# A Bit-Mapping Strategy for Joint Iterative Channel Estimation and Turbo-Decoding

Susanne Godtmann*, Helge Lüders†, Gerd Ascheid*, and Peter Vary†

*Institute for Integrated Signal Processing Systems, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany

†Institute of Communication Systems and Data Processing, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany

***Abstract*— In this paper, we investigate Turbo-coded transmission over a temporally correlated flat Rayleigh fading channel. Conventionally, channel estimation is performed prior to decoding and is solely based on periodically inserted pilot symbols. We here consider the case that the pilot spacing is not small enough to satisfy the Nyquist criterion for the bandlimited channel process. It is demonstrated that iterative code-aided channel estimation can then significantly improve the bit error rate (BER). Furthermore, we show that the convergence speed of the iterative channel estimation can be significantly accelerated without additional computational power by just carefully choosing the mapping strategy of systematic bits to data symbols.**

## I. Introduction

With the invention of Turbo codes [1], iteratively operating receivers have moved more and more into the focus of research. As compared to non-iterative receiver structures, such systems are capable of achieving a certain bit error rate (BER) at significantly reduced signal-to-noise ratios (SNR).

The integration of synchronization into the iterative decoding process, commonly referred to as joint iterative synchronization and decoding [2], Turbo synchronization [3] or code-aided synchronization [4], is a promising approach to face the challenge of accurate synchronization at low SNRs with high spectral efficiency. The basic idea is to use information which is available after each decoding step to compute estimates of the transmitted symbols. With the aid of these estimated symbols, channel estimates are generated and subsequently used to correct the received signal prior to performing further decoding iterations. The synchronization can, thus, be described as code-aided.

In the case of convergence, code-aided channel estimation guarantees very high accuracy and, therefore, also good decoding results. However, code-aided channel estimation only converges within a narrow estimation range and its convergence speed has to be considered carefully. Accurate initialization of the iterative process plays a decisive role in this context, e.g. [5]. It is important that the very first estimate that is used to synchronize the received symbols prior to the first decoding step is of sufficient accuracy in order to enable the convergence of the (Turbo) decoder. The more accurate the estimate is, the faster the code-aided synchronization converges.

Considering the estimation of a stationary time-variant flat Rayleigh fading channel, the accuracy of the initial estimate is determined by the sampling rate of the time-variant process by pilot symbols, the so-called pilot spacing. We here consider the case where the pilot spacing is not small enough to satisfy the Nyquist criterion for the bandlimited channel process and refer to this scenario as the *undersampled* case. Undersampling the bandlimited channel process is especially attractive for systems with multiple transmit antennas, since then the number of necessary pilot symbols scales with the number of transmit antennas. Hence, depending on the channel dynamics and the number of transmit antennas that are involved, quite a significant amount of bandwidth and transmit power needs to be allocated to pilot symbols in order to satisfy the Nyquist criterion for the bandlimited channel process. Not satisfying the Nyquist criterion can therefore be considered an attractive option. In this paper, we only consider the single antenna case, but it should be kept in mind that the results are straight forward to extend for the multiple antenna case. In the context of an undersampled bandlimited channel process, we investigate (iterative) code-aided channel estimation for a Turbo-coded systems and we demonstrate that this approach succeeds in significantly improving the BER performance. Furthermore, it is shown that the positioning of systematic symbols, i.e. data symbols that consist only of systematic bits, within the burst significantly influences the convergence speed for code-aided channel estimation.

The paper is structured as follows: the transmission model and the employed code-aided channel estimator are described in Section II and Section III, respectively. Section IV explores the proposed mapping strategy and simulation results are discussed in Section V. Section VI concludes this paper.

## II. Transmission Model

The transmission model considered in this paper is shown in Fig. 1. Information bits are grouped into packets of $N$ bits, encoded with a rate 1/3 Turbo code, interleaved[1] and mapped onto a modulation alphabet. We limit the modulation alphabets to M-ary PSK, i.e. modulations with constant amplitude.

The resulting data symbols are then transmitted over a flat Rayleigh fading channel that also suffers from additive white Gaussian noise (AWGN). The fading is assumed to be temporally correlated. The maximum Doppler spread $f_d$ normalized to the symbol duration $T_s$ is given by $F_d = f_d T_s$.

This work has been supported by the UMIC (Ultra High-Speed Mobile Information and Communication) research centre. Furthermore, S. Godtmann thanks the Deutsche Telekom Stiftung for its financial support.

