THE CRAMER-RAO BOUND AND ML ESTIMATE FOR DATA-AIDED CHANNEL ESTIMATION IN KSP-OFDM

Heidi Steendam, Marc Moeneclaey, Herwig Bruneel
TELIN Dept., Ghent University,Sint-Pietersnieuwstraat 41, 9000 GENT, BELGIUM

ABSTRACT

In this paper, we derive the Cramer-Rao bound (CRB) for data-aided channel estimation for OFDM with known symbol padding (KSP-OFDM). The pilot symbols used to estimate the channel are distributed over the guard interval and OFDM carriers, in order to keep the guard interval length as small as possible. An analytical expression for the CRB is obtained by performing a proper linear transformation on the observed samples. At low SNR, the CRB corresponds to the low SNR limit of the CRB obtained in [1], where it is assumed that the influence of the data symbols on the channel estimation can be neglected. At high SNR, the CRB is determined by the observations that are independent of the data symbols; the observations that are affected by data symbols are neglected. The CRB depends on the number of pilots and slightly increases with increasing guard interval length, but is essentially independent of the FFT size and the used pilot sequence. Further, a low complexity ML channel estimation technique is derived based on the linear transformation. Although in this estimation technique only a part of the observation is used, the mean squared error (MSE) performance of this estimate reaches the CRB for a large range of SNR, but a high SNR, the MSE reaches an error floor caused by the approximations made in the derivation.

I. INTRODUCTION

Multicarrier systems have received considerable attention for high data rate communications [2] because of their robustness to channel dispersion. To cope with channel dispersion, the multicarrier system uses a guard interval, with a length larger than the channel impulse response, between blocks of data. The most commonly used types of guard interval are cyclic prefix, zero padding and known symbol padding. In cyclic prefix OFDM, the guard interval consists of a cyclic extension of the data block whereas in zero-padding OFDM, no signal is transmitted during the guard interval [3]. In OFDM with known symbol padding (KSP-OFDM), the guard interval type that is considered in this paper, the guard interval consists of a number of known samples [4].

As in KSP-OFDM, the samples in the guard interval are known, they can serve as pilot symbols to obtain a data-aided estimate of the channel. However, as the length of the guard interval is typically small as compared to the FFT length (to keep the efficiency of the multicarrier system as high as possible) the number of known samples is typically too small to obtain an accurate estimate for the channel. To improve the channel estimation, the number of pilot symbols must be increased. This can be done by increasing the guard interval length, which is not favorable as this will reduce the OFDM system efficiency, or by keeping the length of the guard interval constant and replacing in the data part of the signal some data carriers by pilot carriers [1].

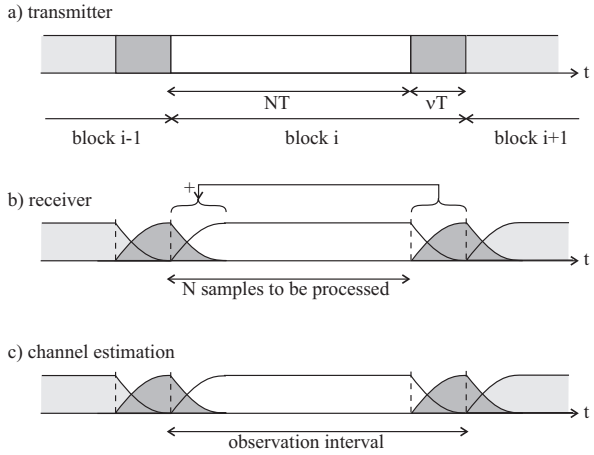

Figure 1: Time-domain signal of KSP-OFDM a) transmitted signal b) received signal and observation interval for data detection c) observation interval for channel estimation.

In this paper, we derive the Cramer-Rao bound (CRB) for channel estimation when the pilot symbols are distributed over the guard interval and pilot carriers. The paper is organized as follows. In section II., we describe the system and determine the CRB. Further, we derive a low complexity ML estimate for the channel. Numerical results are given in section III. and the conclusions are drawn in section IV..

