EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS
Eur. Trans. Telecomms. 2008; **19**:907–916
Published online 18 September 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/ett.1250

Signal Processing

EM based channel estimation method for MIMO-OFDM systems

Chongning Na, Dragan Obradovic* and Andrei Szabo

Siemens AG, Corporate Technology, Information and Communications Department, D-81730, Munich, Germany

SUMMARY

Standard Wiener interpolation for Pilot Aided Channel Estimation (PACE) in OFDM systems does not utilise other available information at the receiver such as the received signals in the non-pilot sub-channels and the knowledge about the transmitted symbol alphabet. In addition, the standard Wiener interpolation method requires the spacing of the pilot symbols to satisfy the *Sampling Theorem*, i.e. at least two pilot sub-channels within the coherence bandwidth. This paper presents two novel channel estimation methods aiming at improving and expanding the PACE. The first method is based on the standard Wiener interpolation, and it achieves better performance by feeding back the posterior probabilities of the detected symbols. The second method is aiming at performing channel estimation with the placement of pilot symbols not necessarily satisfying the *Sampling Theorem*. Both methods use the Expectation and Maximisation (EM) algorithm. Copyright © 2007 John Wiley & Sons, Ltd.

1. INTRODUCTION

The OFDM system has emerged as an appropriate technology to mitigate the effects of frequency selectivity in wide-band mobile communication systems. The use of a Cyclic Prefix (CP) for preventing inter-block interference is known to be equivalent to multiple flat fading parallel transmission channels in the frequency domain [1]. Wide-band communication systems with OFDM modulation can be combined with multiple transmit and receiver antennas (MIMO) to achieve very high data rate transmission.

Typically, channel estimation within the MIMO-OFDM systems is based on the Pilot Aided Channel Estimation (PACE) methods where the pilot symbols are periodically transmitted along the time and/or frequency directions. The estimation in a sub-channel, i.e. frequency bin, with no pilot symbols is then obtained by interpolation. Common algorithms use Si or Cubic interpolators [2, 3]. The most known and optimal in the sense of Minimum Mean Squared Error (MMSE) interpolation approach is the Wiener filter solution [4, 5], which relies on the autocorrelation function of the channel over time and/or frequency. The quality of the channel statistical information together with the signal to noise ratio (SNR) and the spacing of frequency bins with pilot symbols are crucial to the quality of the interpolation process. Therefore, the choice of the frequency bins with pilots has to satisfy the *Sampling Theorem* which is not trivial due to the changing properties of the channel over time. A conservative approach to satisfying the *Sampling Theorem* is to guarantee a very high density of pilot sub-channels.

In Wiener interpolation, estimation of channels at pilot bins is only based on the pilot symbol and the received signals in pilot bins. Other available information, e.g. transmitted symbol alphabet and received signal at non-pilot bins are not used.

In this paper, we present methods for channel estimation which utilise not only the input-output data in bins with pilot symbols but also the received signals in the bins without pilot sequences. These methods are based on Expecta-

* Correspondence to: Dragan Obradovic, Siemens AG, Corporate Technology, Information and Communications Department, D-81730, Munich, Germany. E-mail: dragan.obradovic@siemens.com

Received 6 November 2006
Revised 9 May 2007
Accepted 31 July 2007

tion Maximisation (EM) and local interpolation, and they can be applied even if the *Sampling Theorem* is not satisfied. The benefit of these methods is that the resulting MIMO-OFDM system is robust to changes in the frequency selectivity of the channel and that it, therefore, spares channel capacity. We present these methods in an uncoded transmission system. Although embedding error correcting coding is straightforward, we will not present this part just for simplicity.

The usage of the above-mentioned extra information contained in the received signals in the frequency bins without pilot sequences will, naturally, be different depending whether or not the placing of pilot bins satisfies the *Sampling Theorem* requirements. In the case when the *Sampling Theorem* is satisfied, EM is used to improve the noisy channel estimates in frequency bins with pilot symbols, thus also enhancing the Wiener interpolation in other frequency bins, as shown in Subsection 3.1.

On the other hand, the standard Wiener interpolation based channel estimation method will have very bad results when the *Sampling Theorem* is not satisfied. In this case, EM can be used to perform semi-blind estimation of the channel in each particular frequency bin. The result of EM in this case is sensitive to the initialisation of the algorithm, as shown in Subsection 3.2. Different methods for appropriate initialisation of EM are discussed in the latter section too.

Related works on exploiting the received signals at frequency bins without pilot signals exist. Authors of Reference [6] have presented a framework of an iterative estimation scheme for coded systems in the frequency domain which is not based on the EM algorithm. The introduced method uses the a posteriori probabilities of the symbols after decoding to improve the channel estimate. Relying on the error correcting coding to assist the channel estimation puts additional constraints on the power of the coding scheme used, not necessarily leading to overall resource savings. Due to the coding, the method in Reference [6] can tolerate small deviations in pilot spacing from the case where the *Sampling Theorem* is satisfied. On the other hand, the herein presented EM-based method can deal with arbitrary large spacing between bins with pilot sequences which significantly exceed the spacing required by the *Sampling Theorem*.

Different EM based approaches are reported in References [7–9]. The difference between these approaches and our approach is that the channel is estimated in the time domain, whereas the frequency domain estimates are then obtained by simply making a Fourier transform. As discussed in Reference [10], the frequency spacing of OFDM sub-carriers imposes a finite resolution of the channel delay estimation, which leads to the non-negligible error effect called 'leakage'.

We organise the paper as follows: in Section 2, the system model of the transceiver and the wireless channel is presented. In Section 3, the two different EM based channel estimation methods are described in detail. The performance of these two methods is tested through simulation and the results are presented in Section 4. The mean square error (MSE) of the estimated channel and the bit error rate (BER) of the detected symbol, as a function of SNR are presented to evaluate the performance. Finally, Section 5 concludes the paper.

2. SYSTEM MODEL

We consider a MIMO-OFDM system as depicted in Figure 1.