[1]The systematic bits and the parity bits are interleaved separately.

978-1-4244-1722-3/08/$25.00 ©2008 IEEE.

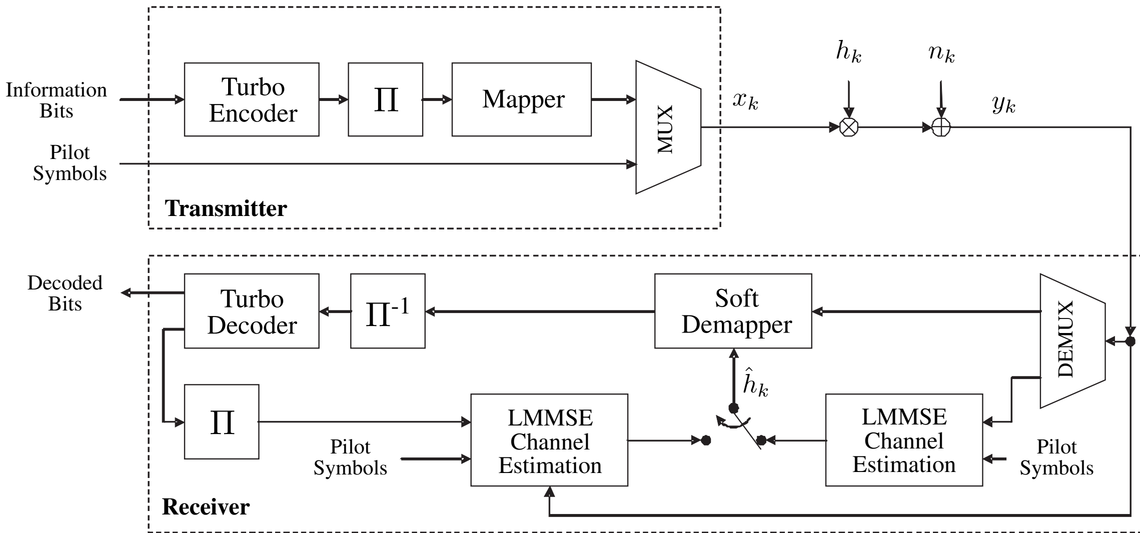



Figure 1. System Model

At transmit side, pilot symbols are multiplexed into the data symbol stream periodically. Two pilot symbols are then separated by $P_S - 1$ data symbols and $P_S$ is also referred to as *pilot spacing*.

The corresponding signal model is given as

$$y_k = h_k \cdot x_k + n_k, \quad (1)$$

where $h_k$ is the complex fading coefficient at time instance $k$ with $\mathrm{E}\{|h_k|^2\} = 1$ and $x_k$ is the transmitted symbol with $|x_k|^2 = 1$. $n_k$ is a sample of complex-valued AWGN with independent real and imaginary part, each having zero-mean and variance $N_0/(2E_s)$.

At receiver side, the received data samples are corrected with channel estimates that are obtained via a Wiener filtering process [6] with filter length $F_{\mathrm{W}}^{[\mathrm{init}]}$ and the subsequently demapped coded bits are deinterleaved and then decoded. Up to here, this is a classical pilot symbol assisted system (PSAM) [7], [8], where only pilot symbols contribute to the channel estimate. After soft-demapping and decoding, estimates of the data samples are now fed back from the decoding unit and used to improve the channel estimation accuracy and also the decoding result. This feedback-loop will be discussed in more detail in Section III.

## III. Code-Aided Channel Estimation

In this section, we shortly summarize the principle of code-aided channel estimation that we make use of in this paper. For a more detailed analysis, the reader is referred to e.g. [9].

For the code-aided estimation of a time-variant channel, information feedback of the channel decoder is used in order to estimate the transmitted symbols [9]. For the case of no feedback from the decoder (initial channel estimation) and the case of perfect feedback from the decoder, the linear minimum mean squared error (LMMSE) channel estimator is optimum [6]. Therefore, in accordance with e.g. [9], we here consider LMMSE channel estimation also for the case of imperfect feedback from the decoder.

For the observation of the channel $\tilde{h}_k$ at the data symbol positions, we then get

$$\tilde{h}_k = y_k \cdot \hat{x}_k^* = h_k \cdot x_k \cdot \hat{x}_k^* + n_k \cdot \hat{x}_k^*, \quad (2)$$

where $\hat{x}_k$ denotes the (hard) estimate of the transmitted symbol at time instant $k$ and $(\cdot)^*$ denotes the complex conjugate. It has been shown by means of simulation, e.g. [9], that soft symbols outperform hard symbols also in the case of code-aided LMMSE channel estimation. However, since we here solely focus on mapping strategies, we will restrict to the usage of hard symbols.

