtual carriers. Unfortunately, such a condition is not always met in practice. When some carriers at the edges of the spectrum are suppressed, the only way to determine the optimal pilots' locations is through exhaustive search. However, some qualitative conclusions can be drawn by inspection of Figs. 2 and 3, which show the BMSE for MLE and MMSEE as a function of the subcarrier index. Solid lines represent theoretical results as given by (30) and (35), whereas vertical bars indicate the pilots' locations. The operating conditions in Fig. 2 are as follows.

i) The channel has an exponential power delay profile $\sigma_k^2 = \exp(-k/10), k = 0, 1, \ldots, L-1$, with $L = 20$.
ii) DFT size $N = 512$.
iii) Modulated subcarriers $2N_\alpha + 1 = 433$.

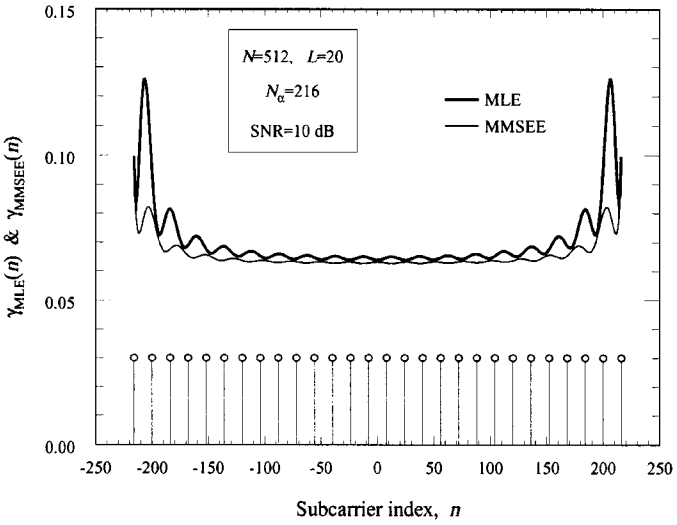

Fig. 2. BMSE versus subcarrier index n with uniformly spced pilots.

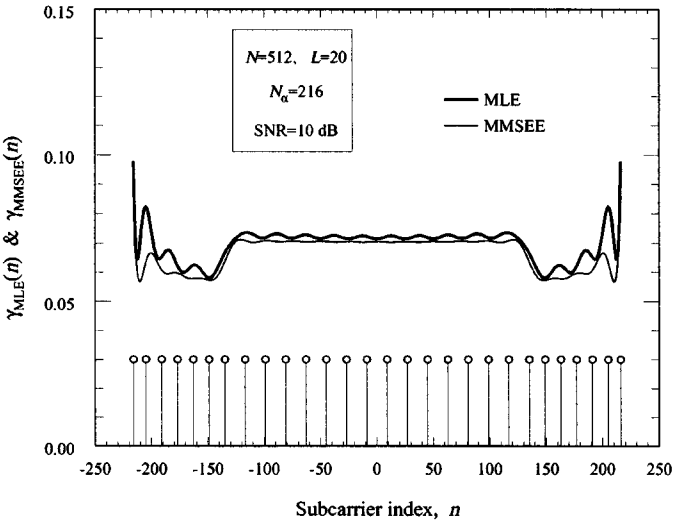

Fig. 3. BMSE versus subcarrier index n with nonuniformly spced pilots.

iv) Twenty eight uniformly spaced pilot tones, i.e.,

$$i_m = -216 + 16m \quad 0 \le m \le 27. \tag{42}$$

v) SNR $= 10$ dB.

It is worth noting that (37) is not met here since the number of suppressed carriers is greater than N/N_p.

It is seen that $\gamma_{\mathrm{MLE}}(n)$ and $\gamma_{\mathrm{MMSEE}}(n)$ are flat in the middle of the signal bandwidth, but $\gamma_{\mathrm{MLE}}(n)$ grows rapidly at the edges. This implies a reduced reliability of the data at the boundaries of the spectrum. A simple way to alleviate this drawback is to decrease the pilots' distance at the edges, as indicated in Fig. 3. In conclusion, in the presence of suppressed carriers, the minimum BMSE can be achieved with nonuniformly spaced pilots.