This system has M_t transmit antennas with each antenna i sending a data stream $s_i(k)$ where k is the time index. At the receiver side, there are M_r receiver antennas receiving the noisy mixture of the transmitted streams. The signal received at antenna j is denoted as $x_j(k)$. The channel between the ith transmit antenna and the jth receiver antenna is h_{ij}. The use of OFDM transceiver is to convert the frequency selective channel into flat fading parallel sub-channels, with each channel transmitting a data stream $s_i^n(k)$, where n is the index of that sub-channel. The input–output vector notation of a flat fading sub-channel n can be written as

$$\mathbf{x}^n(k) = \mathbf{H}(jw^n)\mathbf{s}^n(k) + \eta(k) \tag{1}$$

where $\mathbf{H}(jw^n)$ is the Fourier transform of the channel mixing matrix in the nth sub-channel. The additive Gaussian

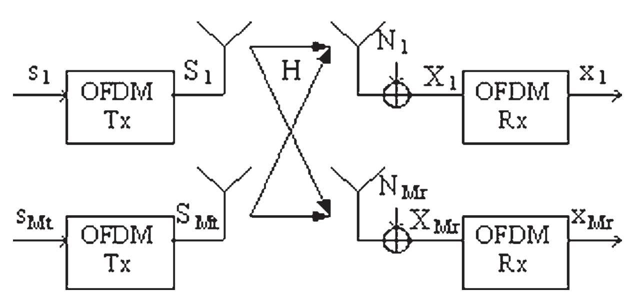

Figure 1. MIMO-OFDM system with M_t transmit and M_r receiver antennas.

DOI:10.1002/ett

noise is $\eta(k)$ and it is assumed to be white, zero mean, uncorrelated and with a known variance σ^2.

3. EM BASED CHANNEL ESTIMATION METHODS

In this section, we will introduce two novel EM based channel estimation methods.

The EM algorithm is a popular and powerful way of solving the likelihood maximisation problem when the gathered data are incomplete in some way. It is an iterative optimisation method with guaranteed convergence, although it may converge to a local optimum. A detailed discussion on EM algorithm can be found in Reference [11].

Before introducing the methods, we should clarify some basic assumptions for our system.

The transmitted symbols are assumed to be temporally white. As OFDM converts the serial transmitted symbols into parallel and transmits them through parallel orthogonal sub-channels; the temporal whiteness is translated into the independence between the symbols transmitted in different sub-channels.

Transmitted symbols are chosen from a fixed symbols set $\{S\} = \{s_1, s_2 \ldots s_M\}$, and the M-ary modulation scheme is known at the receiver side. It is also assumed that the probability of each symbol realisation is known at the receiver and, for simplicity, we assume all symbols to be equally probable. In addition, it is assumed that the data streams from different sources are statistically independent.

3.1. *Wiener interpolation enhancement*

In this sub-section we assume, that the placement of pilot symbols satisfies the *Sampling Theorem* requirement, i.e. at least two frequency bins with pilot symbols within the coherence bandwidth. Hence, Wiener interpolation can be used to estimate the channels in the rest of frequency bins without pilot symbols. The quality of the channel estimation in this case is directly proportional to the quality of channel estimates in frequency bins with pilot symbols, which are influenced by the channel noise.

In standard Wiener interpolation, we estimate the channel in pilot bins based only on the pilot symbol and the signals received in the pilot bins. If we pose this as ML problem, then we can define the log-likelihood as

$$L(\mathbf{h}_\mathrm{P}) = \log p(\mathbf{x}_\mathrm{P}; \mathbf{h}_\mathrm{P}) \tag{2}$$

where $\mathbf{h}_\mathrm{P}$ and $\mathbf{x}_\mathrm{P}$ are the channels and the received signals in the pilot bins. The ML solution is simply

$$\hat{\mathbf{h}}_\mathrm{P} = \left[\frac{x_\mathrm{P}^1}{s_\mathrm{P}^1}, \frac{x_\mathrm{P}^2}{s_\mathrm{P}^2}, \ldots \frac{x_\mathrm{P}^{N_\mathrm{P}}}{s_\mathrm{P}^{N_\mathrm{P}}}\right]^\mathrm{T} \tag{3}$$

where N_P is the number of pilot bins, ${s_\mathrm{P}}^{n_\mathrm{p}}$ and ${x_\mathrm{P}}^{n_\mathrm{p}}, n_\mathrm{p} = 1 \ldots N_\mathrm{P}$ are the pilot symbols and received signals, respectively.

Now, in order to improve the MSE performance, we would like to improve the estimation in pilot bins by exploiting the available information in non-pilot bins. For clarity reasons, we will first do it for the SISO-OFDM case and later extend the result to MIMO-OFDM.

3.1.1. *SISO-OFDM*

Here, we first consider the SISO-OFDM case. The transmission in each OFDM sub-channel can be modelled as

$$x_n = s_n h_n + \eta_n, \quad n = 1 \ldots N \tag{4}$$

where n denotes the frequency bin index, and x, h, and η are the received signal, channel coefficient and noise realisation, respectively.

In Wiener interpolation, the channel at non-pilot frequency bin can be expressed as a weighted linear combination of the channel at pilot bins, i.e.

$$h_n = \sum_{i=1}^{N_\mathrm{P}} w_n^i \cdot h_\mathrm{P}^i = \mathbf{w}_n^\mathrm{T} \mathbf{h}_\mathrm{P} \tag{5}$$

where $\mathbf{w}_n = [w_n^i, \ldots w_n^{N_\mathrm{P}}]^\mathrm{T}$ are the Wiener coefficients for interpolating the channel in the nth frequency bin. $\mathbf{h}_\mathrm{P} = [h_\mathrm{p}^1, \ldots h_\mathrm{p}^{N_\mathrm{P}}]^\mathrm{T}$ are the channels at pilot bins.

