An Information Theoretic Foundation of Synchronized Detection

Jens Baltersee, Gunnar Fock, and Heinrich Meyr, *Fellow, IEEE*

***Abstract*—The constrained capacity of a coherent coded modulation (CM) digital communication system with data-aided channel estimation and a discrete, equiprobable symbol alphabet is derived under the assumption that the system operates on a flat fading channel and uses an interleaver to combat the bursty nature of the channel. It is shown that linear minimum mean square error channel estimation directly follows from the derivation and links average mutual information to the channel dynamics. Based on the assumption that known training symbols are transmitted, the achievable rate of the system is optimized with respect to the amount of training information needed. Furthermore, the results are compared to the additive white Gaussian noise channel, and the case when ideal channel state information is available at the receiver.**

***Index Terms*—Channel capacity, fading channels, interleaved coding, mutual information, synchronized detection.**

I. Introduction

DUE TO complexity constraints, virtually all of today's digital wireless communication systems follow the principle of synchronized detection for which a channel estimate is formed and subsequently used for detection as if it were the true known channel. Here it is assumed that known training symbols are transmitted to the receiver. Usually, the practicing engineer estimates the performance of such a communication system by assuming that the channel estimation error translates directly into a simple signal-to-noise-ratio (SNR) loss. While this approach has proven very useful when designing communication systems, the interesting question of how the principle of synchronized detection affects the channel capacity is still largely unanswered, although similar problems have been treated in the literature; a good overview of the area is given in [13]. The work in [10], for example, analyzes the capacity of a system with a fixed, modified nearest-neighbor decoding rule with respect to errors in the estimation of the channel fading process. Here, since we make no assumption on the decoding rule, optimal decoding is implied. Similar problems are also treated in [11] and [12]. For a discussion, please refer to the remarks at the end of the paper.

In this paper, the constrained capacity is computed for a typical coded modulation (CM) transmission system operating on flat fading channels and using an interleaver to combat the bursty nature of the channel. The term constrained capacity or achievable rate is used throughout the paper in order to indicate that the capacity results are constrained to a discrete, equiprobable symbol alphabet. It is shown that the average mutual information is a function of the channel dynamics through linear minimum mean square error (LMMSE) channel estimation which follows directly from the derivation. The interleaver is an integral part of the system, since most well-known codes have been devised to combat statistically independent channel realizations. Virtually all of today's communication systems are based on that assumption, and therefore this type of channel is of greatest practical interest. The resulting average mutual information is then used to optimize, via Monte Carlo simulation, the achievable rate with respect to the amount of training information needed for a given scenario in terms of channel dynamics and SNR. Furthermore, the results are compared to the additive white Gaussian noise (AWGN) channel, and the case when ideal channel state information (CSI) is available at the receiver. Finally, our approach is compared to the recently published work by Marzetta and Hochwald [8].

Paper approved by R. Raheli, the Editor for Detection, Equalization and Coding of the IEEE Communications Society. Manuscript received June 27, 2000; revised January 9, 2001, and June 1, 2001.

The authors are with the Institute for Integrated Systems in Signal Processing (ISS), Aachen University of Technology (RWTH), 52056 Aachen, Germany (e-mail: balterse@ert.rwth-aachen.de; fock@ert.rwth-aachen.de; meyr@ert.rwth-aachen.de).

Publisher Item Identifier S 0090-6778(01)10613-6.

II. Channel Model

A CM transmission system operating on a flat-fading channel and featuring data-aided (DA) synchronized detection is shown in Fig. 1. The flat fading channel is characterized by a multiplicative fading process $\{c'_k\}$ and AWGN. For synchronized detection, the inner receiver performs DA channel estimation and delivers the estimated channel coefficients $\{\hat{c}'_k\}$ and the received symbols to the outer receiver. The outer transmission system comprises channel coding, modulation (symbol mapping), interleaving/deinterleaving, and decoding. The interleaving/deinterleaving is employed to transform the bursty channel into an independently distributed channel. This is necessary, since most well-known codes have been devised to combat statistically independent channel realizations.

For the interleaved and encoded data symbols $\mathbf{a}'$, the following transmission model results:

$$z'_k = a'_k c'_k + m'_k \tag{1}$$

where symbol power $E\{a'^{*}_k a'_k\} = \sigma_a^2$ is assumed and the noise power is given by $E\{m'^{*}_k m'_k\} = N_0$. The fading process $\{c'_k\}$ is modeled as a stationary, zero mean, circularly symmetric complex Gaussian process of variance σ_c^2 and Doppler spectrum $S(e^{j\omega})$. Thus the average SNR per symbol is $\bar{\gamma}_s = (\sigma_a^2\sigma_c^2)/N_0$.

0090–6778/01$10.00 © 2001 IEEE

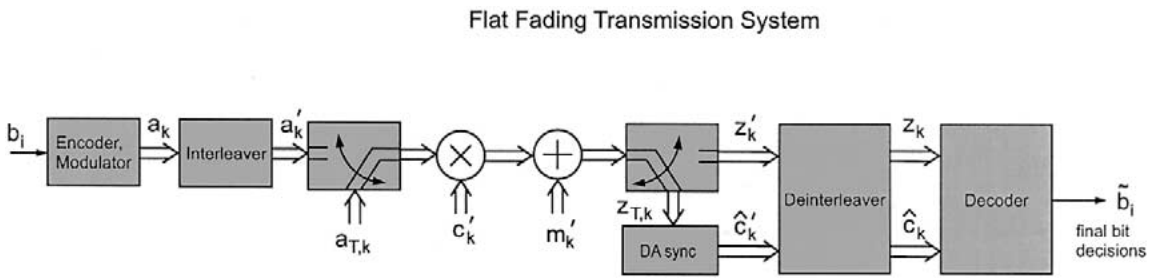

Fig. 1. Transmission system with synchronized detection.

For purposes of DA channel estimation, it is assumed that N_T known training symbols

$$\mathbf{a}_T = \begin{pmatrix} a_{T;k_0} & a_{T;k_1} & \cdots & a_{T;k_{N_T-1}} \end{pmatrix} \tag{2}$$

are introduced at positions $k = k_0, k_1, \ldots k_{N_T-1}$, within the length-$N$ interleaved message $\mathbf{a}'$. Thus, the transmission model for the received training symbols can be written as

$$z_{T;k} = a_{T;k} c'_k + m'_k. \tag{3}$$

Based on the received training symbol, the DA channel estimation produces estimates $\{\hat{c}'_k\}$ of the fading process $\{c'_k\}$. Both the channel estimates and the received data symbols are passed to the de-interleaver which maps $\{z'_k\} \rightarrow \{z_k\}$ and $\{\hat{c}'_k\} \rightarrow \{\hat{c}_k\}$. If we assume an ideal interleaving/de-interleaving operation with a sufficiently deep memory then, the mapping of $\{z'_k\} \rightarrow \{z_k\}$ results in the following channel model:

$$z_k = a_k c_k + m_k \tag{4}$$

where $\{c_k\}$ is an i.i.d. fading process.