Fig. 4 shows the total BMSE versus L for MLE with $N = 512, N_\alpha = 216$ and SNR $= 10$ dB. The upper curve corresponds to 28 pilots at the locations in (42), whereas the lower curve corresponds to 44 pilots at

$$i_m = -215 + 10m \quad 0 \le m \le 43. \tag{43}$$

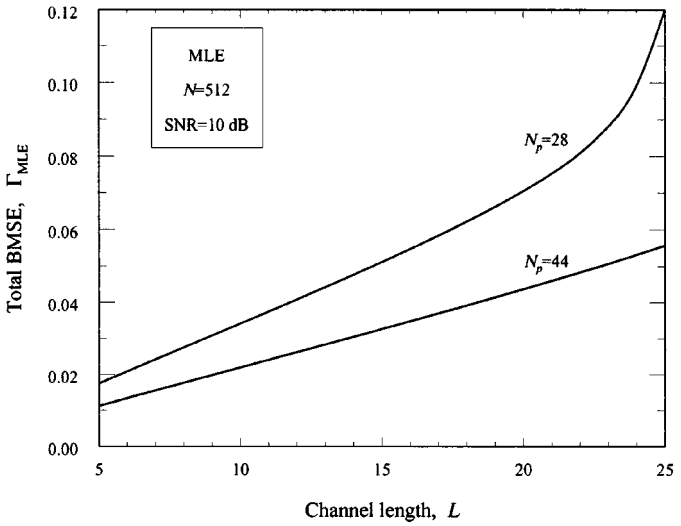

Fig. 4. Γ_{MLE} versus the channel length L with $N_p = 28, 44$.

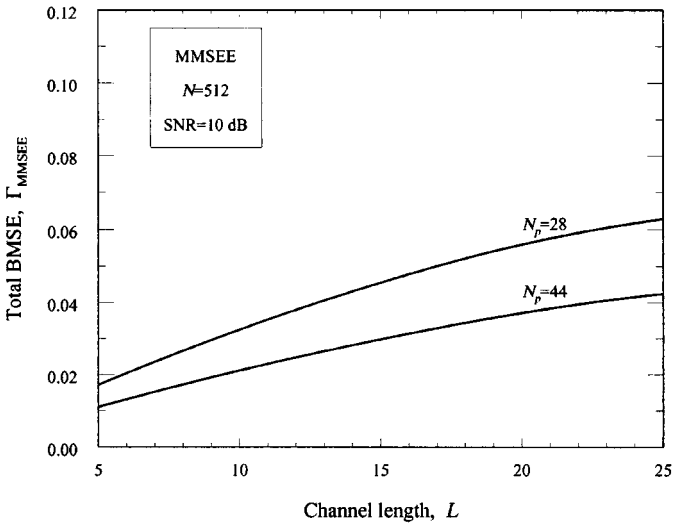

Fig. 5. Γ_{MMSEE} versus the channel length L with $N_p = 28, 44$.

As expected, the estimation accuracy improves as the number of pilots increases and the channel length decreases. Note that Γ_{MLE} shows a rapid growth as L approaches N_p. This indicates that setting $N_p = L$, as is done in [10], may be restrictive.

Fig. 5 shows the total BMSE versus L for MMSEE under the same operating conditions as in Fig. 4. Again, an exponential power delay profile as in Figs. 2 and 3 is assumed. Compared with Fig. 4, it is seen that the growth of BMSE is limited as L approaches N_p.

The ratio $\Gamma_{\mathrm{MLE}}/\Gamma_{\mathrm{MMSEE}}$ versus L is shown in Fig. 6 for $N_p = 28$ with SNR as a parameter. It is seen that MLE and MMSEE have comparable performance at intermediate/high SNR values, provided that the number of pilots is sufficiently larger than the channel length. However, the MMSEE is superior either at low SNR or for L close to N_p. As mentioned in the introduction, this superiority comes from the knowledge of the covariance matrix $\boldsymbol{C}_h$ and the noise level σ^2.