Having the correlation between the sub-channels, as shown in Equation (5), we can rewrite Equation (4) as

$$x_n = s_n h_n + \eta_n = s_n \mathbf{w}_n^\mathrm{T} \mathbf{h}_\mathrm{P} + \eta_n, \quad n = 1 \ldots N \tag{6}$$

In Equation (6), the parameters to be estimated are reduced to the channels in the pilot bins. Now we define the new log-likelihood as a function of channel in pilot bins:

$$L(\mathbf{h}_\mathrm{P}) = \log p(\mathbf{x}; \mathbf{h}) = \log p(\mathbf{x}; \mathbf{W}\mathbf{h}_\mathrm{P}) \tag{7}$$

where $\mathbf{W} = [\mathbf{w}_1 \ldots \mathbf{w}_N]^\mathrm{T} \in \mathbf{C}^{N \times N_\mathrm{P}}$ is the matrix containing the Wiener coefficients for all frequency bins. The difference between Equations (2) and (7) is that in Equation (2), we only use received signals in pilot bins, i.e. $\mathbf{x}_\mathrm{P}$, while in Equation (7), we use all received signals in all sub-channels, including non-pilot bins.

DOI:10.1002/ett

Now the channel estimation problem is posed as the following ML problem:

$$
\begin{aligned}
\hat{\mathbf{h}}_\mathrm{P} &= \arg\max_{\mathbf{h}_\mathrm{P}} L(\mathbf{h}_\mathrm{P}) = \arg\max_{\mathbf{h}_\mathrm{P}} \log p(\mathbf{x};\mathbf{h}) \\
&= \arg\max_{\mathbf{h}_\mathrm{P}} \log \sum_{s_1\in\{S\}} \cdots \sum_{s_N\in\{S\}} p(\mathbf{x},\mathbf{s};\mathbf{h}) \\
&= \arg\max_{\mathbf{h}_\mathrm{P}} \log \sum_{s_1\in\{S\}} \cdots \sum_{s_N\in\{S\}} p(\mathbf{x}|\mathbf{s};\mathbf{h}) \\
&= \arg\max_{\mathbf{h}_\mathrm{P}} \log \sum_{s_1\in\{S\}} \cdots \sum_{s_N\in\{S\}} \prod_{n=1}^{N} p(x_n|s_n;h_n) \\
&= \arg\max_{\mathbf{h}_\mathrm{P}} \log \prod_{n=1}^{N} \sum_{s_n\in\{S\}} p(x_n|s_n;h_n) \\
&= \arg\max_{\mathbf{h}_\mathrm{P}} \sum_{n=1}^{N} \log \sum_{s_n\in\{S\}} p(x_n|s_n;h_n) \\
&= \arg\max_{\mathbf{h}_\mathrm{P}} \sum_{n=1}^{N} \log \sum_{s_n\in\{S\}} p(x_n|s_n;\mathbf{w}_n^\mathrm{T}\mathbf{h}_\mathrm{P})
\end{aligned} \tag{8}
$$

The summation inside the logarithm makes it intractable to find an analytical solution to Equation (8). However, we can use the EM algorithm to solve this ML problem. Here, the hidden variables are the transmitted symbols in the non-pilot bins. We use Equation (3) as the initial guess $\hat{\mathbf{h}}_\mathrm{P}^0$. In each EM iteration, we have

E step: calculate $p(s_n|x_n,h_n) = p(s_n|x_n,\mathbf{w}_n^\mathrm{T}\hat{\mathbf{h}}_\mathrm{P}{}^{(t)})$

M step: maximize the lower bound:

$$
\begin{aligned}
\hat{\mathbf{h}}_\mathrm{P}^{(t+1)} &= \arg\max_{\mathbf{h}_\mathrm{P}^{(t+1)}} B\left(\mathbf{h}_\mathrm{P}^{(t+1)},\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right) \\
&= \arg\max_{\mathbf{h}_\mathrm{P}^{(t+1)}} \sum_{n=1}^{N} \sum_{s_n\in\{S\}} p\left(s_n|x_n,\mathbf{w}_n^\mathrm{T}\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right) \\
&\quad \times \log p\left(x_n|s_n,\mathbf{w}_n^\mathrm{T}\mathbf{h}_\mathrm{P}^{(t+1)}\right)
\end{aligned} \tag{9}
$$

Following the approach in Reference [12] it can be shown (see Appendix) that the solution of the maximisation in the iteration t is

$$
\hat{\mathbf{h}}_\mathrm{P}{}^{(t+1)} = \mathbf{R}_{ss(\mathbf{w})}^{(t)}{}^{-1}\mathbf{r}_{xs}^{(t)} \tag{10}
$$

with:

$$
\mathbf{R}_{ss(\mathbf{w})}^{(t)} = \sum_{n=1}^{N} \sum_{s_n\in\{S\}} p\left(s_n|x_n;\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right) s_n^* s_n \mathbf{w}_n^* \mathbf{w}_n^\mathrm{T} \tag{11}
$$

$$
\mathbf{r}_{xs}^{(t)} = \sum_{n=1}^{N} \sum_{s_n\in\{S\}} p\left(s_n|x_n;\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right) s_n^* x_n \mathbf{w}_n^* \tag{12}
$$

Here $p(s_n|x_n;\hat{\mathbf{h}}_\mathrm{P}{}^{(t)})$ is the posterior probability of the transmitted symbol s_n (from the alphabet $\{S\}$) given the received signal x_n at current estimate $\hat{\mathbf{h}}_\mathrm{P}{}^{(t)}$, which can be calculated in the following way:

$$
p\left(s_n|x_n;\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right) = \frac{p\left(x_n|s_n;\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right)\cdot p(s_n)}{\sum\limits_{s_n\in\{S\}} p\left(x_n|s_n;\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right)\cdot p(s_n)} \tag{13}
$$

If we use PSK modulation, i.e. the transmitted symbol has constant transmission power σ^2, then Equation (11) can be simplified as

$$
\begin{aligned}
\mathbf{R}_{ss(\mathbf{w})}^{(t)} &= \sigma_s^2 \sum_{n=1}^{N} \sum_{s_n\in\{S\}} p\left(s_n|x_n;\hat{\mathbf{h}}_\mathrm{P}^{(t)}\right) \mathbf{w}_n^* \mathbf{w}_n^\mathrm{T} \\
&= \sigma_s^2 \sum_{n=1}^{N} \mathbf{w}_n^* \mathbf{w}_n^\mathrm{T}
\end{aligned} \tag{14}
$$

From Equation (14), we can see that $\mathbf{R}_{ss(\mathbf{w})}$ only depends on the Wiener coefficients, which can be pre-computed and stored. Only $\mathbf{r}_{xs}^{(t)}$ should be updated in each iteration.