For the (offline) calculation of the Wiener filter coefficients, in accordance with [9], we assume perfect feedback from the decoder. The iterative channel estimate for the $k$-th symbol can then be given as:

$$\hat{h}_k = \boldsymbol{w}^{\mathrm{T}} \tilde{\boldsymbol{h}}, \quad (3)$$

where $\tilde{\boldsymbol{h}}$ is the concatenation of the $F_{\mathrm{W}}^{[\mathrm{itr}]}$ channel observations that are centered around $k$. According to LMMSE estimation theory, the filter coefficients $\boldsymbol{w}$ can be obtained as

$$\boldsymbol{w} = \mathbf{C}_{zz}^{-1} \cdot \mathbf{C}_{zh} \quad (4)$$

where $\mathbf{C}_{zz}$ is the $(F_{\mathrm{W}}^{[\mathrm{itr}]} \times F_{\mathrm{W}}^{[\mathrm{itr}]})$ covariance matrix given by

$$\mathbf{C}_{zz} = \mathbf{R}_{hh} + \mathbf{I}\,\sigma_n^2 \quad (5)$$
$$= \begin{pmatrix} r_{hh}(0) + \sigma_n^2 & \cdots & r_{hh}(F_{\mathrm{W}}^{[\mathrm{itr}]} - 1) \\ \vdots & \ddots & \vdots \\ r_{hh}(-(F_{\mathrm{W}}^{[\mathrm{itr}]} - 1)) & \cdots & r_{hh}(0) + \sigma_n^2 \end{pmatrix},$$

and $r_{hh}(l)$ represents the autocorrelation function of the fading process. In the case of the Jakes' model, it is given as:

$$r_{hh}(l) = J_0(2\pi f_d T_s l), \quad (6)$$

where $J_0(\cdot)$ represents the zeroth order bessel function of the first kind. The $(F_{\mathrm{W}}^{[\mathrm{itr}]} \times 1)$ cross covariance vector $\mathbf{C}_{zh}$ captures the influence of the closest $F_{\mathrm{W}}^{[\mathrm{itr}]}$ observations on the channel at time index $k$. It is given as

$$\mathbf{C}_{zh} = \begin{pmatrix} r_{hh}(-\lceil F_{\mathrm{W}}^{[\mathrm{itr}]}/2 - 1\rceil) \\ \vdots \\ r_{hh}(\lfloor F_{\mathrm{W}}^{[\mathrm{itr}]}/2\rfloor) \end{pmatrix}. \tag{7}$$

## IV. Consideration of Different Positions of the Systematic Symbols

When considering a Turbo-coded system, the grouping of bits to symbols and the transmission order of the data symbols (in our system model this functionality is covered by the 'Mapper') is of significance for the convergence speed of the code-aided estimator.

Without loss of generality, we assume that an individual symbol $x_k$ either consists solely of systematic bits, solely of parity bits of the first encoder or solely of parity bits of the second encoder. Due to the principle of Turbo decoding, the reliability of the systematic bits is crucial for the convergence speed of the decoding process, whereas the reliability of the parity bits is crucial for the asymptotic performance. This fact has been made use of in the context of puncturing in e.g. [10].

As the main issue of code-aided synchronization is the joint convergence of the decoding and estimation process, we here successfully make use of this property and significantly improve the convergence speed by better protecting the systematic bits against distortions due to the initial channel estimation error. This is feasible, as the reliability of both parity and systematic bits improves during the channel estimation/decoding process due to the continuing improvement of the channel estimate. Therefore, good asymptotic behavior can be guaranteed nevertheless.