Note that, for ease of explanation, the sets $\{\mathbf{a}, \mathbf{a}_T\}$ and $\{\mathbf{z}, \mathbf{z}_T\}$ were introduced as time-multiplexed sequences. However, the model encompasses code- and frequency-multiplexed transmission as well. We only require symbols in set $\mathbf{a}$ to be transmitted on different time, frequency, or codeword coordinates than the symbols in set $\mathbf{a}_T$, and that the sets $\mathbf{z}$ and $\mathbf{z}_T$ are separable at the receiver. In the case of DS-CDMA systems, this can be accomplished by despreading the received signal with the pseudorandom sequence that is assigned to the pilot channel. For OFDM systems, the pilot tones can be separated using the orthogonality of the subcarriers. Furthermore, we would like to point out that coded modulation is assumed throughout the paper, but that the results obtained using this approach also give a very good indication of the achievable rate of bit-interleaved coded modulation (BICM) schemes, a concept which underlies many practical systems. In [7], it is shown that the achievable rate for BICM over known flat fading channels approaches that of CM very closely, if the modulator employs Gray mapping. The constrained capacity for CM is derived in the following Section.

III. Derivation of the Constrained Capacity

In this section, the average mutual information for the aforementioned channel model is derived. Since the transmission model assumes the usage of pilot symbols in order to estimate the channel, the channel is said to be partially known to the receiver. We denote this channel P-CSI. In contrast, if complete (perfect) channel knowledge is available, we choose to use C-CSI. The P-CSI channel described in the previous section, with input $\{\mathbf{a}\}$, output $\{\mathbf{z}, \mathbf{z}_T\}$, and known parameter $\{\mathbf{a}_T\}$ is completely characterized by the joint channel transition pdf $p(\mathbf{z}, \mathbf{z}_T \,|\, \mathbf{a}, \mathbf{a}_T)$. Hence, for a given, discrete symbol constellation, and under the constraint of equiprobable symbols, the capacity of this channel is given by the average mutual information [15]. We will denote this constrained capacity with C^* as compared to C which is reserved for the true channel capacity that requires the maximization of the average mutual information over the input symbol distribution. Mutual information is measured in bits per second per hertz. If we consider a block of $N = N_D + N_T$ transmitted symbol vectors, of which N_D are usable data symbols vectors, the expression for the average mutual information for this block per channel use is given by

$$\frac{1}{N} \mathcal{I}(\mathbf{z}, \mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T). \tag{5}$$

Compare this to the average mutual information per channel use of the C-CSI channel which is given by [9]

$$\frac{1}{N} \mathcal{I}(\mathbf{z}; \mathbf{a} \,|\, \mathbf{c}) \tag{6}$$

where $\mathbf{c} = (c_0, \ldots, c_{N-1})^T$. Due to training symbol insertion, the achievable rate of the P-CSI channel will be a factor N_D/N lower than the achievable rate of the C-CSI channel. In addition to that, noisy channel estimates will further reduce the achievable rate of the P-CSI channel.

Now, according to the chain rule for mutual information we can rewrite $\mathcal{I}(\mathbf{z}, \mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T)$ as follows:

$$\mathcal{I}(\mathbf{z}, \mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T) = \mathcal{I}(\mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T) + \mathcal{I}(\mathbf{z}; \mathbf{a} \,|\, \mathbf{a}_T, \mathbf{z}_T) \tag{7}$$

$$= \mathcal{I}(\mathbf{z}; \mathbf{a} \,|\, \mathbf{a}_T, \mathbf{z}_T) \tag{8}$$

where $\mathcal{I}(\mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T) = 0$, since $\mathbf{z}_T$ conditioned on $\mathbf{a}_T$ does not convey any information about $\mathbf{a}$. Having that in mind, it is now possible to write

$$\mathcal{I}(\mathbf{z}, \mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T) = E_{\mathbf{z}, \mathbf{z}_T, \mathbf{a}} \left\{ \log \frac{p(\mathbf{z} \,|\, \mathbf{a}, \mathbf{a}_T, \mathbf{z}_T)}{\sum_{\mathbf{a}} p(\mathbf{z} \,|\, \mathbf{a}, \mathbf{a}_T, \mathbf{z}_T) \cdot p(\mathbf{a})} \right\}. \tag{9}$$

Note that, since $\mathbf{a}_T$ is the known training sequence, the expectation does not have to be taken with respect to $\mathbf{a}_T$. The channel characterized by the distribution $p(\mathbf{z} \,|\, \mathbf{a}, \mathbf{a}_T, \mathbf{z}_T)$ is not memoryless, in general, because the fading coefficients $\{c_k\}$ are not perfectly known to the receiver. However, with ideal interleaving,

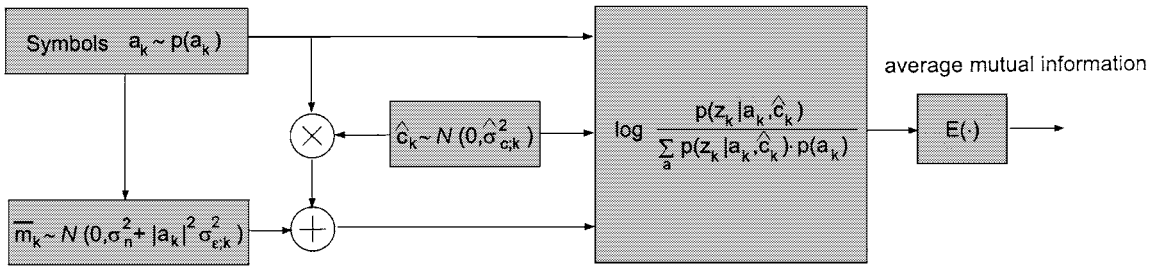

Fig. 2. Monte Carlo evaluation of the achievable rate.

and a finite-index set $\mathcal{N}_D$ with $|\mathcal{N}_D| = N_D$, we can write [7] (Appendix I outlines the effect of the interleaver/de-interleaver on mutual information in more detail)

$$p(\mathbf{z} \,|\, \mathbf{a}, \mathbf{a}_T, \mathbf{z}_T) = \prod_{k \in \mathcal{N}_D} p(z_k \,|\, a_k, \mathbf{a}_T, \mathbf{z}_T) \tag{10}$$

or equivalently

$$\mathcal{I}(\mathbf{z}, \mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T) = \sum_{k \in \mathcal{N}_D} \mathcal{I}(z_k; a_k \,|\, \mathbf{a}_T, \mathbf{z}_T). \tag{11}$$

This is a key assumption in our derivation. It must be mentioned that, exactly because the correlation between received data symbols is not used, some information is actually "thrown away." However, it is also true at the same time that, the better the channel is known, the less information is contained in these correlations. This is plausible, since for a perfectly known channel, the channel is indeed memoryless. A system with a reasonably good channel estimation scheme will therefore nevertheless exploit almost all the available information.