3.1.2. MIMO-OFDM

The extension of the above scheme to MIMO-OFDM systems is possible. In the MIMO system shown in Figure 1, we have M_t transmit antennas and M_r receiver antennas. By using OFDM, we divide the multipath channel between each antenna pair into N orthogonal sub-channels. Then, for each sub-channel, their $M_r \times M_t$ channel coefficients can be written in a matrix $\mathbf{H}$:

$$
\mathbf{H}_n = \begin{bmatrix} h_{(1,1),n} & \cdots & h_{(1,M_t),n} \\ \vdots & \ddots & \vdots \\ h_{(M_r 1),n} & \cdots & h_{(M_r M_t),n} \end{bmatrix} \tag{15}
$$

Here n is the index denoting the nth sub-channel. For lack of knowledge about the spatial correlations, each antenna pair can be regarded as a SISO-OFDM system. Wiener interpolation for the channel between antenna pair (m_t, m_r) can be expressed as

$$
h_{(m_t,m_r),n} = \sum_{i=1}^{N_\mathrm{P}} w_{(m_t,m_r),n}^i h_{(m_t,m_r),\mathrm{P}}^i = \mathbf{w}_{(m_t,m_r),n}^\mathrm{T}\mathbf{h}_\mathrm{P} \tag{16}
$$

A common assumption is that for close together antennas, all the antenna pairs have the same channel statistics.

Hence, we use the same Wiener coefficients $\mathbf{w}_n$ for every antenna pair. Then we can write

$$\mathbf{H}_n = \begin{bmatrix} h_{(1,1),n} & \cdots & h_{(1,M_t),n} \\ \vdots & \ddots & \vdots \\ h_{(M_r,1),n} & \cdots & h_{(M_r,M_t),n} \end{bmatrix} = \sum_{i=1}^{N_{\mathrm{P}}} w_n^i \begin{bmatrix} h^i_{(1,1),\mathrm{i}} & \cdots & h^i_{(1,M_t),\mathrm{i}} \\ \vdots & \ddots & \vdots \\ h^i_{(M_r,1),\mathrm{i}} & \cdots & h^i_{(M_r,M_t),\mathrm{i}} \end{bmatrix} \tag{17}$$

The input–output relation at m_rth receiving antenna can be written as

$$\begin{aligned} x_{m_r,n} &= \left[h_{(m_r,1),n}, \ldots h_{(m_r,M_t),n}\right] \begin{bmatrix} s_{1,n} \\ \vdots \\ s_{M_t,n} \end{bmatrix} + \eta_{m_r,n} \\ &= \sum_{i=1}^{N_{\mathrm{P}}} w_n^i \left[h^i_{(m_r,1),\mathrm{P}}, \ldots h^i_{(m_r,M_t),\mathrm{P}}\right] \begin{bmatrix} s_{1,n} \\ \vdots \\ s_{M_t,n} \end{bmatrix} + \eta_{m_r,n} \\ &= \left[h^1_{(m_r,1),\mathrm{P}} \ldots h^1_{(m_r,M_t),\mathrm{P}}, \ldots, h^{N_{\mathrm{P}}}_{(m_r,1),\mathrm{P}}, \ldots h^{N_{\mathrm{P}}}_{(m_r,M_t),\mathrm{P}}\right] \\ &\quad \times \begin{bmatrix} w_n^1 \mathbf{I}_{M_t\times M_t} \\ \vdots \\ w_n^{N_{\mathrm{P}}} \mathbf{I}_{M_t\times M_t} \end{bmatrix} \begin{bmatrix} s_{1,n} \\ \vdots \\ s_{M_t,n} \end{bmatrix} + \eta_{m_r,n} \end{aligned} \tag{18}$$

Denote

$$\begin{aligned} \mathbf{h}_{m_r,\mathrm{P}} &= \left[h^1_{(m_r,1),\mathrm{P}} \ldots h^1_{(m_r,M_t),\mathrm{P}}, \ldots, h^{N_{\mathrm{P}}}_{(m_r,1),\mathrm{P}}, \ldots h^{N_{\mathrm{P}}}_{(m_r,M_t),\mathrm{P}}\right]^{\mathrm{T}} \\ \mathbf{W}_n &= \left[w_n^1 \mathbf{I}_{M_t\times M_t} \ldots w_n^{N_{\mathrm{P}}} \mathbf{I}_{M_t\times M_t}\right]^{\mathrm{T}} \\ \mathbf{s}_n &= \left[s_{1,n} \ldots s_{M_t,n}\right]^{\mathrm{T}} \end{aligned}$$

Then Equation (18) can be represented as

$$\begin{aligned} x_{m_r,n} &= \mathbf{h}^{\mathrm{T}}_{m_r,\mathrm{P}} \mathbf{W}_n \mathbf{s}_n + \eta_{m_r,n} \\ &= \mathbf{s}_n^{\mathrm{T}} \mathbf{W}_n^{\mathrm{T}} \mathbf{h}_{m_r,\mathrm{P}} + \eta_{m_r,n} \end{aligned} \tag{19}$$

