
© 2001 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Blind Equalization with Iterative Joint Channel and Data Estimation for Wireless DPSK Systems

Xiao-ming Chen† and Peter A. Hoeher
Information and Coding Theory Lab
Faculty of Engineering, University of Kiel
Kaisersstr. 2, D-24143 Kiel, Germany
E-mail: {xc,ph}@tf.uni-kiel.de

Abstract—A joint channel and data estimator suitable for wireless differential phase shift keying (DPSK) systems with short block sizes is proposed. No training data symbols are necessary, but a priori information may be incorporated, if available. The basic structure consists of an iterative approximation of the joint maximum-likelihood (ML) solution, which is enhanced by adaptive filtering, by shift operations, by coefficient annealing, as well as by using a priori information. To our best knowledge, this is the first time that (i) a priori information of the data has been exploited in the context of blind equalization and that (ii) the DPSK/ISI super trellis has been explicitly exploited for blind equalization rather than doing differential decoding for resolving the problem of phase ambiguity.

Keywords—Blind equalization and channel estimation, joint equalization and channel estimation, maximum-likelihood sequence estimation, iterative least-squares processing

I. INTRODUCTION

In digital mobile communications, the two main distortions are intersymbol interference (ISI) caused by time dispersion of the multipath channel, and Doppler spread due to the movement of the mobile receiver. To combat ISI, a variety of equalization algorithms have been developed in the last decades [1], [2]. An obvious problem coupled with wireless equalization is channel estimation.

Nowadays, in typical digital mobile communication systems a training sequence is inserted in each data block for the purpose of channel estimation. For more efficient usage of bandwidth, blind equalization attracts more and more attention. Furthermore, blind detection schemes may be embedded in existing systems as an add-on in order to improve the system performance in difficult environments. Blind equalization techniques may be classified in three categories.

The first category is based on statistical properties of the received signal, which is either sampled at a high rate or oversampled. In the former case, it is necessary to exploit high order statistics [3], [4]. In the latter case, second order statistics are sufficient because fractionally-spaced channel outputs are cyclostationary [5]. Possible problems with these types of statistical approaches are, depending on the algorithm, a slow convergence rate, a possible convergence to local minima, and a lack of robustness against Doppler spread, noise, and interference. On the other hand, dramatic improvements were reported in the last few years, particularly for mobile radio applications [6], [7].

The second category exploits the algebraic structures of oversampled systems, which can be modeled as multichannel systems [5]. In a so-called deterministic approach, the channel impulse responses (CIRs) of the polyphase channels are estimated by solely exploiting the channel outputs rather than estimating statistical terms. The cross correlation between different channel outputs and corresponding CIRs is used to form a set of linear equations [8]. This technique is closely connected to subspace methods for blind equalization [9]. However, these methods have constraints on the channel matrix; conditions on the identifiability of channels are given in [8].

The third category covers trellis-based methods. Classically, an adaptive channel estimator (e.g. based on the LMS or RLS algorithm) is implemented in parallel to a trellis-based equalizer (e.g. a ML sequence estimator). The most likely survivor, given a suitable decision delay, is fed into the channel estimator, i.e., this approach estimates the channel and the data jointly [1]. A further development was the introduction of per-survivor-processing technique [10], where each survivor employs its own channel estimator and no tentative decision delay is afforded. In order to exploit properties of the multipath fading channel and to track the time varying channel, model-fitting algorithms were used in [11], [12], [13], for example. The ML joint data and channel estimator in the absence of training (as well as suboptimum variations thereof) was presented in [14], [15]. While all of these trellis-based techniques are applied for coherent detection, non-coherent blind equalization techniques are an interesting alternative area of current research [16], [17], [18], [19].

Our focus is on trellis-based algorithms, suitable for short block sizes and noisy environments. This goal is approached by an iterative approximation of the ML joint data and channel estimator [14]. Significant improvements are particularly obtained by incorporating adaptive channel estimation [20], [1] into the equalizer, but also by shift operations, by coefficient annealing, and by using a priori information, if available. The possibility of using a priori information of the data symbols (e.g. provided by an outer soft-output channel decoder) is a nice feature of trellis-based algorithms. Training corresponds to a priori information only, whereas a (truly) blind mode corresponds to no a priori information. To our best knowledge, a priori information has not been applied in the context of

†This work is supported by Germany Science Foundation (DFG) under Grant No. Ho 2226/1.
blind equalization yet. As opposed to the common understanding that differential precoding is used to resolve the phase ambiguity both in channel and data estimation, we explicitly use the structure of the DPSK/ISI super trellis.