II. SYSTEM MODEL AND CRAMER-RAO BOUND

In KSP-OFDM, the data symbols to be transmitted are grouped into blocks of N symbols: $\mathbf{a}_i = (a_i(0), \ldots, a_i(N-1))^T$. The data symbols are then modulated on the OFDM carriers using an N-point inverse FFT. The guard interval consisting of ν known samples is inserted after each OFDM symbol (this corresponds to the dark gray area in figure 1(a)), resulting in the time-domain samples $\mathbf{s}_i$ during block i:

$$\mathbf{s}_i = \sqrt{\frac{N}{N+\nu}} \begin{pmatrix} \mathbf{F}^+ \mathbf{a}_i \\ \mathbf{b}_g \end{pmatrix} \tag{1}$$

where $\mathbf{F}$ is the $N \times N$ matrix corresponding to the FFT operation, i.e. $\mathbf{F}_{k,\ell} = \frac{1}{\sqrt{N}} e^{-j2\pi \frac{k\ell}{N}}$, and $\mathbf{b}_g = (b_g(0), \ldots, b_g(\nu-1))^T$ corresponds to the ν known samples of the guard interval. We assume that the data symbols are statistically independent with $E[a_i(n) a_{i'}^*(n')] = E_s \delta_{i,i'} \delta_{n,n'}$. Further, we assume that the known samples $\mathbf{b}_g$ have the same energy per sample as the data symbols: $E[|b_g(n)|^2] = E_s$.

The sequence (1) is transmitted over the dispersive channel and disturbed by additive white Gaussian noise $\mathbf{w}$. The zero-

1-4244-1144-0/07/$25.00©2007 IEEE

mean noise components $w(k)$ have variance N_0. Without loss of generality, we consider the detection of the OFDM block with index $i = 0$. The received time-domain samples can be written as

$$\mathbf{r} = \sqrt{\frac{N}{N+\nu}} \sum_{i=-\infty}^{+\infty} \mathbf{H}^{(i)} \mathbf{s}_i + \mathbf{w} \tag{2}$$

where $\mathbf{H}^{(i)}_{k,k'} = h_i(k-k')$ is the $(N+\nu) \times (N+\nu)$ channel matrix and $\mathbf{h}_i = (h_i(0), \ldots, h_i(L-1))^T$ is the vector of L channel taps corresponding to OFDM block i, i.e. we assume the channel changes slowly as compared to the duration of an OFDM block. To avoid intersymbol interference, we assume that the duration of the guard interval exceeds the duration of the channel impulse response, i.e. $\nu \geq L-1$. For data detection, the known samples are first subtracted from the received signal. Then, the ν samples of the guard interval are added to the first ν samples of the data part of the block, as shown in figure 1(b), and the resulting N samples are applied to an FFT. As the known samples are disturbed by the channel (as can be seen in figure 1(b)), the channel needs to be known in order to remove the contribution of the known samples from the received signal.

To estimate the channel, we assume M pilot symbols are available. As we select the length of the guard interval in function of the channel impulse length and not in function of the precision of the estimation, only ν of the M pilot symbols can be placed in the guard interval. This implies that $M-\nu$ data carriers in (1) are replaced by pilot carriers. The ν pilot symbols located in the guard interval are denoted $\mathbf{b}_g = (b_g(0), \ldots, b_g(\nu-1))^T$ and the $M-\nu$ pilots transmitted on the carriers $\mathbf{b}_c = (b_c(0), \ldots, b_c(M-\nu-1))^T$. We define I_p and I_d as the sets of carriers modulated by the pilot symbols and the data symbols, respectively, with $I_p \cup I_d = \{0, \ldots, N-1\}$. We assume that the pilot symbols $\mathbf{b}_c$ have the same energy per symbol as the data symbols. It can easily be verified that the observation of the $N+\nu$ time-domain samples corresponding to the ith OFDM block (as shown in figure 1(c)) contains sufficient information to estimate $\mathbf{h}_i$. Rewriting (2), we obtain

$$\mathbf{r} = \mathbf{B}\mathbf{h}_0 + \tilde{\mathbf{w}} \tag{3}$$

where $\mathbf{B} = \mathbf{B}_g + \mathbf{B}_c$ is a $(N+\nu) \times L$ matrix. The matrix $\mathbf{B}_g$ contains the contributions from the pilot symbols in the guard interval, and is given by

$$(\mathbf{B}_g)_{k,\ell} = b_g(|k-\ell+\nu|_{N+\nu}) \tag{4}$$

where $|x|_K$ is the modulo-K operation of x yielding a result in the interval $[0, K[$, and $b_g(k) = 0$ for $k \geq \nu$. The matrix $\mathbf{B}_c$ consists of the contributions from the pilots transmitted on the carriers, where

$$(\mathbf{B}_c)_{k,\ell} = s_p(k-\ell). \tag{5}$$

The vector $\mathbf{s}_p$ equals the N-point IFFT of the pilot carriers only, i.e. $\mathbf{s}_p = \mathbf{F}_p \mathbf{b}_c$. The $N \times (M-\nu)$ matrix $\mathbf{F}_p$ consists of a subset of columns of the IFFT matrix $\mathbf{F}^+$ corresponding to the set I_p of pilot carriers. Note that $s_p(k) = 0$ for $k < 0$ or $k \geq N$. The disturbance in (3) can be written as