Equation (19) has a similar form as Equation (6). The definition of likelihood and the derivation of EM algorithm can just follow Equations (7) to (9). Here we only present the result

$$\hat{\mathbf{h}}^{(t+1)}_{m_r,\mathrm{P}} = \mathbf{r}^{(t)}_{x\mathbf{s}} {\mathbf{R}^{(t)}_{\mathbf{ss}(\mathbf{W})}}^{-1} \tag{20}$$

with

$$\begin{aligned} \mathbf{R}^{(t)}_{\mathbf{ss}(\mathbf{W})} &= \sum_{n=1}^{N} \sum_{\mathbf{s}_n \in \{\mathbf{S}\}} p\left(\mathbf{s}_n | x_{m_r,n}; \hat{\mathbf{h}}^{(t)}_{m_r,\mathrm{P}}\right) \mathbf{W}_n^* \mathbf{s}_n^* (\mathbf{W}_n \mathbf{s}_n)^{\mathrm{T}} \\ \mathbf{r}^{(t)}_{x\mathbf{s}(\mathbf{W})} &= \sum_{n=1}^{N} \sum_{\mathbf{s}_n \in \{\mathbf{S}\}} p\left(\mathbf{s}_n | x_{m_r,n}; \hat{\mathbf{h}}^{(t)}_{m_r,\mathrm{P}}\right) x_{m_r,n} \mathbf{W}_n^* \mathbf{s}_n^* \end{aligned} \tag{21}$$

3.2. Sequential EM channel estimation with propagation of the solution

This section investigates the case when the *Sampling Theorem* requirement is not satisfied, i.e. where the spacing of the frequency bins with pilot sequences are large. This case arises, when the characteristics of the system have changed making the channel more frequency selective than before. Hence, the old spacing of the pilot bins is not sufficient to enable the use of the standard Wiener interpolation approach to channel estimation.

In this case we propose a novel approach to channel estimation where the channel estimation in a bin without pilot symbols is solved as a semi-blind iterative problem and the local correlation between adjacent bins is used for the algorithm initialisation.

Recalling the channel model in Equation (1), we will focus on estimating the spatial channel in the nth frequency bin and equalizing the channel. For a simpler notation, we will omit the index 'n' in equations in the remaining part of this section.

In general, the mixing matrix $\mathbf{H}$ is unknown and has to be estimated. If the transmitted streams $s_i(k)$ were known at the receiver, the channel could be estimated by solving the following Maximum Likelihood problem:

$$\arg\max_{\mathbf{H}} \log p(\mathbf{x}|\mathbf{s}; \mathbf{H}) \tag{22}$$

where the probability density function $p(\mathbf{x}|\mathbf{s}; \mathbf{H})$ is given by the noise statistics, a Gaussian vector random variable with zero mean and known covariance matrix $\sigma^2 \cdot \mathbf{I}$ (where $\mathbf{I}$ stands for a unit matrix).

The problem treated in our approach is the case where the transmitted streams $s_i(k)$ are NOT known at the receiver. Therefore, we deal with the so-called 'semi-blind' channel estimation.

Due to the fact that the transmitted streams are not known at the receiver, the channel estimation has to be formulated differently than in Equation (22). The Maximum

DOI:10.1002/ett

Likelihood problem we can formulate in this case is as follows:

$$\arg\max_{\mathbf{H}} \log p(\mathbf{x}; \mathbf{H}) \tag{23}$$

We can solve this ML problem similarly as what we did with Equation (7). The two steps in the tth EM iteration are *E-step*: where the expected value of the conditional log-likelihood is computed as

$$\sum_{\mathbf{s}\in\{\mathbf{S}\}} p\left(\mathbf{s}|\mathbf{x}; \mathbf{H}^{(t)}\right) \cdot \log p(\mathbf{x}|\mathbf{s}; \mathbf{H}) \tag{24}$$

M-step: where Equation (24) is maximized with respect to $\mathbf{H}$. The resulting $\mathbf{H}$ is then denoted as $\mathbf{H}^{(t+1)}$.

This is a semi-blind method since the transmitted streams were not known at the receiver. The problem is nevertheless feasible since the prior knowledge about the symbol alphabet and about the statistical independence of streams suffices for estimation of $\mathbf{H}$. The price we have to pay is the ambiguity of the solution with respect to permutation and phase scaling which are the intrinsic properties of all blind methods. The ambiguity is resolved if the initial channel estimate in the EM algorithm is sufficiently close to the correct solution. The following sections present two methods of exploiting the local correlations between the adjacent bins for the initialisation of the EM algorithm.

3.2.1. Direct EM initialisation with the models of neighbouring bins

If we look at the frequency correlation of the OFDM channel, i.e. the autocorrelation of OFDM sub-channels, we can see that the neighbouring frequency bins are strongly correlated. Using this property, we can initialise the EM channel estimation by using the estimate from the neighbouring bin. Pilot symbols are sent only in a limited number of sub-channels, as they only provide initialisation for the EM channel estimation in the consecutive bins. Figure 2 illustrates the sequential estimation process.

We have N OFDM sub-channels. Bin 1 and bin N are pilot bins where pilot symbols are sent. Channel is estimated in pilot bins by solving the ML problem as formulated in Equation (22). The results in bin 1 and bin N can be used respectively to initialise the EM channel estimation in bin 2 and $N-1$. In the following step, in estimating the channel in frequency bin n, we use the EM algorithm initialised by the result in bin $n-1(n+1$ if backward).

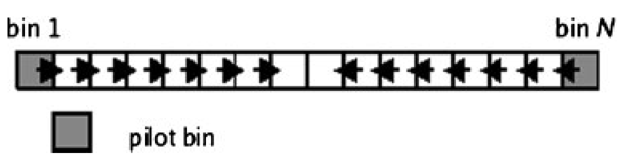

Figure 2. Sequential EM channel estimation with the propagation of solution.

Using sequential initialisation generates error propagation problem as we can see from the simulation result in Section 4. In some instances, neighbouring bin provides a bad initialisation, which may lead EM to converge to a wrong solution. This error will propagate to the estimation in the following bins. In the next sub-section, we will introduce a local Wiener interpolation method, which can improve the initialisation.

3.2.2. EM initialised using local Wiener interpolator

In order to improve the initialisation of EM, we introduce a local Wiener interpolator. We use the previous estimated bins $n-1, n-2, \ldots n-L$ as pseudo pilot bins. The initialisation of current bin n is obtained by interpolating the pseudo pilot bins. In the MMSE sense, the interpolation is given by

$$\mathbf{H}_{ini}^{n} = \sum_{l=1}^{L} w_l \hat{\mathbf{H}}^{n-l} \tag{25}$$

The advantage of doing this is that by a local Wiener interpolation, we exploit the correlation between the neighbouring bins to generate an initialisation which is closer to the true solution than initialisation with the previous bin. This is illustrated in Figure 3.