Remembering that our transmission model is given by $z_k = c_k a_k + m_k$, it becomes obvious that z_k and $\mathbf{z}_T$, conditioned on a_k and $\mathbf{a}_T$, are jointly Gaussian. Therefore, $p(z_k \,|\, a_k, \mathbf{a}_T, \mathbf{z}_T)$ is also normally distributed and completely described by its conditional mean and variance. From estimation theory, we know that the conditional mean $E\{z_k \,|\, a_k, \mathbf{a}_T, \mathbf{z}_T\}$ is the estimator $\hat{z}_k$ of z_k in the minimum mean square error (MMSE) sense. Since the channel model is linear and all associated quantities are Gaussian, the corresponding estimator is itself linear. The conditional mean $\hat{z}_k$ computes as

$$\begin{aligned} \hat{z}_k &= E\{z_k \,|\, a_k, \mathbf{a}_T, \mathbf{z}_T\} \\ &= a_k E\{c_k \,|\, \mathbf{a}_T, \mathbf{z}_T\} \\ &= a_k \hat{c}_k \end{aligned} \tag{12}$$

where it is recognized that $E\{c_k \,|\, \mathbf{a}_T, \mathbf{z}_T\}$ is the linear MMSE channel estimator $\hat{c}_k$. The variance of $p(z_k|a_k, \mathbf{a}_T, \mathbf{z}_T)$ is given by

$$\begin{aligned} \sigma^2_{\hat{z}_k} &= E\{|z_k - \hat{z}_k|^2 \,|\, a_k, \mathbf{a}_T, \mathbf{z}_T\} \\ &= |a_k|^2 E\{|c_k - \hat{c}_k|^2 \,|\, \mathbf{a}_T, \mathbf{z}_T\} + N_0 \\ &= |a_k|^2 \sigma^2_{\varepsilon;k} + N_0 \end{aligned} \tag{13}$$

where for the LMMSE channel estimator it can be shown that the corresponding minimum MSE is given by [4]

$$\sigma^2_{\varepsilon;k} = E\left\{|c_k - \hat{c}_k|^2 \,|\, \mathbf{a}_T, \mathbf{z}_T\right\} = E\left\{|c_k - \hat{c}_k|^2\right\}. \tag{14}$$

Therefore, the distribution $p(z_k|a_k, \mathbf{a}_T, \mathbf{z}_T)$ is normal according to

$$\begin{aligned} p(z_k \,|\, a_k, \mathbf{a}_T, \mathbf{z}_T) &= p(z_k \,|\, a_k, \hat{c}_k) \\ &= \mathcal{CN}\left(a_k \hat{c}_k, |a_k|^2 \sigma^2_{\varepsilon;k} + N_0\right) \end{aligned} \tag{15}$$

where $\mathcal{CN}$ denotes a complex Gaussian distribution. It is noticed that this distribution is a function of the MMSE $\sigma^2_{\varepsilon;k} = E\{|c_k - \hat{c}_k|^2\}$ of the channel estimate which depends on the time index k for which we wish to estimate (interpolate) the channel. However, it will be shown later on that this dependence disappears, if the channel is sampled at least with Nyquist rate. Assuming an adequately sampled channel, the achievable rate of the P-CSI channel can therefore be written as

$$\begin{aligned} C^*_{\mathrm{P-CSI}} &= \frac{1}{N} \sum_{k \in \mathcal{N}_D} \mathcal{I}(z_k; a_k \,|\, \hat{c}_k) \\ &= \frac{N_D}{N} \mathcal{I}(z_k; a_k \,|\, \hat{c}_k). \end{aligned} \tag{16}$$

Note that, because $\hat{c}_k$ is a linear combination of the Gaussian variables of $\mathbf{z}_T$, it is too Gaussian [4]. The corresponding mean and variance are given by

$$\begin{aligned} E\{\hat{c}_k\} &= E\{c_k\} - E\{c_k - \hat{c}_k\} = 0 \\ E\left\{|\hat{c}_k|^2\right\} &= \hat{\sigma}^2_{c;k} = \sigma^2_c - \sigma^2_{\varepsilon;k}. \end{aligned} \tag{17}$$

The first result is due to the fact that for the LMMSE estimator the mean of the error $e_k = c_k - \hat{c}_k$ is zero [4], and the latter result is a consequence of the orthogonality principle.

The average mutual information, as given by (9), can be evaluated numerically by a simple Monte Carlo simulation. Firstly, it is necessary to generate $\{z_k\}$ drawn according to $p(z_k \,|\, a_k, \hat{c}_k)$. This is accomplished by generating $\{a_k\} \sim p(a_k)$, multiplying them with $\{\hat{c}_k\} \sim \mathcal{CN}(0, \hat{\sigma}^2_{c;k})$, and adding $\{\bar{m}_k\} \sim \mathcal{CN}(0, |a_k|^2 \sigma^2_{\varepsilon;k} + N_0)$. The sequences $\{z_k\}, \{a_k\}$, and $\{\hat{c}_k\}$ can then be used to calculate mutual information, or the constrained capacity, as shown in Fig. 2. If we wanted to numerically evaluate the mutual information for the C-CSI channel, we would simply have to replace $\{\hat{c}_k\}$ with $\{c_k\} \sim \mathcal{CN}(0, \sigma^2_c)$, and $\{\bar{m}_k\}$ with $\{m_k\} \sim \mathcal{CN}(0, N_0)$. Comparing these two schemes, two separate effects are noticed.

- The Gaussian process $\{\bar{m}_k\}$ has a higher variance than the channel AWGN $\{m_k\}$, which leads to an effective loss in SNR.
- The optimal LMMSE estimator $\hat{c}_k$ delivers channel estimates that are orthogonal to the estimation error. This implies that the estimated fading coefficients $\{\hat{c}_k\}$ have a

TABLE I
SUMMARY OF P-SCI AND C-CSI CHANNELS

C-CSI	P-CSI
N_0	$\lvert a_k\rvert^2\sigma^2_{\varepsilon;k}+N_0$ $\sigma^2_{\varepsilon;k}=E\{\lvert c_k-\hat{c}_k\rvert^2\}$
c_k	$\hat{c}_k=E\{c_k \mid \mathbf{a}_T,\mathbf{z}_T\}$
$\sigma_c^2=E\{\lvert c_k\rvert^2\}$	$\hat{\sigma}^2_{c;k}=\sigma_c^2-\sigma^2_{\varepsilon;k}$
N no. of data symbols	$(N-N_T)$ no. of data symbols N_T no. of pilot symbols N/N_T channel usage

lower variance than the true fading coefficients, since we have $\hat{\sigma}^2_{c;k}=\sigma_c^2-\sigma^2_{\varepsilon;k}$. This, again, leads to an additional effective SNR loss.

In summary, we have just shown that, when mutual information is evaluated numerically as outlined above, the P-CSI channel is related to the C-CSI channel by making use of the substitutions shown in Table I.