$$\tilde{\mathbf{w}} = \mathbf{H}\mathbf{F}_d \mathbf{a} + \mathbf{w} \tag{6}$$

where $\mathbf{H}_{k,\ell} = h_0(k-\ell)$ is a $(N+\nu) \times N$ matrix. The $N \times (N+\nu-M)$ matrix $\mathbf{F}_d$ consists of a subset of columns of $\mathbf{F}^+$ corresponding to the set I_d of data carriers, and $\mathbf{a}$ is the vector of $N+\nu-M$ data symbols transmitted during the observed OFDM block. Hence, the contribution $\mathbf{s}_d = \mathbf{F}_d \mathbf{a}$ equals the N-point IFFT of the data carriers only.

First note that the number $N+\nu-M$ of data symbols $\mathbf{a}$ is smaller than the number $N+\nu$ of observed samples $\mathbf{r}$. Therefore, it is possible to find an invertible linear transformation $\mathbf{T}$ that maps $\mathbf{r}$ to an $(N+\nu) \times 1$ vector $\mathbf{r}' = [\mathbf{r}_1^T \mathbf{r}_2^T]^T$ where $\mathbf{r}_1$ depends on the transmitted data symbols and $\mathbf{r}_2$ is independent of $\mathbf{a}$. This transform can be found by performing the QR-decomposition of the matrix $\mathbf{H}\mathbf{F}_d$, i.e. $\mathbf{H}\mathbf{F}_d = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q}$ is a unitary matrix $\mathbf{Q}^+ = \mathbf{Q}^{-1}$ and

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_u \\ \mathbf{0} \end{pmatrix} \tag{7}$$

where $\mathbf{R}_u$ is an upper triangular matrix. Taking into account the dimensions of $\mathbf{H}\mathbf{F}_d$, it follows that $\mathbf{H}\mathbf{F}_d$ (and thus $\mathbf{R}$) has rank $\leq N+\nu-M$. Assuming the rank of $\mathbf{H}\mathbf{F}_d$ equals $N+\nu-M-x$, then $\mathbf{R}$ contains $M+x$ zero rows, i.e. $\mathbf{R}_u$ is a $(N+\nu-M-x) \times (N+\nu-M)$ matrix and the all zero matrix $\mathbf{0}$ in (7) is a $(M+x) \times (N+\nu-M)$ matrix. The transform $\mathbf{T}$ is then given by $\mathbf{T} = \mathbf{Q}^+$, and the resulting observations yield

$$\mathbf{r}' = \mathbf{T}\mathbf{r} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} \mathbf{h} + \begin{pmatrix} \mathbf{R}_u \\ \mathbf{0} \end{pmatrix} \mathbf{a} + \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{pmatrix}. \tag{8}$$

In (8), $\mathbf{B}_1$ and $\mathbf{B}_2$ correspond to the first $(N+\nu-M-x)$ and last $(M+x)$ rows of $\mathbf{T}\mathbf{B}$, respectively. Because of the unitary nature of the matrix $\mathbf{Q}$, the noise contributions $\mathbf{w}_1$ and $\mathbf{w}_2$ are statistically independent and have the same mean and variance as the noise $\mathbf{w}$.

Assuming the data symbols $\mathbf{a}$ are zero-mean Gaussian distributed, $\mathbf{r}'$ given $\mathbf{h}_0$ is Gaussian distributed, i.e. $\mathbf{r}'|\mathbf{h}_0 \sim N(\mathbf{Q}^+\mathbf{B}\mathbf{h}_0, \mathbf{R}_{\mathbf{w}'})$ where $\mathbf{R}_{\mathbf{w}'}$ is the autocorrelation matrix of the noise term $\mathbf{w}' = \mathbf{T}\tilde{\mathbf{w}}$ and is given by

$$\mathbf{R}_{\mathbf{w}'} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 \end{pmatrix} \tag{9}$$

where $\mathbf{R}_1 = E_s \frac{N}{N+\nu} \mathbf{R}_u \mathbf{R}_u^+ + N_0 \mathbf{I}_{N+\nu-M-x}$, $\mathbf{R}_2 = N_0 \mathbf{I}_{M+x}$ and $\mathbf{I}_K$ is the $K \times K$ identity matrix.