4. SIMULATION RESULTS

To verify the validity of these EM based channel estimation methods, we test them through simulations. Here we use a realistic frequency selective MIMO channel developed by the 3GPP-SCM group [13]. The main parameters

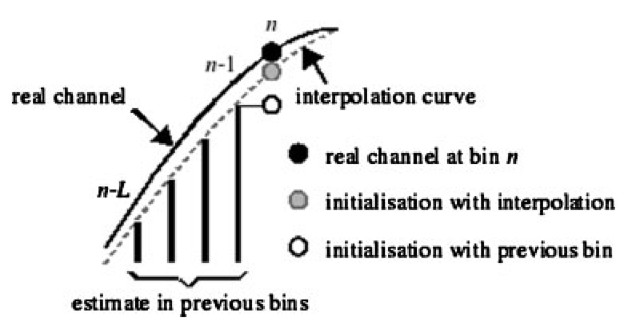

Figure 3. Initialisation with Wiener interpolation.

DOI:10.1002/ett

Table 1. Simulation settings.

Environment	Microcellular
Antenna configuration	2×2, spacing of 4λ
Max delay	1.2 μs
Bandwidth	20 MHz
Number of OFDM bins	256
Number of data streams	2
Pilot positions for Wiener enhancement	1, 6, 11, ...
Pilot positions for sequential EM	1, 32, 64, 96, ...

of the channel and the system settings are summarized in Table 1. The coherence bandwidth is approximately 1/ 1.2 μs = 833 Hz. According to the *Sampling Theorem*, successful interpolation methods require the spacing of pilot bins to be not more than 833/2 = 417 Hz, i.e. 5 bins.

4.1. EM based Wiener interpolation enhancement vs. standard Wiener

In this part we will test the EM based Wiener interpolation enhancement method and compare it with standard Wiener interpolation. The placement of pilot sequences satisfies the *Sampling Theorem*. The simulation results are displayed in Figures 4 and 5.
From the results we can see, by using EM to exploit extra available information, we improve the channel estimation and the symbol detection.

4.2. Sequential EM based channel estimation method

Here we test the validity of sequential EM channel estimations where the *Sampling Theorem* is not satisfied. In the

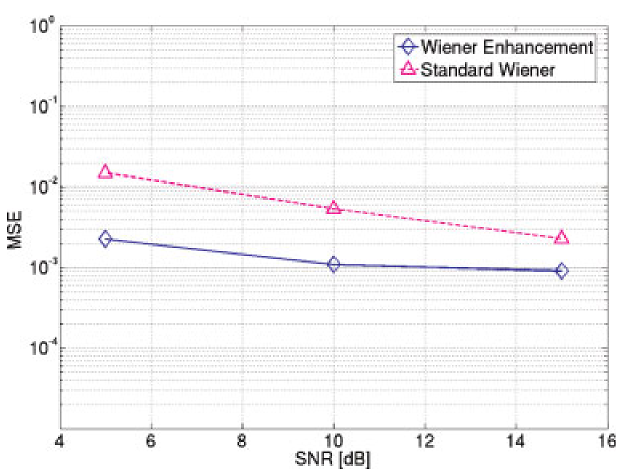

Figure 4. MSE performance of EM based Wiener interpolation enhancement.

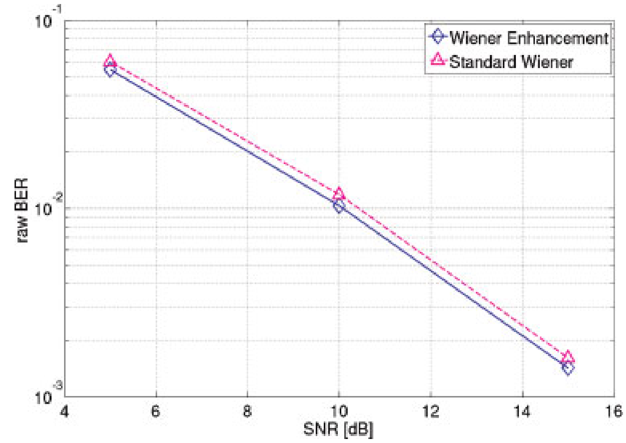

Figure 5. BER performance of EM based Wiener interpolation enhancement.

first simulation, we directly use the estimate from the neighbouring bin for the initialisation of EM. The BER performance is shown in Figure 6.

We can see that the EM channel estimation with direct solution propagation does not work. The neighbouring bin $n-1$ does not always provide a good initialisation for the EM in estimating the current bin n. An improper initialisation leads EM to converge to a wrong solution. If we use this solution to initialise the EM in bin $n+1$, then error propagates. To solve this problem, we use local Wiener interpolation to improve the initialisation, as described in subsection 3.2.2.

Figure 7 shows the BER performance of the sequential EM method with local Wiener interpolator.

With the help of local Wiener interpolator, we could find an initial channel estimate close to the correct solution,

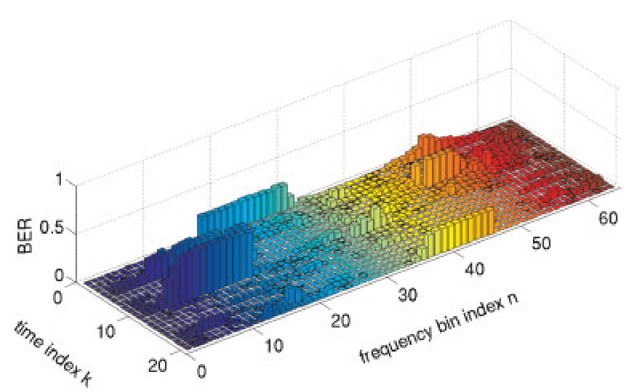

Figure 6. BER in each frequency bin, sequential EM channel estimation in 2×2 MIMO OFDM system.