IV. Bounds to Channel Capacity

It is illustrative to formulate the following transmission model:

$$\begin{aligned} z_k &= a_k c_k + m_k \\ &= a_k\hat{c}_k + a_k(c_k-\hat{c}_k)+m_k \\ &= a_k\hat{c}_k + a_k e_k + m_k \\ &= a_k\hat{c}_k + \tilde{m}_k \end{aligned} \tag{18}$$

where it is noted that the effective noise $\tilde{m}_k$ depends on the channel input a_k and is therefore, in general, not Gaussian. As a result, the expression $p(\mathbf{z}\,|\,\mathbf{a}_T,\mathbf{z}_T)=\sum_{\mathbf{a}}p(\mathbf{z}\,|\,\mathbf{a},\mathbf{a}_T,\mathbf{z}_T)\cdot p(\mathbf{a})$ in the denominator of (9) is not Gaussian either, as would be the case if the fading channel was perfectly known at the receiver. This is exactly the reason why it is easy to obtain an expression for the mutual information in terms of the SNR for the C-CSI channel, and very difficult in our case, for the P-CSI channel. It is interesting to note, however, that in the case of M-PSK signaling the effective noise $\tilde{m}_k$ is in fact Gaussian, since e_k is Gaussian and its statistics are therefore rotationally invariant to a multiplication with a_k ($|a_k|^2=\text{constant}$). Therefore, in that particular case, a channel estimate with a higher MSE will degrade $C^*_{\text{P-CSI}}$ in exactly the same way as a simple SNR loss. We will illustrate this special case with an example in Section V of this paper.

The equivalent transmission model given by (18) can also be used to derive upper and lower bounds on $C^*_{\text{P-CSI}}$. Firstly, we write $\mathcal{I}(z_k;a_k\,|\,\hat{c}_k)$ in terms of its differential entropies.

$$\mathcal{I}(z_k;a_k\,|\,\hat{c}_k)=h(z_k\,|\,\hat{c}_k)-h(z_k\,|\,a_k,\hat{c}_k) \tag{19}$$

$$=h(a_k\,|\,\hat{c}_k)-h(a_k\,|\,z_k,\hat{c}_k). \tag{20}$$

A lower bound can be obtained, as shown by Medard in [11], by upper-bounding $h(a_k|z_k,\hat{c}_k)$ in (20) as

$$h(a_k\,|\,z_k,\hat{c}_k)\le E_{\hat{c}_k}\log\left(\pi e\frac{\sigma_a^4\sigma_\varepsilon^2+\sigma_a^2N_0}{|\hat{c}_k|^2\sigma_a^2+\sigma_a^2\sigma_\varepsilon^2+N_0}\right). \tag{21}$$

Since a_k and $\hat{c}_k$ are independent, we have $h(a_k\,|\,\hat{c}_k)=h(a_k)$. Furthermore, assuming Gaussian signaling ($a_k\sim\mathcal{CN}(0,\sigma_a^2)$) which maximizes $h(a_k)$, it is possible to obtain a lower bound on the true channel capacity of the P-CSI channel.

$$\begin{aligned} C_{\text{P-CSI}} &\ge \frac{N_D}{N}\left(E_{\hat{c}_k}\log\left(|\hat{c}_k|^2\sigma_a^2+\sigma_a^2\sigma_\varepsilon^2+N_0\right)\right. \\ &\quad\left.-\log\left(\sigma_a^2\sigma_\varepsilon^2+N_0\right)\right) \\ &= \frac{N_D}{N}\cdot E_{\hat{c}_k}\log\left(1+\frac{|\hat{c}_k|^2\sigma_a^2}{\sigma_\varepsilon^2\sigma_a^2+N_0}\right). \end{aligned} \tag{22}$$

Equation (19), on the other hand, can be used to obtain an upper bound on $C^*_{\text{P-CSI}}$ by noting that the entropy $h(z_k|\hat{c}_k)$ is upper bounded by the entropy of a Gaussian variable with variance $E\{|z_k|^2\,|\,\hat{c}_k\}$. Since z_k given a_k and $\hat{c}_k$ is Gaussian and described by $p(z_k\,|\,a_k,\hat{c}_k)$ of (15), we have that

$$\begin{aligned} C^*_{\text{P-CSI}} &\le \frac{N_D}{N}\left(E_{\hat{c}_k}\log\left(|\hat{c}_k|^2\sigma_a^2+\sigma_a^2\sigma_\varepsilon^2+N_0\right)\right. \\ &\quad\left.-E_{a_k}\log\left(|a_k|^2\sigma_\varepsilon^2+N_0\right)\right) \\ &= \frac{N_D}{N}E_{\hat{c}_k}\log\left(1+\frac{|\hat{c}_k|^2\sigma_\varepsilon^2}{\sigma_a^2\sigma_\varepsilon^2+N_0}\right) \\ &\quad+\underbrace{\frac{N_D}{N}E_{\hat{a}_k}\log\left(\frac{\sigma_a^2\sigma_\varepsilon^2+N_0}{|a_k|^2\sigma_\varepsilon^2+N_0}\right)}_{\Delta\ge 0}. \end{aligned} \tag{23}$$

It is noted that upper and lower bound are related via Δ where, by Jensen's inequality, we have that $\Delta\ge 0$. For Gaussian signaling, Δ is therefore a measure for the tightness of the upper and lower bounds. Furthermore, it is seen that the smaller the estimation error σ_ε^2 is, the tighter both bounds become, and for $\sigma_\varepsilon^2\to 0$, apart from the linear factor N_D/N, both converge to the capacity expression of the perfectly known (C-CSI) fading channel. Additionally, it is noted that, for M-PSK signaling, the upper bound of (23) is identical to the lower bound of (22) for Gaussian signaling.

V. Channel Estimation

In the previous section, it was shown that LMMSE channel estimation follows directly from the derivation of the average mutual information for the synchronized detection (P-CSI) channel. In a receiver, LMMSE channel estimation would be performed in the so-called inner receiver, as illustrated by Fig. 1. For the purpose of channel estimation, known training symbols are inserted into the data message $\mathbf{a}'$. There are several possible formats for pilot insertion. Here, for the capacity analysis, we assume pilot symbols which are distributed with a period of T_s symbols, as indicated by way of example in Fig. 3. LMMSE channel estimation is then performed by firstly computing maximum likelihood (ML) channel snapshots $x_k=z_{T;k}\cdot(a^*_{T;k}/|a_{T;k}|^2)$; one might also say that the channel is thus sampled at a normalized rate of $1/T_s$. Secondly, a

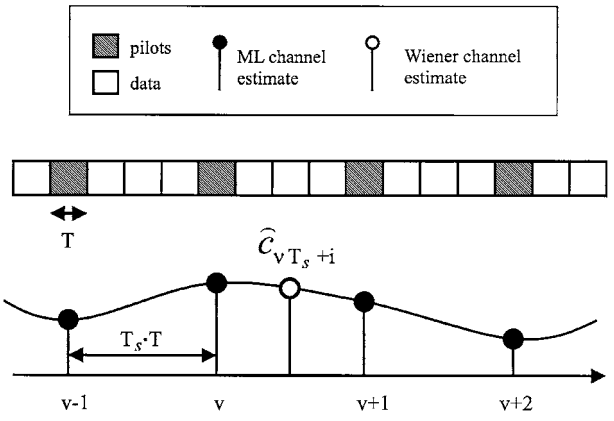

Fig. 3. LMMSE (Wiener) channel estimation.