This principle has already been discussed in [11] in the context of carrier frequency estimation. There, the improvement that can be obtained by carefully placing the systematic symbols is rather slight. For the case of (code-aided) LMMSE channel estimation the gain is far more significant, if the channel sampling is below the Nyquist rate of the channel fading process, i.e.

$$P_S > 1/(2 \cdot F_d). \tag{8}$$

Let us first consider the case, where the pilot spacing is small enough, so that the channel is sampled with Nyquist rate. For this case, it is shown in [12] that the channel estimation error is independent of the time index $k$. Therefore, the distortion on the received samples after initial channel estimation is comparable, irrespective of the symbol position within the burst. Hence, there are no particularly well-protected positions within one burst. If the pilot spacing is increased beyond Nyquist rate, i.e. $P_S > 1/(2 \cdot F_d)$, this is no longer the case. It is then impossible to acquire a sufficient statistic of the fading process in the first iteration and the expected value of the channel estimation error then depends on the data symbol position, cf. Fig. 2. The channel estimation error is low in the vicinity of the pilot symbols, whereas it is rather high in the middle between two pilot symbols. The difference in the protection level increases, the higher the product $2 \cdot P_S \cdot F_d$ increases beyond 1.

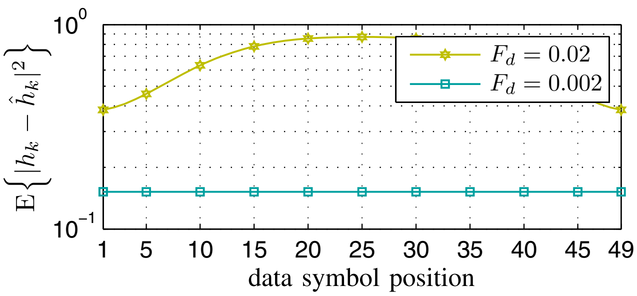
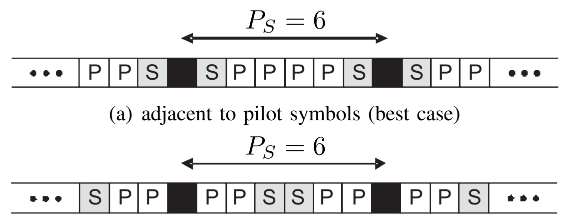



Figure 2. Analytical MSEE for PSAM channel estimation vs. data symbol position; $P_S = 50$, $E_s/N_0 = 5\,\mathrm{dB}$, $F_{\mathrm{W}}^{[\mathrm{init}]} = 10$




(a) adjacent to pilot symbols (best case)

(b) away from pilot symbols (worst case)

Figure 3. Different Positions of Syst. Symb. for a Rate-1/3 Turbo Code $\boxed{\mathrm{S}}$: systematic symbol, $\boxed{\mathrm{P}}$: parity symbols

Making use of the afore-described property of Turbo codes, we conjecture that transmitting systematic symbols adjacent to pilot symbols improves the convergence properties of code-aided LMMSE channel estimation, if the inverse of the pilot spacing equals less than the Nyquist rate of the channel fading process (cf. Fig. 3). It should be stressed again that this mapping strategy does not entail any degradations, when the asymptotic system behavior is considered. This is due to the fact that in the case of convergence, each data symbol can function as a pilot symbol, making the residual channel estimation error completely independent of the data symbol position (for $F_d < 0.5$).

## V. Simulation Results

In this section, we evaluate the mapping strategy described in the previous section by means of semi-analytic method and by means of Monte-Carlo BER simulation. We here consider the scenario of an 8PSK modulation with Gray Mapping. The component codes of the Turbo decoder are recursive convolutional codes with the generator polynomial $1 + D^2$ and the feedback-polynomial $1 + D + D^2$. The correlation properties of the fading process are determined by the Jakes' spectrum with normalized maximum Doppler shift set to $F_d = 0.02$ and 5 Turbo decoding iterations are performed per channel estimation iteration. The filter length for initial (PSAM) channel estimation is set to $F_{\mathrm{W}}^{[\mathrm{init}]} = 10$, whereas the filter