The Cramer-Rao bound is defined by $\mathbf{R}_{\mathbf{h}-\hat{\mathbf{h}}} - \mathbf{J}^{-1} \geq 0$ [5], where $\mathbf{R}_{\mathbf{h}-\hat{\mathbf{h}}}$ is the autocorrelation matrix of the estimation error $\mathbf{h}_0 - \hat{\mathbf{h}}_0$, $\hat{\mathbf{h}}_0$ is an estimate of $\mathbf{h}_0$ and the Fisher information matrix $\mathbf{J}$ is defined as

$$\mathbf{J} = E_{\mathbf{r}'} \left[\left(\frac{\partial}{\partial \mathbf{h}_0} \ln p(\mathbf{r}'|\mathbf{h}_0) \right)^+ \left(\frac{\partial}{\partial \mathbf{h}_0} \ln p(\mathbf{r}'|\mathbf{h}_0) \right) \right]. \tag{10}$$

Hence, the MSE of an estimator is lower bounded by $E[||\mathbf{h}_0 - \hat{\mathbf{h}}_0||^2] = \text{trace}(\mathbf{R}_{\mathbf{h}-\hat{\mathbf{h}}}) \geq \text{trace}(\mathbf{J}^{-1})$. As $\mathbf{r}_1$ and $\mathbf{r}_2$ given $\mathbf{h}_0$ are statistically independent, it can easily be verified that the Fisher information matrix is given by $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$, where

$$\mathbf{J}_i = E_{\mathbf{r}_i} \left[\left(\frac{\partial}{\partial \mathbf{h}_0} \ln p(\mathbf{r}_i|\mathbf{h}_0) \right)^+ \left(\frac{\partial}{\partial \mathbf{h}_0} \ln p(\mathbf{r}_i|\mathbf{h}_0) \right) \right] \tag{11}$$

with $i = 1, 2$. Taking into account (8), it follows that

$$\ln p(\mathbf{r}_i|\mathbf{h}_0) = C - \frac{1}{2}\ln[\![\mathbf{R}_i]\!] - (\mathbf{r}_i - \mathbf{B}_i\mathbf{h}_0)^+\mathbf{R}_i^{-1}(\mathbf{r}_i - \mathbf{B}_i\mathbf{h}_0) \tag{12}$$

where C is an irrelevant constant and $[\![\mathbf{R}_i]\!]$ is the determinant of $\mathbf{R}_i$. Note that as the transform $\mathbf{T}$ depends on the channel taps $\mathbf{h}_0$ to be estimated, we need the derivates of $\mathbf{B}_i$, $[\![\mathbf{R}_i]\!]$ and $\mathbf{R}_i^{-1}$ with respect to $\mathbf{h}_0$ to obtain the Fisher information matrix, and hence the CRB.

To simplify the analysis, we approximate the data contribution $\mathbf{HF}_d\mathbf{a}$ in (6) by $\tilde{\mathbf{F}}\tilde{\mathbf{H}}\mathbf{a}$ where the $(N+\nu)\times(N+\nu-M-x)$ matrix $\tilde{\mathbf{F}}_{k,\ell} = \frac{1}{\sqrt{N}}e^{j2\pi\frac{kn_\ell}{N}}$, $\tilde{\mathbf{H}}$ is a $(N+\nu-M-x)\times(N+\nu-M-x)$ diagonal matrix with diagonal elements H_{n_ℓ}, $n_\ell \in I_d$ and

$$H_m = \sum_{k=0}^{N-1} h_0(k)e^{-j2\pi\frac{km}{N}}. \tag{13}$$

In this approximation, we have neglected the distortion of the data samples at the edges of the observed block of samples. In that case, the transform $\mathbf{T} = \mathbf{Q}^+$ can be obtained from the QR-decomposition of $\tilde{\mathbf{F}}$, and therefore is independent of the channel to be estimated.

First, we determine $\mathbf{J}_2$. As the observation $\mathbf{r}_2 = \mathbf{B}_2\mathbf{h}_0 + \mathbf{w}_2$ is independent of the data symbols, and $p(\mathbf{r}_2|\mathbf{h}_0) \sim N(\mathbf{B}_2\mathbf{h}_0, N_0\mathbf{I}_{M+x})$, where $\mathbf{B}_2$ is independent of $\mathbf{h}_0$, it can easily be found that

$$\mathbf{J}_2 = \frac{1}{N_0}\mathbf{B}_2^+\mathbf{B}_2. \tag{14}$$

Note that the CRB of an estimation can not increase by using more observations. Hence, the CRB obtained from the observation $\mathbf{r}_2$ only is an upper bound for the CRB obtained from the whole observation $\mathbf{r}'$.