Wiener interpolator is lowpass-filtering these channel snapshots in order to calculate the LMMSE channel estimates at any desired time instant k within the data message $\mathbf{a}'$. This process is illustrated in the lower part of Fig. 3. For a detailed discussion on Wiener channel interpolation, please refer to [1, ch. 14].

The quality of the LMMSE channel estimate is characterized by the error covariance $\sigma^2_{\varepsilon;k}$ which is directly related to the channel dynamics via the Doppler spectrum $S(e^{j\omega})$ and the normalized channel sampling rate $1/T_s$. On mobile channels, the fading process $\{c'_k\}$ is often modeled to exhibit the U-shaped Jakes Doppler spectrum (or Clarke Doppler spectrum, following the original publication) [6], [5]. The Jakes spectrum, which is based on the isotropic scattering assumption, accentuates instantaneous Doppler shifts near the cutoff frequency F (F is normalized to the symbol rate $1/T$). However, since the actual shape of the Doppler spectrum has no noticeable effect on the estimator performance [1, pp. 651 and 658], for the purpose of the capacity calculations, the Jakes Doppler spectrum may as well be replaced by an ideal low-pass spectrum with the same cutoff frequency F, i.e., we have

$$S(e^{j\omega}) = \begin{cases} \frac{\sigma_c^2}{2F} & \text{if} |\omega| < 2\pi F \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Assuming, in an information theoretic framework, an infinite length pilot symbol vector $\mathbf{a}_T$, frequency-domain Wiener filter theory becomes applicable. Denoting the optimal Wiener filter $W(e^{j\omega})$, the error covariance $\sigma^2_{\varepsilon;k}$ is given by

$$\begin{aligned} \sigma^2_{\varepsilon;k} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} S(e^{j\omega})(1 - W(e^{j\omega}))\, d\omega \\ &= \frac{2FT_s}{2FT_s + \bar{\gamma}_s} \cdot \sigma_c^2 \end{aligned} \tag{25}$$

with

$$W(e^{j\omega}) = \frac{S\left(e^{j\omega}\right)}{S(e^{j\omega}) + T_s \cdot (N_0/\sigma_a^2)}. \tag{26}$$

The error covariance is independent of the time index k as long as the fading process is strictly bandlimited and the channel is sampled at least with Nyquist rate. Therefore, when computing $C^*_{\mathrm{P-CSI}}$, no additional averaging over error covariances is necessary. The derivation of the Wiener filter and its error covariance is outlined in Appendix II. Furthermore, note that (25) refers to the error covariance of the channel estimates for the fading process $\{c'_k\}$ which is characterized by its Doppler spectrum $S(e^{j\omega})$. Since the de-interleaver maps $\{\hat{c}'_k\} \rightarrow \{\hat{c}_k\}$, we have that

$$\sigma^2_{\varepsilon} = E\{|c_k - \hat{c}_k|^2\} = E\{|c'_k - \hat{c}'_k|^2\}. \tag{27}$$

Before the de-interleaver, the channel estimates are the colored sequence $\{\hat{c}'_k\}$ with power spectrum $S(e^{j\omega})$ and, after de-interleaving, the channel estimates become the white sequence $\{\hat{c}_k\}$. Since the pilot symbols are inserted into the transmitter after the interleaver, the received pilots and the received data are correlated via $S(e^{j\omega})$, whereas the de-interleaver has destroyed any correlation between received data symbols.

Before the Monte Carlo simulation results are discussed in the next section, we would like to give an outlook as to what kind of results can be expected. Therefore, we briefly consider the SNR loss as a function of the normalized channel sampling period T_s for a simple M-PSK system with $|a_k|^2 = \text{constant}$. The SNR of the P-CSI channel is given by

$$\mathrm{SNR}_{\mathrm{P-CSI}} = \bar{\gamma}_s \cdot \frac{1}{1 + (1 + \bar{\gamma}_s) 2FT_s} = \bar{\gamma}_s \cdot \eta_C \tag{28}$$

and hence we may define an SNR loss factor η_C such that $\eta_C \cdot \mathrm{SNR}_{\mathrm{P-CSI}} = \mathrm{SNR}_{\mathrm{C-CSI}} = \bar{\gamma}_s$. From the previous equation, it is noted that, in order for η_C to vanish (i.e., $\eta_C \rightarrow 1$) as $F \rightarrow 0$, it is necessary that $1/(2T_s) \rightarrow 0$ faster than F. This condition is satisfied if we choose

$$\frac{1}{T_s} \geq (2F)^{1-p} \tag{29}$$

where p is a real number in the open set $\{0, 1\}$. The SNR loss is minimized for some p in this interval. Due to pilot symbol insertion, the total amount of usable data symbols is $N_D = N - N_T$. Since the channel is sampled with a rate $1/T_s$, the fraction of training symbols to data symbols is given by

$$\eta_{T_s} = \frac{N_D}{N} = \frac{T_s - 1}{T_s}. \tag{30}$$

Therefore, in this example, the constrained capacity can be written as

$$C^*_{\mathrm{P-CSI}} = \eta_{T_s} \cdot C^*_{\mathrm{C-CSI}}(\eta_C \bar{\gamma}_s). \tag{31}$$

In other words, in the case of M-PSK modulation, the achievable rate of the P-CSI channel can be written as a function of the achievable rate of the C-CSI channel which is subject to an SNR loss η_C and scaled proportional to the fraction of inserted training symbols $\eta_{T_s} = 1 - 1/T_s$. Furthermore, since for the appropriate choice of T_s, both η_{T_s} and η_C approach 1 as $F \rightarrow 0$, it can be concluded that the mutual information of the P-CSI channel approaches the mutual information of the C-CSI channel for $F \rightarrow 0$. Additionally, it is noticed that the SNR loss η_C of the flat fading channel is strongly dependent on the channel dynamics (via F), but only weakly dependent on the SNR itself. This indicates that, for rapidly varying channels,

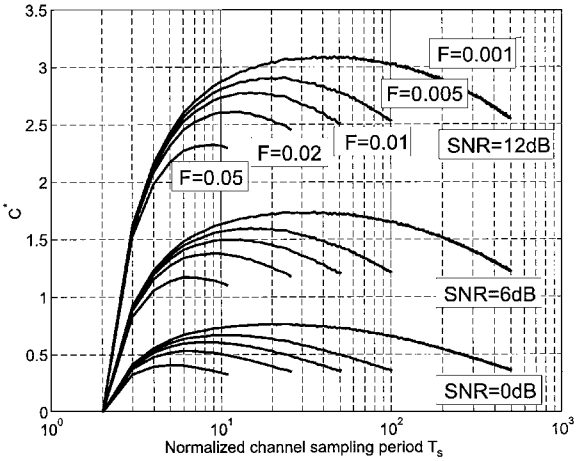

Fig. 4. $C^{*}_{\mathrm{P-CSI}}$ versus normalized channel sampling period T_s.