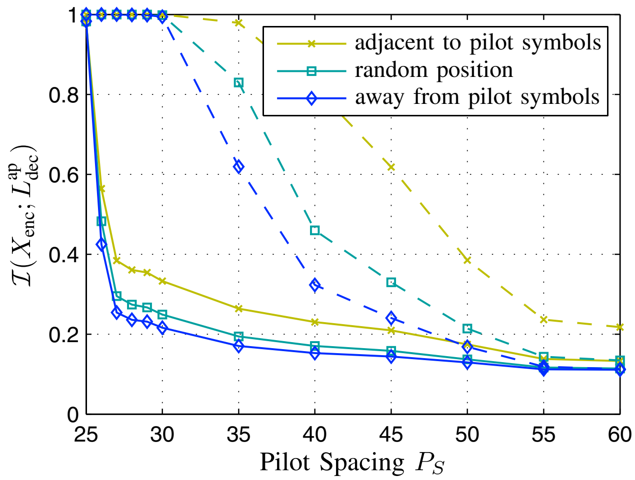



Figure 4. $\mathcal{I}(X_{\text{enc}}; L_{\text{dec}}^{\text{ap}})$ vs. $P_S$ for diff. positions of the systematic symb., $N = 250002$, $F_d = 0.02$, $E_b/N_0 = 7.5\,\text{dB}$, 8PSK Gray, $\mathcal{C} = \{1, 5/7\}_8$, 5 Dec. Iterations, $F_{\text{W}}^{[\text{init}]} = 10$, $F_{\text{W}}^{[\text{itr}]} = 100$ (———) after 1st iteration; (– – –) after 2nd iteration.

length for iterative channel estimation is set to $F_{\text{W}}^{[\text{itr}]} = 100$.

We first consider the benefit with respect to the mutual information $\mathcal{I}(X_{\text{enc}}; L_{\text{dec}}^{\text{ap}})$ between the encoded bits $X_{\text{enc}}$ and the a posteriori log-likelihood ratios (LLR) of these bits at the Turbo decoder output $L_{\text{dec}}^{\text{ap}}$ after a specific number of Turbo decoding iterations (here: 5 Turbo decoding iterations). This measure provides a valuable description of the convergence behavior of the system (cf. [13]).

The average mutual information for one frame can be obtained as:

$$\mathcal{I}(X_{\text{enc}}; L_{\text{dec}}^{\text{ap}}) = 1 - \frac{1}{\text{ld}(M)K} \sum_{k=1}^{K} \sum_{i=1}^{\text{ld}(M)} \text{ld}\left[1 + \exp(-(L_{\text{dec}}^{\text{ap}})_k^i \cdot (X_{\text{enc}})_k^i)\right], \tag{9}$$

where $\text{ld}(\cdot)$ is the binary logarithm. $(X_{\text{enc}})_k^i$ denotes the bipolar coded bit that was mapped onto the $i$-th bit position on the $k$-th symbol and $(L_{\text{dec}}^{\text{ap}})_k^i$ represents the corresponding a posteriori LLR. We choose this measure, as it is reliability information on the coded bits that is exchanged during code-aided synchronization. The results are depicted in Fig. 4. We here consider three options of placing the systematic symbols: adjacent to the pilot symbols (cross markers) (cf. Fig. 3 (a)), as far away as possible from the pilot symbols (diamond markers) (cf. Fig. 3 (b)), and random positioning (rectangular markers). Since the maximum normalized Doppler frequency is set to $F_d = 0.02$, the pilot spacing $P_S = 25$ still obeys the Nyquist rate. The solid lines represent the mutual information at the decoder output after initial channel estimation that is solely based on pilot symbols, whereas the dashed lines show the mutual information at the decoder output after the second iteration, where already one code-aided channel estimation has been carried out.

It can be seen in Fig. 4 that the benefit in choosing well protected positions for the systematic bits only exists if the pilot spacing $P_S$ is larger than $1/(2 \cdot F_d)$. It also becomes obvious that the convergence speed of code-aided channel estimation increases with increasing accuracy of the initialization. E.g. for $P_S < 30$, convergence already seems to be obtained after the second channel estimation procedure – irrespective of the mapping scheme. Furthermore, we note that the larger the pilot spacing, the larger is the relative gain of the proposed mapping scheme over the conventional scheme (random positions).

Fig. 5 depicts the BER versus the SNR for code-aided iterative synchronization. The dashed curve corresponds to the BER result after (initial) PSAM channel estimation. Subfigures (a) – (c) correspond to the result after 5, 10 and 15 channel estimation iterations. First, it can be seen that if one decides to choose a rather high pilot spacing (here $P_S = 50$) in order to guarantee a high spectral efficiency, the improvements in terms of the SNR gain that can be achieved by code-aided channel estimation are significant. Furthermore, we see that carefully placing the systematic symbols significantly accelerates the convergence speed of code-aided channel estimation. It should be stressed that, we do not claim asymptotic gains for the proposed mapping scheme, but solely faster convergence.

## VI. Conclusion

In this paper, it is pointed out that in a Turbo-coded system, systematic bits have a larger impact on the convergence properties of code-aided channel estimation than parity bits. Systematic bits should, therefore, be grouped into separate symbols and be transmitted at positions in the burst that suffer from the least distortion due to channel estimation inaccuracies. For LMMSE channel estimation, these positions are well known in advance, and do only exist if the channel is not sampled with sufficiently dense pilot symbols. It is demonstrated that in this scenario, carefully choosing the mapping strategy leads to a rather significant gain in terms of the convergence speed.