Based on the observation $\mathbf{r}_2$, we can easily obtain the ML estimate of $\mathbf{h}_0$ [1]:

$$\hat{\mathbf{h}}_{0,ML} = (\mathbf{B}_2^+\mathbf{B}_2)^{-1}\mathbf{B}_2^+\mathbf{r}_2. \tag{15}$$

The mean squared error of this estimate is given by

$$MSE = E[||\mathbf{h}_0 - \hat{\mathbf{h}}_{0,ML}||^2] = \text{trace}\left(N_0(\mathbf{B}_2^+\mathbf{B}_2)^{-1}\right). \tag{16}$$

Hence, the MSE of this estimate reaches the CRB=trace$(\mathbf{J}_2^{-1})$, i.e. the estimate is a minimum variance unbiased (MVU) estimate. However, when the approximation $\mathbf{HF}_d \approx \tilde{\mathbf{F}}\tilde{\mathbf{H}}$ no longer holds, the observation $\mathbf{r}_2$ will be affected by a residual contribution of the data symbols. In that case, the MSE of the estimate (15) yields

$$MSE = \text{trace}\left(\mathbf{D}\mathbf{R}_{\tilde{\mathbf{w}}}\mathbf{D}^+\right) \tag{17}$$

where $\mathbf{R}_{\tilde{\mathbf{w}}} = N_0\mathbf{I}_{N+\nu} + \frac{N}{N+\nu}E_s\mathbf{HF}_d\mathbf{F}_d^+\mathbf{H}^+$ is the autocorrelation matrix of the noise term $\tilde{\mathbf{w}}$ (6), $\mathbf{D} = (\mathbf{B}_2^+\mathbf{B}_2)^{-1}\mathbf{B}_2^+\mathbf{T}_2$ and $\mathbf{T}_2$ consists of the last $M+x$ rows of $\mathbf{T}$. As the transform $\mathbf{T}$ is obtained by the QR-decomposition of $\tilde{\mathbf{F}}$, and $\tilde{\mathbf{F}}$ is known when the positions of the data symbols are known, $\mathbf{B}_2$ only depends on the known pilot symbols and the known positions of the data carriers and the pilot carriers. Hence, $\mathbf{B}_2$ is known at the receiver and can be precomputed. Therefore, the estimate (15) can be obtained with low complexity. Note that the estimate (15) is only an ML estimate as long as the approximation $\mathbf{HF}_d \approx \tilde{\mathbf{F}}\tilde{\mathbf{H}}$ is valid.

Next, we determine $\mathbf{J}_1$, based on the observation $\mathbf{r}_1 = \mathbf{B}_1\mathbf{h}_0 + \mathbf{R}_u\mathbf{a} + \mathbf{w}_1$ only. Note that, although $\mathbf{B}_1$ is independent of $\mathbf{h}_0$, the autocorrelation matrix $\mathbf{R}_1$ is not. Let us consider the special case that the rank of $\tilde{\mathbf{F}}$ is $N+\nu-M$, i.e. $x = 0$[1]. When $M - \nu \ll N$, $\mathbf{R}_1$ can be approximated by

$$\mathbf{R}_1 = \mathbf{T}_1\tilde{\mathbf{F}}\mathbf{\Delta}\tilde{\mathbf{F}}^+\mathbf{T}_1^+ \tag{18}$$

where $\mathbf{\Delta}$ is a diagonal matrix with elements $N_0 + \frac{N}{N+\nu}E_s|H_{n_\ell}|^2$ and $\mathbf{T}_1$ consists of the $N+\nu-M$ first rows of $\mathbf{T}$. Because $\tilde{\mathbf{F}}$ has rank $N+\nu-M$, $\mathbf{T}_1\tilde{\mathbf{F}}$ is a full rank square matrix. When $\mathbf{A}$ and $\mathbf{B}$ are square matrices, it follows that $[\![\mathbf{AB}]\!] = [\![\mathbf{A}]\!][\![\mathbf{B}]\!]$. Hence, $\ln[\![\mathbf{R}_1]\!]$ reduces to

$$\ln[\![\mathbf{R}_1]\!] = \ln[\![\mathbf{T}_1\tilde{\mathbf{F}}\tilde{\mathbf{F}}^+\mathbf{T}_1^+]\!] + \sum_{n_\ell\in I_d}\ln\left(N_0 + \frac{N}{N+\nu}E_s|H_{n_\ell}|^2\right). \tag{19}$$

Further, as $\mathbf{T}_1\tilde{\mathbf{F}}$ has full rank, the inverse of $\mathbf{R}_1$ (18) can easily be computed:

$$(\mathbf{R}_1)^{-1} = (\tilde{\mathbf{F}}^+\mathbf{T}_1^+)^{-1}\mathbf{\Delta}^{-1}(\mathbf{T}_1\tilde{\mathbf{F}})^{-1}. \tag{20}$$