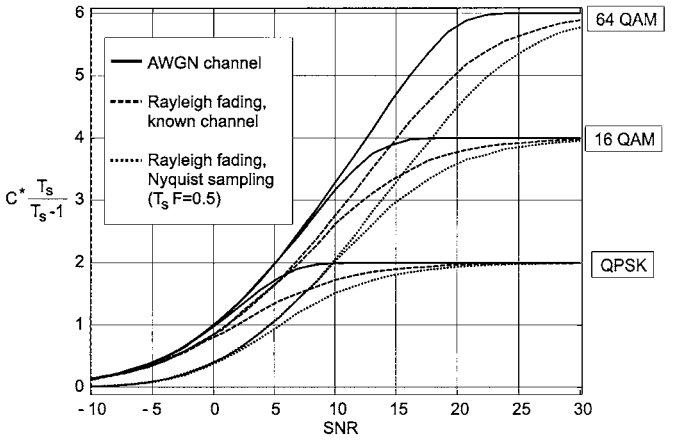

Fig. 5. Achievable rates of the AWGN channel, the C-CSI channel, and the P-CSI channel with Nyquist rate channel sampling.

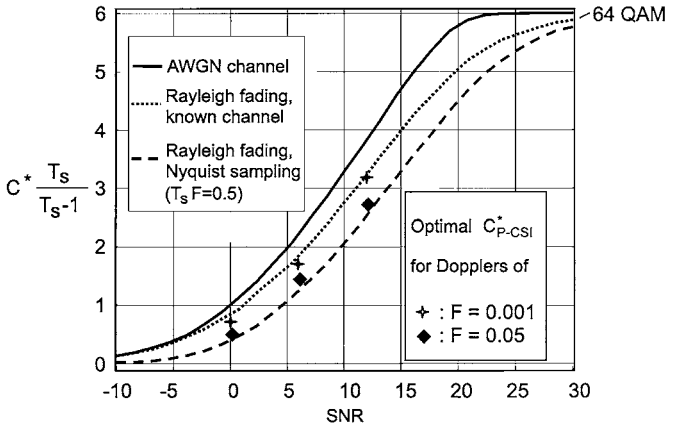

Fig. 6. Achievable rates of the P-CSI channel with optimal channel sampling period T_s for two different Doppler frequencies and the C-CSI channel.

the mutual information of the C-CSI channel overestimates the achievable rate, since it is entirely independent of the channel dynamics.

VI. Constrained Capacity Analysis

With the results outlined in the previous two sections, the constrained capacity $C^{*}_{\mathrm{P-CSI}}$ as given by (5) was calculated via Monte Carlo simulation. All simulations assume a M-ary communication system with QAM modulation and equally probable symbols, i.e., $p(a_k) = 1/M$. It is important to note that mutual information is a function of F, the Doppler frequency of the channel fading process, and T_s, the channel sampling period. Therefore, it becomes possible to optimize the constrained channel capacity for a flat fading transmission system with respect to the channel sampling rate. Fig. 4 shows the result for $M = 64$, plotting $C^{*}_{\mathrm{P-CSI}}$ versus the normalized channel sampling period T_s for several typical SNRs and channel dynamics. In the figure, the curves are only shown for channel sampling periods T_s smaller or equal to the Nyquist rate $T_s F = 0.5$. The results indicate clearly that for slow fading channels the difference between the optimal T_s and Nyquist rate channel sampling is much larger than for the fast fading channels. In that sense, an optimal choice of T_s is much more important for a slow fading channel than for a fast fading channel, because the capacity gains are larger. It is also observed that the maximum for the slower fading channels becomes flatter, indicating that, provided one is in the proper range, the constrained channel capacity is not very sensitive to the choice of T_s. Therefore, it would be possible to choose a sampling period which is somewhat larger than the optimum without incurring any noticeable degradation. This finding is a very useful guideline for receiver design, since a lower channel sampling rate translates directly into a reduced computational effort required for the synchronization tasks of the inner receiver. Finally, another interesting conclusion that can be drawn from Fig. 4 is that sampling the channel at the Nyquist rate of $FT_s = 0.5$ never results in an optimal constrained channel capacity.

Next, in Fig. 5, the achievable rate is shown as a function of SNR for three different types of channels: The AWGN channel, the Rayleigh fading channel with perfectly known fading process, and Rayleigh fading with channel estimation based on (suboptimally) sampling the channel at Nyquist rate ($FT_s = 0.5$). In this case the function $\mathcal{I}(\mathbf{z}, \mathbf{z}_T; \mathbf{a} \,|\, \mathbf{a}_T)$ solely depends on the average SNR $\bar{\gamma}$. Note, that the curves show $C^{*}_{\mathrm{P-CSI}} \cdot T_s/(T_s - 1)$, thereby ignoring any losses directly due to pilot symbol insertion. This simply means that for a specific Doppler frequency F, the curve has to be compressed correspondingly by $1 - 1/T_s$ to arrive at the corresponding achievable rate. If, for example, we assume a fairly slowly varying fading process with $F = 0.005$, the capacity shown in Fig. 5 has to be adjusted by $1 - 1/T_s = 0.99$. The Nyquist rate is the slowest sensible rate at which to sample the channel. Therefore, this representation allows us to illustrate the range of typical SNR losses for the synchronized detection channel (P-CSI) as compared to the AWGN channel and the Rayleigh fading channel with perfectly known fading process (C-CSI). Comparing the results of the C-CSI channel with the results of the P-CSI channel, the figure demonstrates, that the SNR loss is limited to approximately 3 dB for slowly varying channels, but it can be somewhat higher for rapidly varying channels, since then the factor $1 - 1/T_s$ cannot be neglected. In Fig. 6, the results of the channel sampling rate optimization are shown in conjunction with the constrained capacity curves for 64-QAM of Fig. 5. For clarity, only the two extreme Doppler

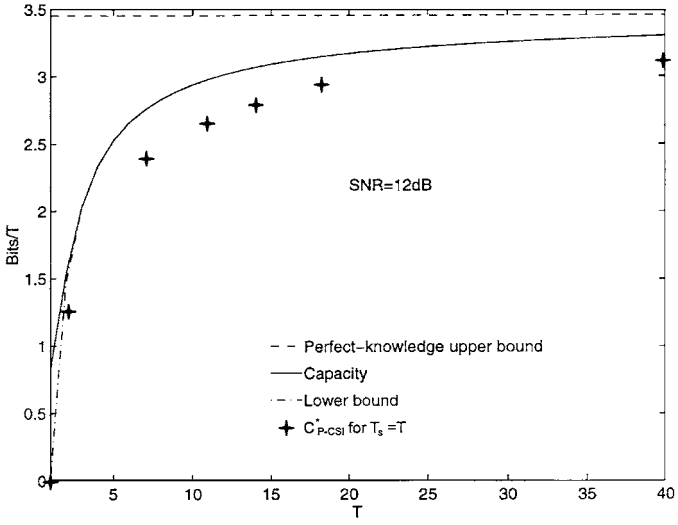

Fig. 7. $C^{*}_{\mathrm{P-CSI}}$ versus Marzetta's approach.