## References


[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. of IEEE International Conference on Communications (ICC)*, Geneva, Switzerland, May 1993.

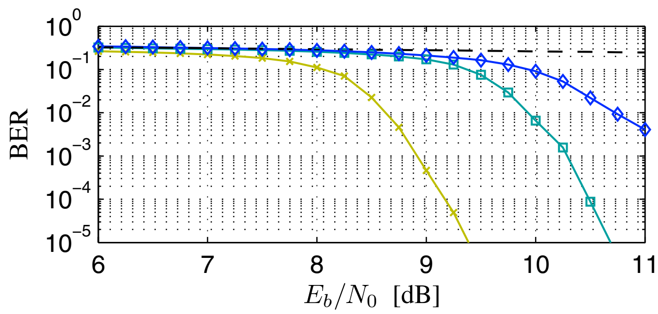
[2] C. Komninakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 9, pp. 1706–1717, Sept. 2001.

[3] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "Turbo synchronization: An EM algorithm interpretation," in *Proc. of IEEE International Conference on Communications (ICC)*, Anchorage, Alaska, USA, May 2003.

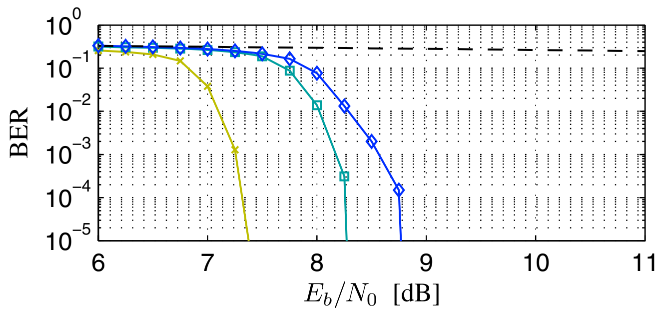
[4] C. Herzet, "Code-aided synchronization for digital burst communications," Ph.D. dissertation, Communications Laboratory, Université catholique de Louvain, April 2006.

[5] C. Herzet, H. Wymeersch, M. Moeneclaey, and L. Vandendorpe, "On maximum-likelihood timing estimation," *IEEE Transactions on Communications*, vol. 55, no. 6, pp. 1116–1119, June 2007.

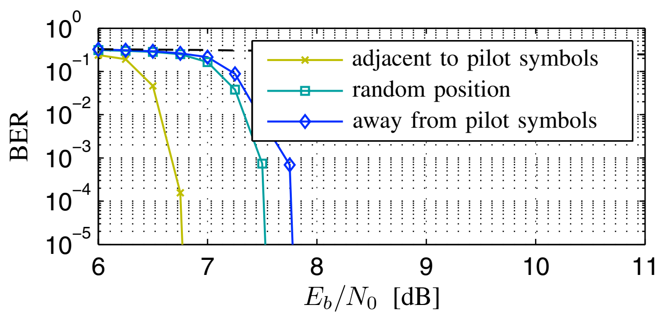
[6] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall PTR, 1993.




(a) 5 Channel Estimation Iterations




(b) 10 Channel Estimation Iterations




(c) 15 Channel Estimation Iterations

Figure 5. BER vs. $E_b/N_0$ for diff. positions of the systematic symb., $N = 250002$, $F_d = 0.02$, $P_S = 50$, 8PSK Gray, $\mathcal{C} = \{1, 5/7\}_8$, 5 Dec. Iterations, $F_W^{[\mathrm{init}]} = 10$, $F_W^{[\mathrm{itr}]} = 100$, (— — —) PSAM Channel Estimation.


[7] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[8] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*, 1st ed. New York, NY: John Wiley & Sons, 1998.

[9] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted Turbo codes over flat-fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 9, pp. 1697–1705, Sep. 2001.

[10] F. Babich, G. Montorsi, and F. Vatta, "Design of rate-compatible punctured serial concatenated convolutional codes," in *Proc. of IEEE International Conference on Communications (ICC)*, June 2004.

[11] S. Godtmann, A. Pollok, N. Hadaschik, G. Ascheid, and H. Meyr, "On the influence of pilot symbol and data symbol positioning on Turbo synchronization," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Dublin, Ireland, 2007.

[12] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2115–2123, Dec. 2001.

[13] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.