DOI:10.1002/ett

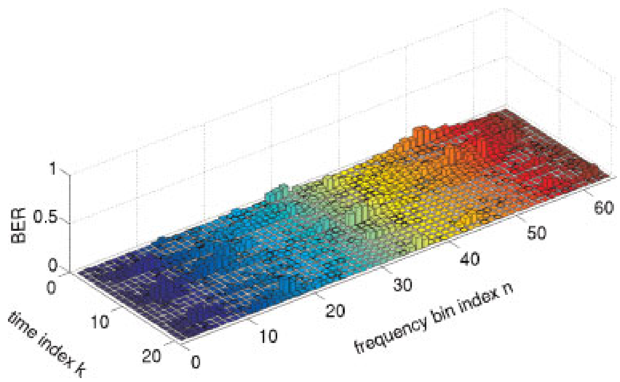

Figure 7. BER in each frequency bin, sequential EM channel estimation with local Wiener interpolator.

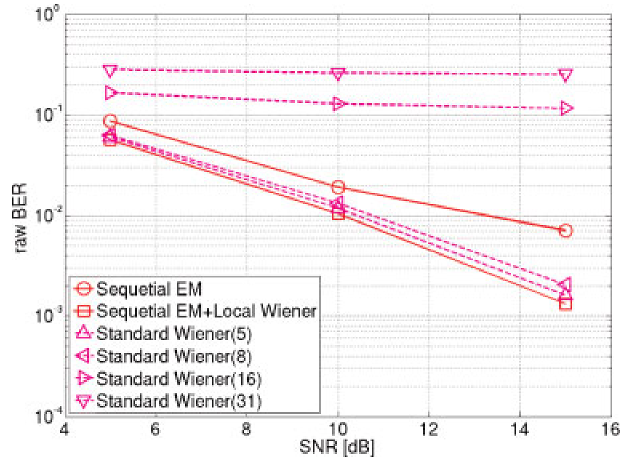

Figure 9. BER performance of sequential EM based channel estimation.

which avoids EM converging to a wrong solution. Hence, we observe much lower error propagation in Figure 7.

To further investigate the properties of the above-mentioned sequential EM based methods, we compared them with the standard Wiener interpolation. In order to illustrate how the channel estimation and BER degrade with the reduction of the rate of occurrence of frequency bins with pilots, we will increase the distance between neighbouring frequency bins from 5 where the *Sampling Theorem* was satisfied to 8, 16 and 31 where it is not satisfied and nevertheless apply the standard Wiener interpolation. The MSE and BER performance of different methods in 2×2 MIMO OFDM systems are displayed in Figures 8 and 9.

The simulation results show that the sequential EM based channel estimation methods perform well even with very sparse pilot placement, where standard Wiener interpolation fails. Sequential EM estimation initialised by neighbouring bins generates error propagation problem. By introducing a local Wiener interpolator for the initialisation, this problem can be solved. The presented methods open the possibility of capacity improvement in OFDM/MIMO-OFDM systems by reducing the signalling overhead needed for pilots and channel estimation.

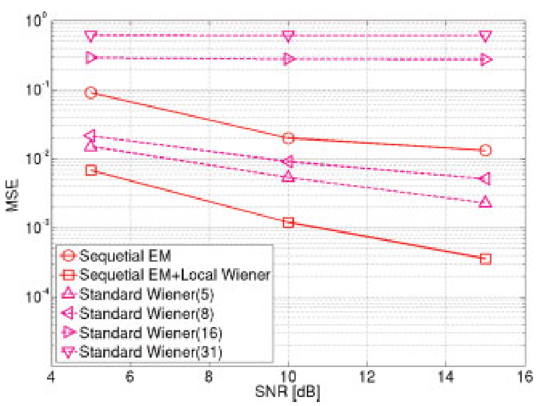

Figure 8. MSE performance of sequential EM based channel estimation.

5. CONCLUSIONS

In standard Wiener interpolation, only the transmitted symbols and received signals at pilot bins are used for the channel estimation. By including the information at non-pilot bins, the estimation performance is largely improved. Knowledge of the transmitted symbol alphabet and the received signal at non-pilot bins is exploited with a maximum likelihood approach. Analytical solution could not be directly found. Suitable expression has been derived to apply the EM algorithm

In this paper, we have presented two EM based channel estimation methods. The Wiener enhancement method is based on standard Wiener interpolation. It requires the *Sampling Theorem* requirement satisfied. When the *Sampling Theorem* is not satisfied, then a sequential method is applied, where the channel is estimated bin by bin. The result from the neighbouring bin is used to initialise the EM in the current bin. A further improvement is the design of a local Wiener interpolation, which provides a better initialisation, reducing the risk of converging to a permuted constellation which would lead to error propagation. The method is novel considering the case where the *Sampling Theorem* is not satisfied.

APPENDIX

In this appendix we derive the solution for Equation (14). We start from Equation (9) and compute:

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{P}}{}^{(t+1)} &= \arg\max_{\mathbf{h}_{\mathrm{P}}^{t+1}} B\left(\mathbf{h}_{\mathrm{P}}{}^{(t+1)}, \hat{\mathbf{h}}_{\mathrm{P}}{}^{(t)}\right) \\ &= \arg\max_{\mathbf{h}_{\mathrm{P}}^{t+1}} \sum_{n=1}^{N} \sum_{s_n \in \{S\}} p\left(s_n | x_n, \mathbf{w}_n^{\mathrm{H}} \hat{\mathbf{h}}_{\mathrm{P}}{}^{(t)}\right) \log\, p\left(x_n | s_n, \mathbf{w}_n^{\mathrm{H}} \mathbf{h}_{\mathrm{P}}{}^{(t+1)}\right) \\ &= \arg\max_{\mathbf{h}_{\mathrm{P}}^{t+1}} \sum_{n=1}^{N} \sum_{s_n \in \{S\}} p\left(s_n | x_n, \mathbf{w}_n^{\mathrm{T}} \hat{\mathbf{h}}_{\mathrm{P}}{}^{(t)}\right) (x_n - \mathbf{w}_n^{\mathrm{T}} \mathbf{h}_{\mathrm{P}}{}^{(t+1)} s_n)^{\mathrm{H}} \\ &\quad \times \left(x_n - \mathbf{w}_n^{\mathrm{T}} \mathbf{h}_{\mathrm{P}}{}^{(t+1)} s_n\right)\end{aligned}$$