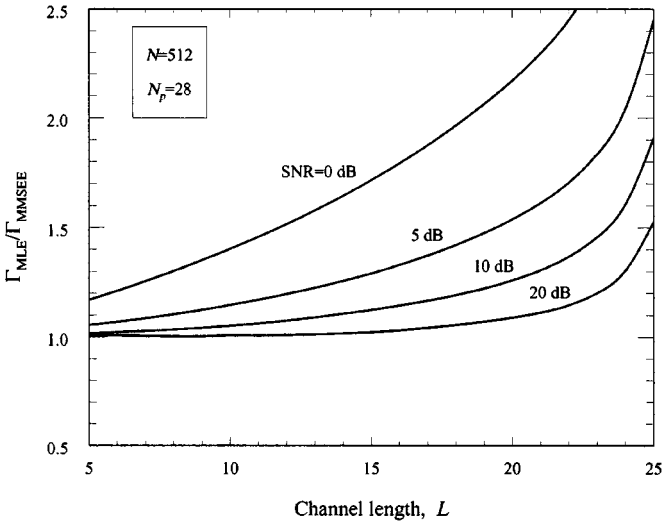

Fig. 6. $\Gamma_{\mathrm{MLE}}/\Gamma_{\mathrm{MMSEE}}$ versus the channel length L with $N_p = 28$.

V. Simulation Results

Computer simulations have been run to check and extend the analytical results of the previous section. The simulated system is as follows.

A. System Parameters

i) DFT size $N = 2048$.
ii) Modulated subcarriers $2N_\alpha + 1 = 1801$.
iii) There is a cyclic prefix of 50 samples.
iv) The channel has L paths, with path delays of $0, 1, \ldots, L-1$ samples. The amplitude A_k of each path varies independently of the others, according to a Rayleigh distribution with an exponential power delay profile, i.e.,

$$E\left\{A_k^2\right\} = \exp(-k/10) \quad k = 0, 1, \ldots, L-1. \tag{44}$$

Correspondingly, we have

$$\boldsymbol{C}_h = \mathrm{diag}\left\{\sigma_0^2, \sigma_1^2, \ldots, \sigma_{L-1}^2\right\} \tag{45}$$

where $\sigma_k^2 = E\{A_k^2\}$. The phase shift on each path is uniformly distributed over $[0, 2\pi)$, and the channel length is $L = 40$.

v) The channel is static over the OFDM symbol duration [7]. A new channel is generated at each simulation run, and the system performance is averaged over the CIR realizations.
vi) Two configurations are considered for the pilots.
   1) $N_p = 151$, with pilots' locations at

$$i_m = -900 + 12m \quad 0 \le m \le 150. \tag{46}$$

   2) $N_p = 76$, with pilots' locations at

$$i_m = -900 + 24m \quad 0 \le m \le 75. \tag{47}$$

   The first case corresponds to the standard for digital TV [15].
vii) The energy of the useful signal (including the pilot tones) at the DFT output is normalized to unity. Correspondingly, the SNR equals $1/\sigma^2$, where σ^2 is the variance of the Gaussian noise.

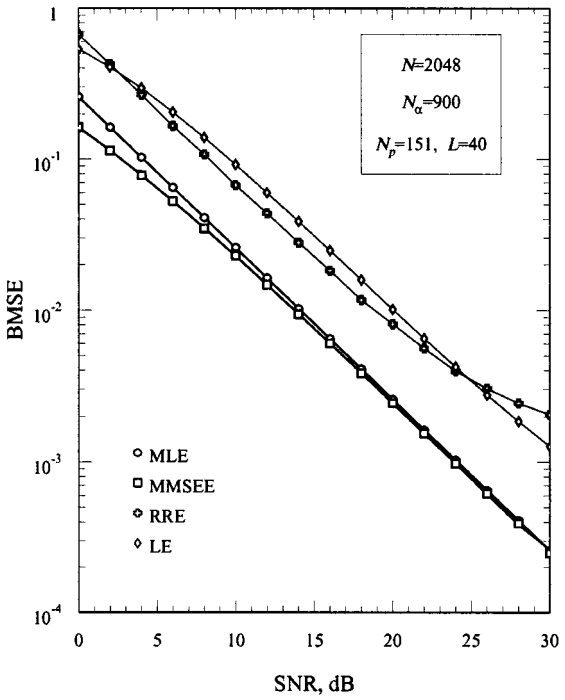

Fig. 7. Total BMSE versus SNR with $N_p = 151$.