In Section II, we describe the system model under investigation and formulate the joint estimation problem. The proposed receiver structure is presented in Section III, which also discusses the distinct novel features. Simulation results are provided in Section IV. Finally, conclusions are drawn in Section V.

II. System Model

Throughout this paper we use the complex baseband notation.

A. Transmitter

We consider an uncoded M-ary DPSK system. The output symbols of the differential encoder can be written as

$$x(k) = x(k-1) \cdot d(k), \quad x(0) = +1,$$  (1)

where $d(k)$, $k \geq 0$, are $M$-ary PSK symbols and $x(0)$ serves as a reference symbol.

B. Channel Model

The pulse shaping filter, multipath fading channel, receiving filter, and the sampling can be represented by a tapped delay-line baud-rate model [1]. (We restrict ourselves to baud-rate sampling. An extension to fractionally-spaced sampling is straightforward.) For simplicity, the channel is assumed to be time-invariant within a data block. (The proposed receiver is able to track a time-varying channel, however.) The corresponding channel output can be written as

$$y(k) = \sum_{l=0}^{L} x(k-l) \cdot h_l + n(k),$$  (2)

where $\{h_l\}$, $0 \leq l \leq L$, are the coefficients of the equivalent discrete-time channel model (in the following referred to as the channel coefficients), $L$ is the effective channel memory (after suitable truncation), and $\{n(k)\}$ is a sequence of additive noise samples. In the following, $\{n(k)\}$ is assumed to be an i.i.d additive white Gaussian noise sequence with one-sided power spectral density $N_0$. $L$ is assumed to be known at the receiver.

In most mobile communication systems the data symbols are transmitted block by block. If we assume $K$ data symbols per block (excluding the reference symbol), we can write (2) in matrix/vector notation as:

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{h} + \mathbf{n},$$  (3)

where

$$\mathbf{y} = [y(0), y(1), ..., y(K)]^T,$$
$$\mathbf{h} = [h_0, h_1, ..., h_L]^T,$$
$$\mathbf{n} = [n(0), n(1), ..., n(K)]^T,$$

$$\mathbf{X} = \begin{bmatrix}
  x(0) & 0 & \cdots & 0 \\
  x(1) & x(0) & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  x(L) & x(L-1) & \cdots & x(0) \\
  \vdots & \vdots & \ddots & \vdots \\
  x(K) & x(K-1) & \cdots & x(K-L)
\end{bmatrix}. \quad (4)$$

(If the data symbols are not transmitted on a block-by-block basis, if the block size is large, or if the channel is fastly time-varying, $K$ may define a subblock.)

C. Receiver

The task of the receiver is twofold: Primarily, we are interested in an estimate of the data vector $\mathbf{d} = [d(1), \ldots, d(K)]^T$. A coherent receiver must also obtain estimates of each element of $\mathbf{h}$ in amplitude and phase.

In a coherent DPSK receiver, joint channel and data estimation may be based on the ISI trellis (followed by differential decoding), or may be based on the super trellis, which combines the differential encoder and the ISI trellis. In the following, we operate on the DPSK/ISI super trellis, unless otherwise stated. Note that the DPSK/ISI super trellis and the ISI trellis have the same number of states!

Denoting $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{h}}$ as hypotheses for the data vector and the channel coefficient vector, respectively, we can formulate the joint channel/data estimation problem according to the ML-criterion in the presence of AWGN as

$$\left(\tilde{\mathbf{x}}, \tilde{\mathbf{h}}\right) = \arg \max_{\mathbf{x}, \mathbf{h}} \left\{ p \left( \mathbf{y} | \mathbf{x}, \mathbf{h} \right) \right\}$$
$$= \arg \min_{\mathbf{x}, \mathbf{h}} \sum_{k=0}^{K} \left| y(k) - \sum_{l=0}^{L} \tilde{x}(k-l) \cdot \tilde{h}_l \right|^2$$
$$= \arg \min_{\mathbf{x}, \mathbf{h}} \| \mathbf{y} - \tilde{\mathbf{X}} \cdot \tilde{\mathbf{h}} \|^2_F,$$  (5)