Taking the derivative with respect to $(\mathbf{h}_{\mathrm{P}}{}^{(t+1)})^{\mathrm{H}}$ and setting it to 0:

$$\sum_{n=1}^{N} \sum_{s_n \in \{S\}} p\left(s_n | x_n, \mathbf{w}_n^{\mathrm{T}} \hat{\mathbf{h}}_{\mathrm{P}}{}^{(t)}\right) \left(s_n \mathbf{w}_n^{\mathrm{T}}\right)^{\mathrm{H}} \times \left(x_n - s_n \mathbf{w}_n^{\mathrm{T}} \mathbf{h}_{\mathrm{P}}{}^{(t+1)}\right) = 0$$

Solving this equation, we obtain

$$\left(\sum_{n=1}^{N} \sum_{s_n \in \{S\}} p\left(s_n | x_n, \mathbf{w}_n^{\mathrm{T}} \hat{\mathbf{h}}_{\mathrm{P}}{}^{(t)}\right) s_n^* s_n \mathbf{w}_n^* \mathbf{w}_n^{\mathrm{T}}\right) \mathbf{h}_{\mathrm{P}}{}^{(t+1)} = \sum_{n=1}^{N} \sum_{s_n \in \{S\}} p\left(s_n | x_n, \mathbf{w}_n^{\mathrm{T}} \hat{\mathbf{h}}_{\mathrm{P}}{}^{(t)}\right) s_n^* x_n \mathbf{w}_n^*$$

which gives us the results in Equations (15), (16) and (17).

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Ruxandra Lupas Scheiterer for the help in reviewing the paper and for the useful comments.

REFERENCES

1. Bingham JAC. Multicarrier modulation for data transmission: an idea whose time has come. *IEEE Communications Magazine* 1990; **28**(5):4–14.
2. Rohling H, Grünheid H, Galda D. OFDM transmission technique for the 4th generation of mobile communication systems. In *Proceedings of International Symposium on Telecommunications (IST)*, 2003.
3. Auer G. Channel estimation in two dimensions for OFDM systems with multiple transmit antennas. In *Proceedings of IEEE Global Tele-communications Conference (GLOBECOM)*, 2003.
4. Hoeher P, Kaiser S, Robertson P. Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering. In *Proceedings of IEEE ICASSP'97*, 1997; 1845–1848.
5. Kaiser S. Multi-carrier CDMA mobile radio systems — analysis and optimization of detection, decoding, and channel estimation. *Fortschritts-Berichte VDI, Reihe IO, Nr.531*, VDI-Verlag, Dusseldorf, 1998.
6. Sanzi F, Jelting S, Speidel J. A comparative study of iterative channel estimators for mobile OFDM systems. *IEEE Transactions on Wireless Communications* 2003; **2**(5):849–859.
7. Ma X, Kobayashi H, Schwartz SC. EM-based channel estimation algorithms for OFDM. *EURASIP Journal on Applied Signal Processing* 2004; **10**:1460–1477.
8. Xie Y, Georghiades CN. Two EM-type channel estimation algorithms for OFDM with transmitter diversity. *IEEE Transactions on Communications* 2003; **51**(1):106–115.
9. Sgraja C, Lindner J. EM-based decision-directed channel estimation for OFDM: a new interpretation. In: *Proceedings of 10th International OFDM-Workshop 2005 (InOWo'05)*, 2005.
10. van de Beek J, Edfors O, Sandell M, Wilson S, Beorjesson P. On channel estimation in OFDM systems. In *Proceedings of 45th IEEE Vehicular Technology Conference (VTC)*, 1995; 815–819.
11. Dellaert F. *The Expectation Maximization Algorithm.* Technical Report. GIT. 2002.
12. Belouchrani A, Cardoso J. Maximum likelihood source separation for discrete sources. In *Proceedings of EUSIPCO94*, 1994.
13. *3GPP TSG RAN Spatial Channel Model for MIMO simulations (Release 6)*, TR 25.996, v6.1.0.

AUTHORS' BIOGRAPHIES

Chongning Na received the B.E. degree in Mechanical Engineering from Tsinghua University, Beijing, China, in 2002, and the M.S. degree (with an academic award) in Electrical Engineering from Technical University of Munich, Germany, in 2005. He is now pursuing the Ph.D. degree in Electrical Engineering from Technical University of Munich. His Ph.D. work is funded by Siemens AG, Corporate Technology department. He was with Microsoft Research (MSR) Asia, Beijing, during the fall in 2002 and with Siemens AG, Communications department from summer 2004 to spring 2005, and with Siemens AG, Corporate Technology department from spring 2005 to present. His research interests are in the area of information theory, probabilistic inference, machine learning and their application in telecommunication, control system and signal processing.

Dragan Obradovic received the B.E. and M.E. degrees in Mechanical Engineering in 1980 and 1985, respectively, from the University of Belgrade, Yugoslavia. In 1990, he received the Ph.D. degree in Mechanical Engineering from the Massachusetts Institute of Technology, Cambridge, USA. Since 1992, Dr. Obradovic has been with the Corporate Technology of Siemens in Munich, Germany where he currently holds a position of Principal Research Scientist in the Information and Communications department. His research involves, among others, the development and application of machine learning methods for system control and signal processing. He has published several dozens of journal and international conference papers and the book *'An Information-Theoretic Approach to Neural Computing'* (New York, Springer, 1996) and is author of several granted and pending patents. Dr. Obradovic also holds an Industrial Lecturer position at Munich University of Technology, where he has been teaching since 2001 in the international Masters of Science in Communication Engineering programme within the EE Department.

Andrei Szabo received his Dipl.Ing. degree from the University 'Politehnica' Bucharest in 2000. He is currently with the Corporate Technology of Siemens AG in Munich, in the Information and Communications department, while in parallel, pursuing the Ph.D. degree at the Technical University of Munich. His research interests are learning systems with application to telecommunication, multiple antenna systems and localisation applications.