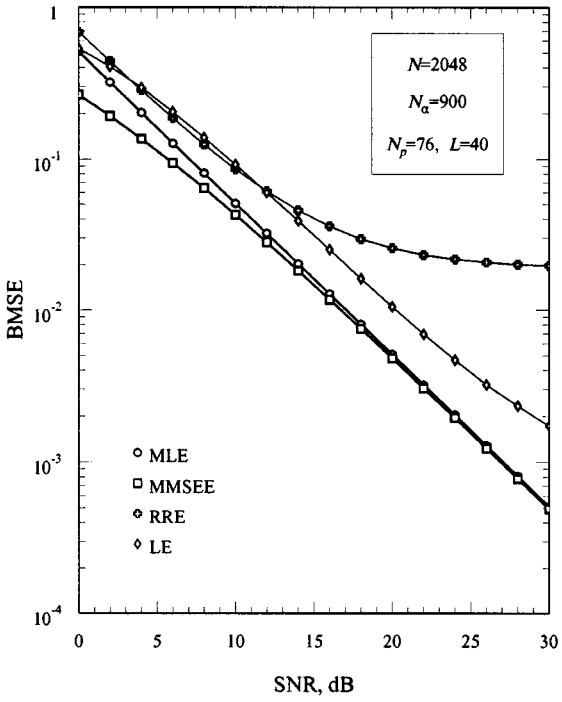

Fig. 8. Total BMSE versus SNR with $N_p = 76$.

B. Performance Comparisons

Performance comparisons have been made in terms of total BMSE and symbol error rate (SER). Figs. 7 and 8 show the total BMSE versus SNR for MLE, MMSEE, and the estimators proposed by Rinne and Renfors (RRE) [6] and by Li (LE) [9]. The RRE makes a piecewise-linear interpolation of the type

$$\hat{H}(n) = Z(m-1) + \frac{Z(m) - Z(m-1)}{i_m - i_{m-1}}(n - i_{m-1}) \qquad i_m \le n \le i_{m-1}. \quad (48)$$

The estimator in [9] holds for a general time-varying channel. With a static channel, it boils down to a cardinal interpolation between pilots

$$\hat{H}(n) = \sum_{m=0}^{N_p-1} Z(m) p_{\mathrm{LE}}(n - m\Delta) \quad (49)$$

with

$$p_{\mathrm{LE}}(n) = \frac{1}{1+\sigma^2} \times \frac{\sin(\pi n/\Delta)}{\pi n/\Delta} \quad (50)$$

and Δ is the pilots' spacing. Note that the LE needs knowledge of the noise variance σ^2. In the simulations shown in Figs. 7–10, it is assumed that σ^2 is perfectly known.

The pilots in Fig. 7 are 151, whereas in Fig. 8, they are 76. Marks indicate simulations. Solid lines represent theoretical results only for MLE and MMSEE, whereas for RRE and LE, they are drawn to make them easier to read. It is seen that MMSEE has the best performance at low SNRs. At intermediate and high SNRs, on the other hand, MLE and MMSEE are comparable. The loss of LE from MLE and MMSEE is approximately 6 dB in Fig. 7 and 3 dB in Fig. 8. As for the RRE, it exhibits a floor that worsens as the number of pilots decreases.