Using (19) and (20), the derivates of $\ln[\![\mathbf{R}_1]\!]$ and $(\mathbf{R}_1)^{-1}$ with respect to $\mathbf{h}_0$ can easily be computed. Defining

$$\alpha_\ell = N_0 + \frac{N}{N+\nu}E_s|H_{n_\ell}|^2 \tag{21}$$

$$\gamma_{k,\ell} = \frac{N}{N+\nu}E_sH_{n_\ell}^*e^{-j2\pi\frac{kn_\ell}{N}} \tag{22}$$

$$\beta_k = -\frac{1}{2}\sum_{n_\ell\in I_d}\frac{\gamma_{k,\ell}}{\alpha_\ell}, \tag{23}$$

it follows after tedious computations that the Fisher information matrix $\mathbf{J}_1$ is given by

$$(\mathbf{J}_1)_{k,k'} = (\mathbf{B}_1^+\mathbf{R}_1^{-1}\mathbf{B}_1)_{k,k'} + \beta_k^*\beta_{k'} + \sum_{n_\ell\in I_d}\frac{\gamma_{k,\ell}^*\gamma_{k',\ell}}{|\alpha_\ell|^2}. \tag{24}$$

Combining (14) and (24), the total Fisher information matrix, based on the observation of both $\mathbf{r}_1$ and $\mathbf{r}_2$, is given by

$$(\mathbf{J})_{k,k'} = (\mathbf{B}^+\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}\mathbf{B})_{k,k'} + \beta_k^*\beta_{k'} + \sum_{n_\ell\in I_d}\frac{\gamma_{k,\ell}^*\gamma_{k',\ell}}{|\alpha_\ell|^2} \tag{25}$$

III. Numerical Results

In this section, we evaluate the CRB's obtained from the whole observation $\mathbf{r}_1$ and $\mathbf{r}_2$ (25) and the data-free observation $\mathbf{r}_2$ only (14). Without loss of generality, we assume the comb-type pilot arrangement [6] is used for the pilots transmitted on the carriers. We assume that the pilots are equally spaced over the carriers, i.e. the positions of the pilot carriers are

[1] This assumption is valid in most practical cases, especially when $N \gg 1$.

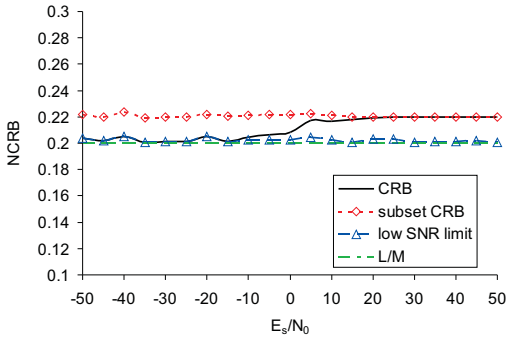

Figure 2: Normalized CRB, $\nu = 7$, $N = 1024$, $M = 40$.

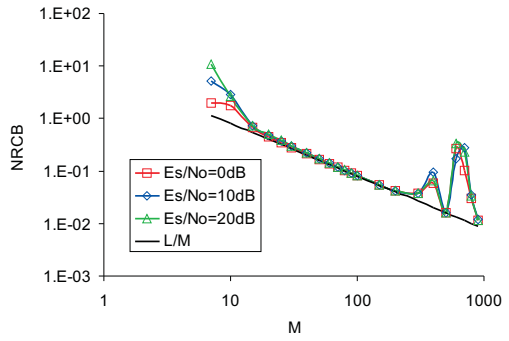

Figure 3: Influence of the number of pilots M on the CRB, $\nu = 7$, $N = 1024$.

$I_p = \{n_0 + m\delta | m = 0, \ldots, M - \nu - 1\}$, where $\delta = \left\lfloor \frac{N}{M-\nu} \right\rfloor$. Note however that the results can easily be extended for other types of pilot arrangements. Further, we assume $L = 8$ and the channel impulse response linearly decreases and is normalized: $\sum_{\ell=0}^{L-1} |h_0(\ell)|^2 = 1$. The pilot symbols are randomly generated and BPSK modulated.