where we assume that the noise variance is constant within a block and where we (at the moment) assume that no a priori information about the data sequence is available. Here, $p \left( \mathbf{y} | \mathbf{x}, \mathbf{h} \right)$ is the pdf of the received vector conditioned on data and channel hypotheses, and $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{h}}$ are the corresponding estimates. $\| \cdot \|^2_F$ is the Frobenius norm with $\| \mathbf{A} \|^2_F = \text{tr}(\mathbf{A}^H \cdot \mathbf{A})$, which is equivalent with the Euclidean distance if $\mathbf{A}$ is a column vector. Given no a priori information about the channel coefficients, the optimal solution for the joint estimation problem (5) is a least-squares channel estimator for each possible hypothesis $\tilde{\mathbf{x}}$.

The ML-estimate is the pair $(\tilde{\mathbf{x}}, \tilde{\mathbf{h}})$ which minimizes the Euclidean distance. The complexity of this exhaustive search approach inhibits its practical applications, however.

The joint minimization problem in (5) can also be for-
mulated as a two-step minimization:

\[
\left( \hat{\chi}, \hat{h} \right) = \arg \left\{ \min_{\hat{\chi}} \left[ \min_{\hat{h}} \| y - \hat{\chi} \cdot \hat{h} \|_F^2 \right] \right\}.
\]

(6)

Equation (6) motivates an intuitive iterative procedure, where data estimation is performed given channel estimates from the last iteration:

\[
\hat{x}^{(i+1)} = \arg \min_{\hat{\chi}} \| y - \hat{\chi} \cdot \hat{h}^{(i)} \|_F^2,
\]

and where afterwards the channel coefficients are evaluated based on the data estimates from (7):

\[
\hat{h}^{(i+1)} = \arg \min_{\hat{h}} \| y - \hat{x}^{(i+1)} \cdot \hat{h} \|_F^2.
\]

(8)

This procedure may be continued until a convergence is observed:

\[
\hat{h}^{(i+1)} = \hat{h}^{(i)}
\]

(9)

Note, however, that this two-step iterative scheme is suboptimal since not all hypotheses are tested. Traditionally, (7) is solved by means of the Viterbi algorithm (VA) and (8) is solved by means of LS channel estimation:

\[
\hat{h}^{(i+1)} = \left( \hat{x}^{(i+1)} \cdot \hat{x}^{(i+1)} \right)^{-1} \cdot \hat{x}^{(i+1)} \cdot y.
\]

(10)

This two-step iterative method is referred to as the conventional iterative scheme in the following. Given initial channel estimates obtained by training, this technique has been investigated in [21]. Excellent results have been reported in this context. In the context of blind equalization, however, we observed that the convergence speed is much slower and that the possibility of a convergence to local minima is much higher.

III. PROPOSED RECEIVER STRUCTURE

A. Novel Features

In order to improve the system performance with respect to performance and complexity, we propose to solve (7) and (8) in just one step. The main features are as follows:

- We incorporate an adaptive channel estimator within a trellis-based equalizer operating on the super trellis. Suitable channel estimators are based on the LMS algorithm or the RLS algorithm, which is faster than the LMS algorithm at the price of additional complexity [20]. Suitable equalizers are sequence oriented, such as the VA, reduced-state sequence estimators (RSSE), or the MaxLog-APP algorithm [23]. At the moment, we assume a ML sequence estimator. The main advantage of the joint equalizer/channel estimator is that the convergence rate (i.e., the acquisition behavior) improves significantly compared to the conventional iterative scheme. Simultaneously, the adaptive channel estimator is capable to track time-varying channels.

- We tackle the problem of local minima. An important observation is as follows: Typically, blind equalizers are only able to provide estimates of \( x(k-l) \cdot h_l \) for all \( 0 \leq l \leq L \). In the absence of decision errors, in the binary case this implies that \( \hat{x}(k) = \pm \hat{x}(k) \cdot h_\kappa \), where \( \kappa = 0, \pm 1, \pm 2, \ldots \). The sign ambiguity is well known as the phase ambiguity problem, which is solved by differential precoding. The shift ambiguity appears to be little known in the literature, however. Since the corresponding channel estimates are \( \hat{h}_l = \pm \hat{h}_{l+1} \), any \( \kappa \neq 0 \) corresponds to a local minimum. Besides this class of minima, other local minima exists as well.