frequencies $F = 0.001$, and $F = 0.05$ where chosen, and their corresponding optimal capacities for SNRs of 0, 6, and 12 dB are plotted. Again, we do not include the factor $1 - 1/T_s$. This figure reinforces the results shown earlier, demonstrating that, for the slowly varying channel $(F = 0.001)$, the capacity of the perfectly known channel is almost achieved, whereas for the rapidly varying channel $(F = 0.05)$, the SNR loss comes close to that of Nyquist rate channel sampling.

Finally, we would like to compare our approach to some aspects of the recently published work by Marzetta and Hochwald [8]. To our knowledge, theirs is the first publication which considers channel capacity as a function of channel dynamics. In [8], the capacity calculations neither assume a known channel nor a given symbol constellation, as in our case, and indeed perform the required optimization over the input symbol distribution. In order to keep the analysis of this complicated problem tractable, Marzetta *et al.* assume that the fading process, which neither the transmitter nor the receiver knows, are constant for T symbol periods, after which it changes to a new independent random value which they maintain for another T symbol periods, and so on. The resulting capacity is displayed in Fig. 7 as a function of T, along with the perfect channel knowledge upper bound and a lower bound, for an SNR of 12 dB. In the same figure, the stars underneath the capacity curves indicate the constrained capacity obtained from Fig. 4 for the different Fs and their respective optimal channel sampling period T_s. The problem, however, are the incompatible channel models used, and hence it is difficult to find a meaningful correspondence between T and F. Here, since T_s is a function of F, we just arbitrarily set $T = T_s$. As a consequence, the absolute values of the resulting capacities cannot be compared, but it nevertheless shows that both approaches lead to the same qualitative result.

VII. Conclusion

The achievable rate, or constrained capacity, of a typical digital communication system following the principle of synchronized detection was derived under the assumption that the system operates on a flat fading channel and uses an interleaver to combat bursty error events. This framework was used to calculate the optimal channel sampling rate $1/T_s$ with respect to the achievable rate. It was shown that the optimal channel sampling rate is always above Nyquist rate $(FT_s = 0.5)$. For slowly varying channels, a proper choice of $1/T_s$ is very important in order to limit any losses in the achievable rate. In the case of rapidly varying channels, the optimum sampling interval translates into constrained capacities which do not differ much from what would be obtained by sampling the channel at Nyquist rate. Furthermore, for the more slowly varying channels, a flat maximum means that a somewhat lower than optimal sampling rate can be used which reduces receiver complexity. Furthermore, it was shown, that in most cases the principle of synchronized detection leads to SNR losses no more than approximately 3 dB, except channels that suffer from fast fading. For slower fading, the achievable rate of the perfectly known fading channel is almost achieved. Finally, it was shown that the results presented here are of the same qualitative nature as the results recently published in [8].

For clarity, in this paper we focused solely on so-called single input-single output communication systems. However, it is very straightforward to extend the principles presented here to the more general case of multiple input-multiple output communication systems which recently has been paid so much attention to in the literature. In the case of MIMO systems the correct choice of training data becomes much more important, since a lot more channels have to be estimated which implies a greatly increased need of training data. A corresponding paper has been published [14].

VIII. Remarks to [11] and [12]

The reviewers brought to our attention [11] and [12] which were published after submission of our paper. We would like mention here that, even if these papers consider a similar problem, the approach taken is entirely different. Here, the achievable rate of a flat fading communication system with perfect interleaving is computed, whereas [11] derives capacity bounds for a channel without interleaving, and [12] uses these bounds in the framework of block-fading channels. Furthermore, it is shown here that the MMSE channel estimator and the corresponding error covariance follow directly from the derivation of mutual information, whereas in [11] channel estimation is introduced in an ad-hoc fashion and not an outcome of the derivation. The quantitative results presented here assume a continuously time-varying (Rayleigh fading) channel model unlike the block-fading model used in [12].

Appendix I
Mutual Information of a Flat Fading Channel With Interleaver/De-Interleaver

If the chain rule for information [2] is applied to (8), it can be rewritten as

$$\mathcal{I}(\mathbf{z};\mathbf{a}\,|\,\mathbf{a}_T,\mathbf{z}_T) = \sum_{j\in\mathcal{N}_D}\sum_{k\in\mathcal{N}_D}\mathcal{I}(z_j;a_k\,|\,z_0\ldots z_{j-1},a_0\ldots a_{k-1},\mathbf{a}_T,\mathbf{z}_T). \quad (32)$$

Furthermore, it is assumed that the interleaving/de-interleaving operation has completely broken up the channel memory. Consequently, a received symbol z_k does not convey any information about $a_{j\neq k}$. Thus, for $j \neq k$, we have that

$$\underbrace{\sum_{j\in\mathcal{N}_D}\sum_{k\in\mathcal{N}_D}}_{j\neq k} \mathcal{I}(z_j; a_k \,|\, z_0 \ldots z_{j-1}, a_0 \ldots a_{k-1}, \mathbf{a}_T, \mathbf{z}_T) = 0. \tag{33}$$

Likewise, since any z_k or a_k do not depend on either $z_0, \ldots, z_{k-1}$ nor on $a_0, \ldots, a_{k-1}$, for $k = j$, we arrive at

$$\begin{aligned}
&\underbrace{\sum_{j\in\mathcal{N}_D}\sum_{k\in\mathcal{N}_D}}_{j=k} \mathcal{I}(z_j; a_k \,|\, z_0 \ldots z_{j-1}, a_0 \ldots a_{k-1}, \mathbf{a}_T, \mathbf{z}_T) \\
&\quad = \sum_{k\in\mathcal{N}_D} \mathcal{I}(z_k; a_k \,|\, z_0 \ldots z_{k-1}, a_0 \ldots a_{k-1}, \mathbf{a}_T, \mathbf{z}_T) \\
&\quad = \sum_{k\in\mathcal{N}_D} \mathcal{I}(z_k; a_k \,|\, \mathbf{a}_T, \mathbf{z}_T).
\end{aligned} \tag{34}$$

The flat fading channel with interleaver/de-interleaver is therefore completely characterized by the distribution $p(z_k|a_k, \mathbf{a}_T, \mathbf{z}_T)$.