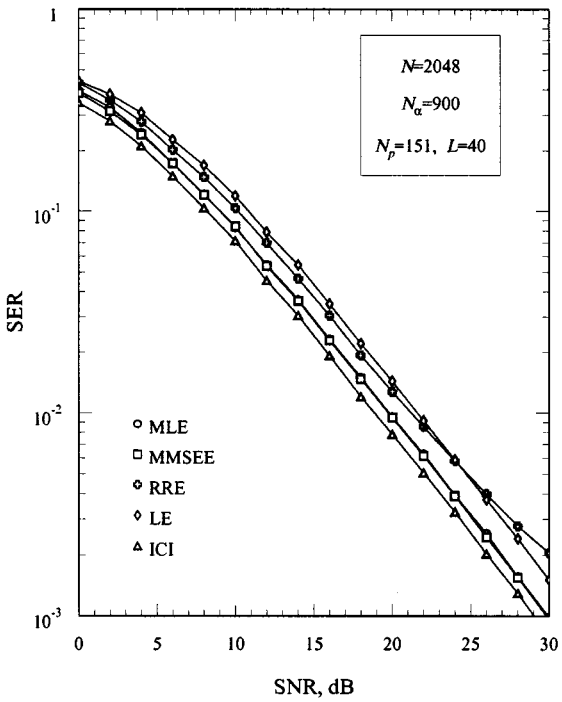

Fig. 9. SER versus SNR with $N_p = 151$.

Figs. 9 and 10 show the influence of the estimation accuracy on the SER. The symbols are taken from a QPSK constellation, and the operating conditions are the same as in Figs. 7 and 8. The curve labeled ideal channel information (ICI) corresponds to perfect knowledge of the channel transfer function at the receiver. It is seen that MLE and MMSEE have virtually the same performance at all SNRs. The loss from ICI is 1 dB in Fig. 9 and 2 dB in Fig. 10.

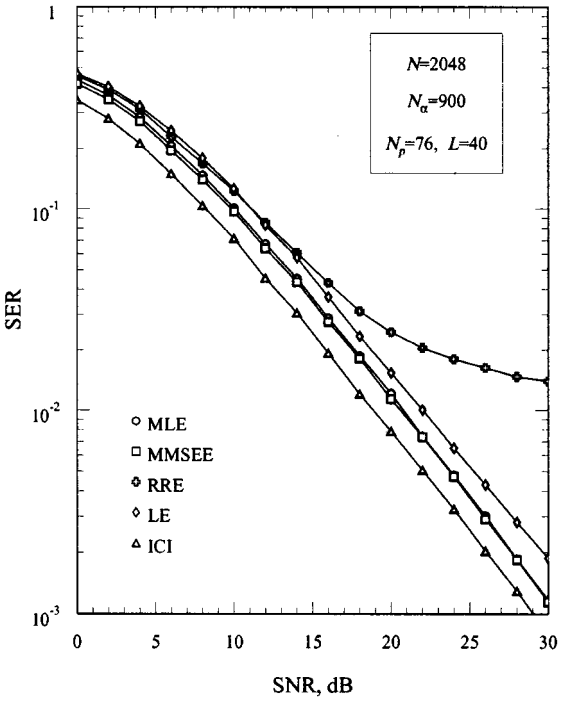

Fig. 10. SER versus SNR with $N_p = 76$.

VI. Conclusion

Two channel estimation schemes for OFDM systems have been compared. The main advantage of MLE over MMSEE is that it does not require knowledge of the channel statistics and the SNR, and therefore, it is simpler to implement. On the other hand, under certain operating conditions, the MMSEE has better accuracy as it exploits prior information about the channel. Specifically, the following has been found.

i) The channel estimates at the edges of the bandwidth are worse than those in the middle. A possible remedy is to adopt a denser pilot spacing at the edges.
ii) MMSEE performs better than MLE at low SNR.
iii) At intermediate and high SNRs, the two schemes have comparable performance, provided that the number of pilots is sufficiently larger than the duration of the CIRs.

Comparisons have also been made with the estimators proposed in [6] and [9]. It turns out that the loss in performance of LE with respect to MLE and MMSEE is limited, whereas that of RRE may be significant, unless the number of pilots is sufficiently high.