The main problem related to the shift ambiguity is that channel coefficients are shifted out of the observation interval \( L+1 \). This problem is easily solved by extending the observation interval and by shifting the estimated impulse so that the energy within the observation interval \( L+1 \) is maximized or that the signal/noise ratio is maximized. Secondly, to get rid of local minima, we introduce a simple coefficient annealing technique, which is different from simulated annealing [22] in that this annealing technique is performed in a more deterministic sense and that it is much simpler than simulated annealing.

- We take a priori information into account, if available. This also helps to solve the problem of local minima.

B. Proposed Algorithm

In the following, \( \xi^{(i)}(k) \) denotes a value at time \( k \) in the \( i \)th iteration, where \( \xi \) may be a scalar, a vector, or a matrix. With this notation, a formal description of the proposed receiver is as follows:

1. Initialization:

   - In the absence of training data, the channel coefficient vector is initialized by a fixed vector, e.g., with elements \( \hat{h}_l^{(i)}(0) = 1/\sqrt{L+1} \), \( l = 0, \cdots, L' \), where \( L' \geq L \) is an expanded channel order used in sequence and channel estimation. In case that the RLS algorithm is applied for adaptive channel estimation, the inverse correlation matrix is initialized as \( P^{(i)}(0) = \delta^{-1} \cdot I \), where \( I \) is an \( L' + 1 \) by \( L' + 1 \) identity matrix and \( \delta \) is a small positive constant.

2. Adaptive equalization:

   - An adaptive channel estimator is incorporated in the data sequence estimator. The equalizer delivers tentative decisions to the channel estimator. In case of RLS adaptation, the \( i \)th iteration, \( 1 \leq i \leq N_{iter} \), can be written as [20]

   \[
ge^{(i)}(k) = y(k) - \hat{h}^{(i)} \cdot T(k) \cdot \hat{x}^{(i)}(k),
\]

   (12)

   \[
\hat{h}^{(i)}(k) = \hat{h}^{(i)}(k-1) + g^{(i)}(k) \cdot e^{(i)}(k),
\]

   (13)

   \[
P^{(i)}(k) = \beta^{-1} \cdot P^{(i)}(k-1) - g^{(i)}(k) \cdot \hat{x}^{(i)}(k) \cdot T(k) \cdot P^{(i)}(k-1),
\]

   (14)

   with \( g^{(i)}(k), \hat{x}^{(i)}(k), \hat{h}^{(i)}(k) \in G^{(L'+1)} \)

   \[
P^{(i)}(k-1) \in G^{(L'+1) \times (L'+1)}.
\]
where \( \mathbf{g}^{(i)}(k) \) is the gain vector at time \( k \), \( P^{(i)}(k) \) is the inverse correlation matrix at time \( k \), \( \beta \) is the forgetting factor, and \( 1 < \alpha < K \). \( \mathbf{x}^{(i)}(k) \), \( \mathbf{h}^{(i)}(k) \) and \( \mathbf{c}^{(i)}(k) \) are the tentative data vector (which is obtained by tracing back the trolis from the best state at time \( k \) given a certain decision delay), the estimated channel coefficient vector and the corresponding a priori estimation error, respectively. In case of LMS channel estimation, the gain vector is replaced by a scalar.

At the end of the block \( (k = K) \), the most likely path is traced back to obtain an estimate \( \hat{\mathbf{x}}^{(i)} \) of the data vector. The last estimated channel coefficient vector is declared as a tentative channel estimate: \( \hat{\mathbf{h}}^{(i)}(K) \).

Processing of step 2 is repeated until a convergence is observed, compare (9), or the maximum number of iterations, \( N_{\text{iter}} \), is reached.