In figure 2, we show the normalized CRB (NCRB=$\frac{N+\nu}{N}\frac{N_0}{E_s}$CRB) as function of the SNR= E_s/N_0 for the total observation and the subset $\mathbf{r}_2$ of observations only. Further, the low SNR limit of the CRB is shown. As expected, for low SNR (< -10 dB), the CRB of the total observation coincides with the low SNR limit of the CRB. At high SNR, the CRB reaches the CRB for the subset observation, indicating that at high SNR, only the observations that are independent of the data symbols are taken into account, whereas the observations affected by the data symbols are neglected. Further, it can be observed that the low SNR limit of the CRB is essentially equal to L/M. This indicates that the CRB is inversely proportional to the number of pilots M.

In figure 3, the NCRB is shown as function of M for different values of the SNR. The (N)CRB is inversely proportional to M for a wide range of M. At low and high values of M, the NCRB is increased as compared to L/M. This can be explained by figure 4. In figure 4, the influence of the pilot sequence is shown on the CRB. In this figure, the CRB is computed for 50 randomly generated pilot sequences. Further, the average of the CRB over the random pilot sequences is shown. Note that the CRB is only related to the values of the pilots through the first term in (25). At high values of M, the pilot spacing $\delta = 1$, such that the pilots on the carriers are no longer evenly spread over the carriers. This effect causes the peak in the curve at high M. The CRB in this case clearly depends on the values of the pilots: we observe an increase of the variance. The effect disappears when M is further growing: the spreading of the pilots over the carriers becomes again more uniform. Also at low values of M, the average value of and the variance of the CRB are increased. At low M, the contribution of the guard interval pilots is dominant. From simulations, it follows that this contribution strongly depends on the values of the pilots in the guard interval, and has large outliers when the guard interval pilots are badly chosen. The effect disappears when the number of guard interval pilots is smaller than the number of pilot carriers. However, the variation is in most cases rather small, so that we can conclude that the CRB is essentially independent of the pilot sequence.

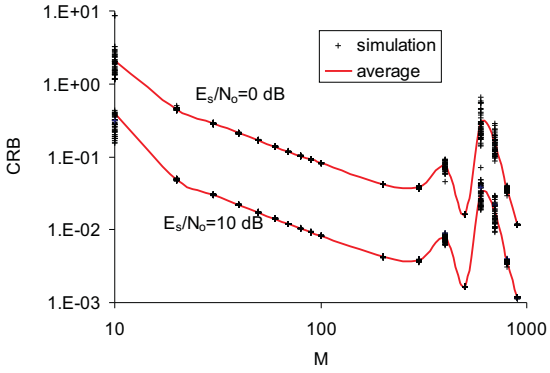

Figure 4: Influence of the pilot sequence on the CRB, $\nu = 7$, $N = 1024$.

Figure 5 shows the dependency of the NCRB on the guard interval length. It is observed that the NCRB slightly increases for increasing guard interval length. This can be explained as when ν increases, the number of guard interval pilots increases while the number of pilot carriers decreases. Hence, when ν increases, the relative importance of the contribution of the guard interval pilots will increase. As in figure 4, this will cause an increase of the CRB.

The dependency of the CRB on the FFT size N is shown in figure 6. The CRB is constant over a wide range of N. Only at low values of N, the CRB slightly increases: the assumption $M - \nu \ll N$ does not hold, so the approximation (18) is no longer valid. However, for the range of N for which the approximation (18) holds, we can conclude that the CRB is independent of N.

In figure 7, we show the CRB for both the total observation and the subset observation, along with the low SNR limit of the CRB. Although it follows from figure 2 that the CRB and the subset CRB are larger than the low SNR limit of the CRB, the difference is small: the curves in figure 7 essentially coincide.

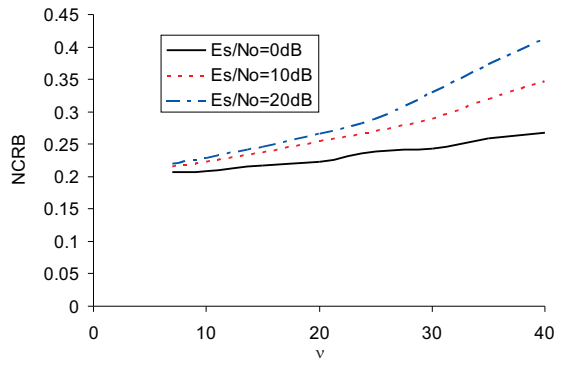

Figure 5: Influence of the guard interval length ν on the CRB, $M = 40$, $N = 1024$.