Appendix A

In this Appendix, we compute the MSE for MLE. Substituting (13) into (12) and using (8) produces

$$\hat{\boldsymbol{H}}_{\mathrm{MLE}} = \boldsymbol{G}\boldsymbol{h} + \boldsymbol{G}\boldsymbol{D}^{-1}\boldsymbol{B}^H\tilde{\boldsymbol{w}}. \tag{A1}$$

Then, bearing in mind that $\tilde{\boldsymbol{w}}$ has zero mean, from (A1) and (10), it is seen that $\hat{\boldsymbol{H}}_{\mathrm{MLE}}$ is unbiased.

Using (A1) and (14), the covariance matrix of $\hat{\boldsymbol{H}}_{\mathrm{MLE}}$ is found to be

$$\boldsymbol{C}_{\hat{H}} = \sigma^2\boldsymbol{G}\boldsymbol{D}^{-1}\boldsymbol{G}^H. \tag{A2}$$

Hence, the MSE for $H(n)$ is

$$\gamma_{\mathrm{MLE}}(n) = \sigma^2\sum_{k=0}^{L-1}\sum_{m=0}^{L-1}[\boldsymbol{D}^{-1}]_{k,m}e^{j2\pi n(m-k)/N}. \tag{A3}$$

Next, we show that $\hat{\boldsymbol{H}}_{\mathrm{MLE}}$ achieves the CRLB. To this end, we observe that $\boldsymbol{H}$ is a linear transformation of $\boldsymbol{h}$ [see (10)]. As estimation efficiency is maintained over linear transformations [13, p. 47], it suffices to demonstrate that $\hat{\boldsymbol{h}}_{\mathrm{MLE}}$ is an efficient estimator. Accordingly, we first compute the CRLB for $\boldsymbol{h}$, and then, we compare it with the MSE of $\hat{\boldsymbol{h}}_{\mathrm{MLE}}$.

Call $\boldsymbol{h}_R$ and $\boldsymbol{h}_I$ the real and imaginary components of $\boldsymbol{h}$, and define $\boldsymbol{\varphi} = (\boldsymbol{h}_R^T\boldsymbol{h}_I^T)^T$. The components of the Fisher information matrix for $\boldsymbol{\varphi}$ are given by

$$[\boldsymbol{F}]_{i,j} = -E\left\{\frac{\partial\ln\Lambda(\boldsymbol{Z};\boldsymbol{\varphi})}{\partial\varphi(i)\partial\varphi(j)}\right\} \tag{A4}$$

where $\Lambda(\boldsymbol{Z};\boldsymbol{\varphi})$ is the probability density function of $\boldsymbol{Z}$ for $\boldsymbol{Z}$, given $\boldsymbol{\varphi}$

$$\Lambda(\boldsymbol{Z};\boldsymbol{\varphi}) = \frac{1}{(\pi\sigma^2)^{N_p}} \times\exp\left\{-\frac{1}{\sigma^2}[\boldsymbol{Z}-\boldsymbol{B}\boldsymbol{h}]^H[\boldsymbol{Z}-\boldsymbol{B}\boldsymbol{h}]\right\}. \tag{A5}$$

Substituting (A5) into (A4) yields

$$\boldsymbol{F} = \frac{2}{\sigma^2}\begin{bmatrix}\mathrm{Re}\{\boldsymbol{D}\} & -\mathrm{Im}\{\boldsymbol{D}\}\\ \mathrm{Im}\{\boldsymbol{D}\} & \mathrm{Re}\{\boldsymbol{D}\}\end{bmatrix} \tag{A6}$$

where $\boldsymbol{D}$ is the matrix defined in (14). Then, we have

$$\boldsymbol{F}^{-1} = \frac{\sigma^2}{2}\begin{bmatrix}\mathrm{Re}\{\boldsymbol{D}^{-1}\} & -\mathrm{Im}\{\boldsymbol{D}^{-1}\}\\ \mathrm{Im}\{\boldsymbol{D}^{-1}\} & \mathrm{Re}\{\boldsymbol{D}^{-1}\}\end{bmatrix}. \tag{A7}$$