3. Processing of the estimated channel coefficients: If a convergence is observed in step 2, i.e., \( \hat{\mathbf{h}}^{(i)} = \hat{\mathbf{h}}^{(i-1)} \), firstly compute the squared error:

\[
\gamma^{(i)} = \| \mathbf{y} - \hat{\mathbf{x}}^{(i)} \cdot \mathbf{h}^{(i)} \|_F^2.
\]

By minimizing the energy of the estimated channel coefficients inside a sliding window of length \( L + 1 \), we can mitigate the shift ambiguity problem:

\[
\alpha = \arg \max_{\tilde{\alpha}} \sum_{l=\tilde{\alpha}}^{L+\tilde{\alpha}} | \hat{h}^{(i)}_l |^2, \quad 0 \leq \tilde{\alpha} \leq L'.
\]

Then, a new squared error \( \gamma_{\alpha}^{(i)} \) is computed with \( \hat{\mathbf{h}}_{\alpha}^{(i)} = [\hat{h}^{(i)}_{\alpha}, \ldots, \hat{h}^{(i)}_{\alpha+L}, 0, \ldots, 0]^T \) and the corresponding shifted data vector \( \hat{\mathbf{x}}^{(i)}_{\alpha} \) using (15). After comparing two squared errors \( \gamma^{(i)} \) and \( \gamma_{\alpha}^{(i)} \), the channel and data estimates with smaller error are stored.

After that, coefficient annealing is performed as follows: The proposed annealing technique is based on the sensitivity of the squared error w.r.t. the \( l \)-th estimated channel coefficient. If the sensitivity is large, then the change of this coefficient is relative small, otherwise the change should be large. Let us denote the estimated channel and data vectors corresponding to the smaller squared error as \( \hat{\mathbf{h}}^{(i)} \) and \( \hat{\mathbf{x}}^{(i)} \), respectively. The sensitivities are computed as:

\[
s_j^{(i)}(k) = \frac{\partial \gamma^{(i)}}{\partial \hat{h}_j^{(i)}}, \quad 0 \leq j \leq L'.
\]

\[
= -\sum_{k=0}^{K} \left( y(k) - \sum_{l=0}^{L} \hat{z}^{(i)}(k-l) \cdot \hat{h}_l^{(i)} \right) \cdot \hat{z}^{(i)}(k-j).
\]

Then, coefficient annealing is performed using a linear function (other functions are also possible):

\[
Re\{\hat{h}_l^{(i)}\} = Re\{\hat{z}^{(i)}\} + \frac{k_2 \cdot s_{\text{max}} - k_1 \cdot s_{\text{min}} + (k_1 - k_2) \cdot Re\{s_l^{(i)}\}}{s_{\text{max}} - s_{\text{min}}},
\]

where \( s_{\text{min}} \) and \( s_{\text{max}} \) are the minimum and maximum amplitudes of the real part of the calculated sensitivities, respectively. \( k_2 \) and \( k_1 \) determine the range of absolute increment. Note that \( v \in \{ \pm 1 \} \) is a random variable, which determines the direction for annealing. The same procedure should be carried out for the imaginary part, which is omitted here for simplicity.

**Else** go to 4.

4. Test of stopping criteria:

**If** the maximum number of iterations is reached, i.e., \( i = N_{\text{iter}} \), go to step 5.

**Else** initialize \( \hat{\mathbf{h}}^{(i+1)}(0) = \hat{\mathbf{h}}^{(i)}(0) = \delta^1 \cdot \mathbf{I} \), and go to step 2.

5. Final data estimation:

**If** a convergence happened during step 2, select the channel coefficients with minimum \( \lambda \) as the channel estimate and the corresponding data sequence as the final data estimate.

**Else** the last estimated channel coefficients are selected as the channel estimate and the corresponding data sequence as the final data estimate.

If processing is done in the DPSK/ISI super trellis, there is a one-to-one mapping between the differentially encoded data sequence and the info sequence. Otherwise, differential decoding has to be performed.

As mentioned before, a nice feature of trellis-based blind equalization is the possibility to make use of a priori information. If a priori information is available, we replace the ML sequence estimator assumed so far by a MAP sequence estimator, i.e., the branch metrics are modified as [24], [23]:

\[
\gamma (S^{<i,j>}(k) \rightarrow S^{<j,i>}(k)) = -\frac{1}{2\sigma_n^2} \left[ y(k) - \sum_{l=0}^{L} x^{<i-j>}(k-l) \cdot \hat{h}_l(k-1) \right]^2 + \log p(d^{<i-j>}(k)),
\]

where \( (S^{<i,j>}(k-1) \rightarrow S^{<j,i>}(k)) \) denotes a state transition, and \( \sigma_n^2 \) is the noise variance.