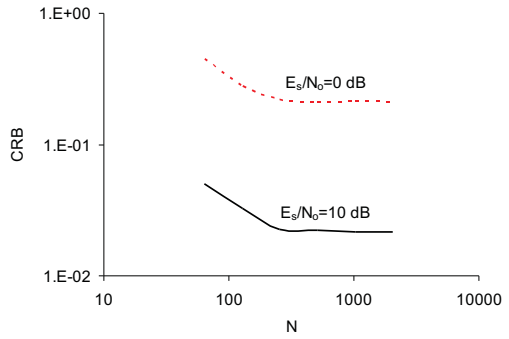

Figure 6: Influence of the FFT length N on the CRB, $M = 40$, $\nu = 7$.

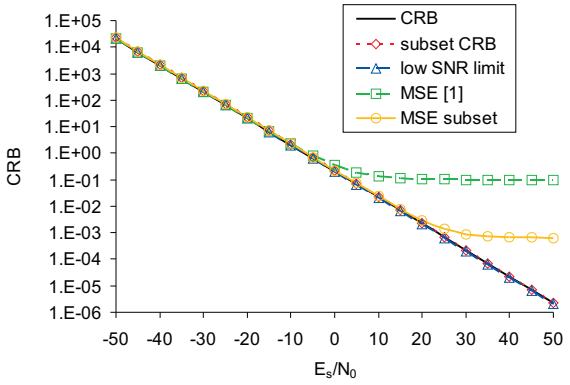

Figure 7: CRB and MSE, $N = 1024$, $M = 40$, $\nu = 7$.

In figure 7, we also show the MSE (17) of the proposed subset based estimator. As can be observed, the MSE coincides with the CRB for a large range of SNR. Only for large SNR (>20 dB), the MSE shows an error floor, indicating that the approximation $\mathbf{H}\mathbf{F}_d \approx \tilde{\mathbf{F}}\tilde{\mathbf{H}}$ is no longer valid. Further, we show in figure 7 the MSE of a suboptimal ML based estimator for the channel, derived in [1] and based on the estimator given in [7]. It is clear that the estimator proposed in this paper outperforms the estimator from [1]. Further, in the latter estimator it is assumed that the autocorrelation matrix $\mathbf{R}_{\tilde{\mathbf{w}}}$ is known (e.g. by estimating it from the received signal). Therefore, the complexity of the estimator from [1] is much higher than that of the proposed estimator.

IV. Conclusions

In this paper, we have derived the Cramer Rao bound for data-aided channel estimation in KSP-OFDM, when the pilot symbols are distributed over the guard interval and pilot carriers. An analytical expression for the CRB is found by applying a proper linear transformation to the received samples. It turns out that the CRB is essentially independent of the FFT length, the guard interval and the pilot sequence, but is inversely proportional to the number of pilots. At low SNR, the CRB obtained in this paper coincides with the low SNR limit of the CRB, derived in [1]. At high SNR, the CRB reaches the CRB corresponding to the data-independent subset of the observation, indicating that at high SNR, observations affected by data symbols are neglected. Further, we have compared the MSE of the ML subset estimator with the obtained CRB and with the MSE of the ML-based channel estimator from [1]. The proposed estimator coincides with the CRB for a large range of SNR. Only at large SNR, the MSE shows an error floor. However, the proposed estimator outperforms the estimator from [1], both on complexity and performance.

References

[1] H. Steendam, M. Moeneclaey, "Different Guard Interval Techniques for OFDM: Performance Comparison", *in Proc. of MC-SS'07*, Herrsching, Germany, May 7-9, 2007.

[2] J. A. C. Bingham, "Multicarrier modulation for data transmission, an idea whose time has come," *IEEE Comm. Mag.*, Vol. 31, pp. 514, May 1990.

[3] B. Muquet, Z. Wang, et. Al., "Cyclic Prefixing or Zero Padding for Wireless Multicarrier Transmissions?," *IEEE Trans. on Comm.*, Vol. 50, no 12, Dec 2002, pp. 2136-2148.

[4] L. Deneire, B. Gyselinckx, M. Engels, "Training Sequence versus Cyclic Prefix – A New Look on Single Carrier Communication," *IEEE Comm. Letters*, Vol.5, no 7, Jul 2001, pp. 292-294.

[5] H.L. Van Trees, *Detection, Estimation and Modulation Theory*, New York, Wiley, 1968.

[6] F. Tufvesson, T. Maseng, "Pilot Assisted Channel Estimation for OFDM in Mobile Cellular Systems," *in Proc. of IEEE Veh. Tech. Conf.*, VTC'97, Phoenix, U.S.A, May 4-7, 1997, pp. 1639-1643.

[7] O. Rousseaux, G. Leus, M. Moonen, "Estimation and Equalization of Doubly Selective Channels Using Known Symbol Padding," *IEEE Trans. on Signal Proc.*, Vol. 54, no 3, March 2006.