The CRLB is given by

$$\mathrm{CRLB}(\boldsymbol{h}) = \mathrm{tr}\{\boldsymbol{F}^{-1}\} \tag{A8}$$

where $\mathrm{tr}\{\cdot\}$ indicates the trace of a matrix. Substituting (A7) into (A8) and bearing in mind that $\boldsymbol{D}^{-1}$ is Hermitian produces

$$\mathrm{CRLB}(\boldsymbol{h}) = \sigma^2\mathrm{tr}\{\boldsymbol{D}^{-1}\}. \tag{A9}$$

On the other hand, from (8) and (13), it follows that $\hat{\boldsymbol{h}}_{\mathrm{MLE}}$ has covariance matrix

$$\boldsymbol{C}_{\hat{h}} = \sigma^2\boldsymbol{D}^{-1}. \tag{A10}$$

Comparing (A9) with (A10), it follows that the variance of $\hat{\boldsymbol{h}}_{\mathrm{MLE}}$ coincides with the CRLB, and this concludes the proof.

References

[1] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Trans. Commun.*, vol. COMM-19, pp. 628–634, Oct. 1971.
[2] L. J. Cimini Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COMM-33, pp. 665–675, July 1985.
[3] B. Le Floch, R. Halbert-Lasalle, and D. Castellain, "Digital audio broadcasting to mobile receivers," *IEEE Trans. Consum.Electron.*, vol. 35, no. 3, pp. 493–503, Aug. 1989.
[4] H. Sari, G. Karam, and J. Janclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 36, pp. 100–109, Feb. 1995.

[5] A. S. Macedo and E. S. Sousa, "Coded OFDM for broadband indoor wireless systems," in *Proc. Int. Commun. Conf.*, Montreal, QC, Canada, June 1997.
[6] J. Rinne and M. Renfors, "Pilot spacing in orthogonal frequency division multiplexing systems on practical channels," *IEEE Trans. Consum. Electron.*, vol. 42, pp. 959–962, Nov. 1996.
[7] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, pp. 931–939, July 1998.
[8] Y. Li, L. J. Cimini Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, July 1998.
[9] Y. Le, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1207–1215, July 2000.
[10] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, pp. 1122–1128, Aug. 1998.
[11] J. A. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, pp. 5–14, May 1990.
[12] M. Luise, R. Reggiannini, and G. M. Vitetta, "Blind equalization/detection for OFDM signals over frequency-selective channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1568–1578, Oct. 1998.
[13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[14] J. D. Markel, "FFT pruning," *IEEE Trans. Audio Electroacoust.*, vol. AV-19, pp. 305–311, Dec. 1971.
[15] *Digital Broadcasting Systems for Television, Sound and Data Services*, Sept. 1996.

Michele Morelli was born in Pisa, Italy, in 1965. He received the Laurea degree (cum laude) in electrical engineering and the "Premio di Laurea SIP" degree from the University of Pisa, in 1991 and 1992, respectively. From 1992 to 1995, he was with the Department of Information Engineering, University of Pisa, where he received the Ph.D. degree in electrical engineering.

He is currently a Research Fellow at the Centro Studi Metodi e Dispositivi per Radiotrasmissioni, Italian National Research Council (CNR), Pisa. His interests are in digital communication theory with emphasis on synchronization algorithms.

Umberto Mengali (M'69–SM'85–F'90) received his education in electrical engineering from the University of Pisa, Pisa, Italy. In 1971, he received the Libera Docenza degree in telecommunications from the Italian Education Ministry.

Since 1963, he has been with the Department of Information Engineering, University of Pisa, where he is a Professor of Telecommunications. In 1994, he was a Visiting Professor at the University of Canterbury, Christchurch, New Zealand, as an Erskine Fellow. His research interests are in digital communications and communication theory with emphasis on synchronization methods and modulation techniques. He has published approximately 80 journal papers and has co-authored the book *Synchronization Techniques for Digital Receivers* (New York: Plenum, 1997).

Prof. Mengali is a member of the Communication Theory Committee and was an associate editor of the IEEE TRANSACTIONS ON COMMUNICATIONS from 1985 to 1991. He is now Editor for Communication Theory of the *European Transactions on Telecommunications*. He is listed in *American Men and Women in Science*.