Appendix II
Derivation of the Wiener Channel Interpolator

In order to compute a simple, analytic expression for the error covariance, it proves convenient to apply frequency-domain Wiener filter theory. Therefore, in an information theoretic framework, an infinite length pilot symbol vector $\mathbf{a}_T$ is assumed. Denoting the ML channel estimates $\{x_k\}$, and the Wiener filter $\{\tilde{w}_{i;k}\}$, the channel estimate can be expressed as

$$\hat{c}_k = \hat{c}_{\nu T_s+i} = \sum_{n=-\infty}^{\infty} \tilde{w}_{i;n} x_{\nu T_s - nT_s} \tag{35}$$

where $k = \nu T_s + i$ is used to indicate that the channel is interpolated at time index i relative to the ML channel samples which are available at time indices νT_s. Applying the orthogonality principle and the discrete-time Fourier transform, and provided that the channel is sampled at least with Nyquist rate, the optimal Wiener filter in the frequency domain is given by

$$\begin{aligned}
\tilde{W}_i(e^{j\omega}) &= e^{j\omega i} \cdot \frac{S\left(e^{\frac{j(\omega \bmod 2\pi)}{T_s}}\right)}{S\left(e^{\frac{j(\omega \bmod 2\pi)}{T_s}}\right) + T_s \cdot (N_0/\sigma_a^2)} \\
&= e^{j\omega i} \cdot W\left(e^{\frac{j(\omega \bmod 2\pi)}{T_s}}\right).
\end{aligned} \tag{36}$$

It is recognized that $\tilde{W}_i(e^{j\omega})$ has the same spectral characteristics for every relative time index i, except for a phase shift $e^{j\omega i}$, which justifies the second step in the above equation. Therefore, rewriting (35), the channel estimate can now be expressed as

$$\hat{c}_{\nu T_s+i} = \sum_{n=-\infty}^{\infty} w_{nT_s+i} x_{\nu T_s - nT_s} \tag{37}$$

with

$$W(e^{j\omega}) = \frac{S(e^{j\omega})}{S(e^{j\omega}) + T_s \cdot (N_0\sigma_a^2)}. \tag{38}$$

Similarly, defining the channel autocorrelation function as $\alpha^*_{kT_s+i} = E\{c_{lT_s-kT_s} c^*_{lT_s+i}\}$, we can compute the error covariance $\sigma^2_{\varepsilon;\nu T_s+i}$ using the orthogonality principle and applying Parseval's theorem.

$$\begin{aligned}
\sigma^2_{\varepsilon;\nu T_s+i} &= E\left\{\left|c_{\nu T_s+i} - \hat{c}_{\nu T_s+i}\right|^2\right\} \\
&= \alpha^*(0) - \sum_{k=-\infty}^{\infty} w_{kT_s+i} \alpha^*_{kT_s+i} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} S(e^{j\omega})(1 - W(e^{j\omega}))\, d\omega \\
&= \frac{2FT_s}{2FT_s + \bar{\gamma}_s} \cdot \sigma_c^2
\end{aligned} \tag{39}$$

where the rectangular Doppler spectrum $S(e^{j\omega})$ of (24) was assumed. It is important to note that the error covariance does not depend on the relative time index i (or the absolute time index k) at which we wish to interpolate the channel. This independence is a direct result of applying Parseval's theorem which is invariant to a time shift.

Acknowledgment

The authors would like to thank one anonymous reviewer for the constructive comments which greatly improved the paper.

References

[1] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*. New York, NY: Wiley, 1998.
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY: John Wiley and Sons, 1991.
[3] J. G. Proakis, *Digital Communications*, 2nd ed. New York, NY: McGraw-Hill, 1989.
[4] S. M. Kay, *Fundamentals of Statistical Signal Processing—Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[5] W. C. Jakes, *Microwave Mobile Communications*. New York, NY: Wiley, 1974.
[6] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, vol. 47, pp. 957–1000, July/Aug. 1968.
[7] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.
[8] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.
[9] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.
[10] A. Lapidoth and S. Shamai (Shitz), "Fading Channels: How perfect need 'perfect side information' be?," in *Proc. Information Theory and Communications Workshop*, 1999, pp. 36–38.
[11] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, pp. 933–946, May 2000.
[12] A. Sabharwal, E. Erkip, and B. Aazhang, "On channel state information in multiple antenna block fading channels," in *Proc. Int. Symp. Information Theory and Its Applications*, Honululu, HI, Nov. 2000.
[13] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.
[14] J. Baltersee, G. Fock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," IEEE J. Select. Areas Commun., to be published.
[15] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 55–67, Jan. 1982.

Jens Baltersee received the M.Eng. degree in electrical and electronic engineering from Imperial College of Science, Technology, and Medicine, London, U.K., in 1996. He is currently working toward the Ph.D. degree at the Aachen University of Technology, Aachen, Germany.

He is a Research Assistant at the Institute for Integrated Signal Processing Systems (ISS), Aachen University of Technology. His current research interests include information theoretic apsects of space–time processing and digital wireless receiver design with emphasis on synchronization and detection.

Gunnar Fock received the Dipl.-Ing. degree in electrical engineering from Aachen University of Technology (RWTH Aachen), Germany, in 1996. Currently, he is a Research Assistant at the Institute for Integrated Signal Processing Systems (ISS), RWTH Aachen, working toward the Ph.D. degree. His current research interests include information theoretic aspects of MIMO systems, and digital receiver design with emphasis on synchronization, estimation, and equalization.

Heinrich Meyr (M'75–SM'83–F'86) received the M.S and Ph.D. degrees from ETH Zurich, Switzerland.

He spent over 12 years in various research and management positions in industry before accepting a professorship in Electrical Engineering at Aachen University of Technology (RWTH Aachen), Aachen, Germany, in 1977. He has worked extensively in the areas of communication theory, synchronization, and digital signal processing for the last thirty years. His research has been applied to the design of many industrial products. At RWTH Aachen, he heads an institute involved in the analysis and design of complex signal processing systems for communication applications. He was a co-founder of CADIS GmbH (acquired 1993 by Synopsys, Mountain View, CA), a company that commercialized the tool suite COSSAP. He is currently serving on the board of two companies in the communications industry. Recently he was also appointed as a member of the technical advisory board of MorphIC's, Cupertino, CA. He has published numerous IEEE papers and holds many patents. He is the author (together with Dr. G. Ascheid) of the book *Synchronization in Digital Communications*(New York: Wiley, 1990) and of the book *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing* (together with Dr. M. Moeneclaey and Dr. S. Fechtel) (New York: Wiley, 1997). In 1998, he was a Visiting Scholar at the University of California at Berkeley's (UC Berkeley) Wireless Research Centre (BWRC). He was elected the "Mc Kay Distinguished Lecturer" at the Electrical Engineering department of UC Berkeley for the spring term 2000. The lecture series was on "Design and Implementation of Advanced Digital Receivers for Wireless Communications." He is also the recipient of the prestigious "Mannesmann Innovation Prize" for the year 2000, which is awarded for outstanding contribution to the area of wireless communication. He has also served as Vice President for International Affairs of the IEEE Communications Society.