For binary systems, (19) can be rewritten as:

\[
\gamma (S^{<i,j>}(k-1) \rightarrow S^{<j,i>}(k)) = -\frac{1}{2\sigma_n^2} \left[ y(k) - \sum_{l=0}^{L} x^{<i-j>}(k-l) \cdot \hat{h}_l(k-1) \right]^2 + \frac{1}{2} L (d(k)) \cdot d^{<i-j>}(k),
\]

where \( L (d(k)) \) denotes the given log-likelihood ratio (LLR) of symbol \( d(k) \), before differential encoding. (Symbol-by-symbol MAP estimation is not recommendable here due to the lack of survivors; the locally best surviving paths are necessary for channel estimation.) The significance of (20) is a generic receiver structure, which is the same for the wide range of blind equalization without a priori information (where \( L (d(k)) = 0 \) for all \( k \)) to a training-based equalizer (where \( |L (d(k))| \to \infty \) for some \( k \)).
IV. Simulation Results

The performance of the proposed receiver has been tested and optimized by Monte Carlo simulations. In order to emulate a system which is similar to the GSM system, we investigated a binary DPSK system with a block length of $K = 148$ symbols. Within this paper, the channels under consideration are the two worst case static channels reported in [1] for $L = 2$ and $L = 3$, respectively:

$$
\begin{align*}
\mathbf{h}_a &= [0.5, 0.707, 0.5]^T; \\
\mathbf{h}_b &= [0.38, 0.6, 0.6, 0.38]^T.
\end{align*}
$$

The following design parameters were chosen:
- expanded channel order: $L' = 5$;
- RLS adaptation: $\delta = 0.02$ and $\beta = 0.95$;
- coefficient annealing: $k_2 = 0.3$ and $k_1 = 0.8$;
- equalization by means of a VA with $2L'$ states;
- decision delay in VA: 5 symbols.

To investigate the effect of a priori information, for simplicity we have chosen a constant LLR $|L(d(k))| = W$ in (20), i.e., $L(d(k)) = \text{sgn}(d(k))W$. Simulation results for $N_{iter} = 2$ are plotted in Fig. 1 and Fig. 2 as a function of $W$. Even with just two iterations and no a priori information, the loss is less than 0.7 dB/1.0 dB compared to a known CIR. With more iterations and/or a priori information, the loss can be made arbitrarily small.

![Fig. 1. BER vs. SNR for channel $\mathbf{h}_a$, $N_{iter} = 2$](image1.png)

![Fig. 2. BER vs. SNR for channel $\mathbf{h}_b$, $N_{iter} = 2$](image2.png)

V. Conclusions

A low-cost blind equalizer/channel estimator is proposed which approximates the performance of a blind MLSE [14], [15]. As opposed to the conventional two-step approach with alternating equalization and channel estimation, (i) an adaptive channel estimator is incorporated into a trellis-based equalizer and (ii) techniques to reduce the problem of local minima are proposed. Hence, the number of iterations and the complexity can be significantly reduced. With just two iterations, the gap w.r.t. a reference system with known CIR is less than 1 dB for the channels under investigation. With more a priori information, the gap can be made arbitrarily small. Instead of full-state hard-output equalization treated so far, the VA may be replaced by a reduced-state sequence estimator or by a soft-output equalizer such as the Max-Log-APP equalizer, both of which also use the concept of survivors. The use of a priori information and the possibility to deliver reliability information to subsequent processing stages are powerful features of trellis-based equalizers.

Some interesting observations are as follows:
- Equalization based on the DPSK/ISI super-trellis is as complex as equalization based on the ISI trellis. The number of states is exactly the same and there is a one to one mapping between the states of the super trellis and the states of the ISI trellis.
- Blind equalization and the use of a priori knowledge of the data are not exclusive. In effect, a priori information helps to reduce the problem of local minima and helps to reduce the average number of iterations. The proposed receiver is suitable for the wide range from blind equalization to training based equalization without changing the receiver structure.
- Blind equalization is not only affected by a phase ambiguity, but also by a shift ambiguity.
- Our simulation results demonstrate the strength of blind coherent equalization/channel estimation techniques.

ACKNOWLEDGMENT

We would like to thank Dietmar Stiller, University of Kiel, for preparing initial simulations.

REFERENCES


