Comparison of Receiver Training Methods in Joint Iterative Channel Estimation and Decoding in Flat Fading Channels

Parastoo Sadeghi, Predrag B. Rapajic
School of Electrical Eng. and Telecomm.
University of New South Wales
Sydney NSW 2052 Australia
Email: parastoo@student.unsw.edu.au, p.rapajic@unsw.edu.au

Abstract—**In this paper we compare two different methods for receiver training in flat fading channels. The first method is the traditional way in which periodic training sequences are sent to the receiver (explicit training). In the second method, recently proposed, the information source emits bits with unequal probabilities of being '0' and '1'. This method is called implicit training, since the training is implied in the non-symmetrical source structure. BPSK signaling is used as the simplest example of constant-envelope phase modulations. We map the phase component of the flat fading channel response to a simple two-state Markov model. Then joint iterative trellis-based maximum a posteriori probability (MAP) method is used for channel state estimation and decoding. The results of computer simulations for the receiver Bit Error Rate (BER) performance in various channel fading rates and information rates are presented. The results indicate superior performance of implicit training. In slow fading conditions, the gain is 4 dB at information rate of 0.15 bits/channel use and 1.2 dB for information rate of 0.25 bits/channel use.**

Index Terms—**Flat Fading Channels, MAP Estimation.**

I. INTRODUCTION

Receivers that are designed to work in unknown time-selective correlated fading channels perform channel estimation for reliable communications. The quality of channel estimation is an important factor that affects the receiver performance. Most current mobile communication systems use constant envelope phase modulations. For this class of modulation schemes, estimation of the phase component of the channel response is essential and poses a great challenge for the receiver design, especially at high fading rates. Usually, to achieve reasonable precision in phase estimation, explicit periodic training bits are inserted among channel symbols. These bits are predetermined and convey no detectable information. The main advantage of this method is that during the training period good precision of channel estimation is attainable. Explicit training drawbacks are summarized as:

- The receiver does not get any supervision between the training periods and if the channel state changes during this interval, performance is degraded.
- Better channel estimation precision in the training period does not necessarily result in better channel estimation in the tracking period. In fact we have recently shown that the precision of channel state information degrades exponentially for two-state Markov channels as we move further from the last training bit [10].

Here we compare explicit training with a recently devised method in which the training is hidden or implied in the structure of the information source, i.e. its non-symmetrical bit distribution [5]. The advantages of implicit training are:

- Training is present for every channel symbol.
- Since training does not take extra channel symbols, for fixed end-to-end channel interleaver/deinterleaver delay, larger blocks of data can be transmitted. Equivalently, if the block size is fixed, smaller end-to-end system delay is achievable.
- Channel symbols are used both for channel estimation and data decoding, allowing the two mechanisms benefit from each other.

Motivated by the last feature, we investigate and compare these two methods in joint iterative Maximum a posteriori Probability (MAP) channel estimation and data decoding [6]. This is discussed in detail in the section II. We have performed extensive computer simulations for various fading channel scenarios and information rates. Detailed graphic results of the simulations can be found in section III, that are also summarized here as:

- Joint iterative MAP channel estimation and decoding is an immediate improvement to the method used in [5]. For example, in [5] implicit training outperforms explicit training by 2dB at information rate of 0.15 bits/channel use in slow fading channel ($f_dT = 0.01$). Here the relative improvement is 4 dB that also happens at 5 dB lower Signal to Noise Ratio (SNR). The main reason for this is that in [5] channel estimation result is passed to the decoder in hard estimation format. In the present work, soft estimates about the channel states and decoded bits are passed between the modules in an iterative form. Iterative method takes advantage of the implicit training nature, mentioned above, to the full extent.
- We have extended the results to medium speed fading conditions ($f_dT = 0.05$). The simulation results again indicate better performance of implicit training by 1.5 dB for the information rate of 0.15 bits/channel use.
- We have also extended the results to higher information rate of 0.25 bits/ channel use in the slow fading channel. The

This work was supported by ARC discovery grant scheme.

0-7803-7822-9/03/$17.00

simulation results show about 1.2 dB gain in using implicit training method.

Joint iterative MAP channel estimation and data decoding is based on the Forward-Backward (BCJR) algorithm [1]. Application of BCJR algorithm for channel estimation requires that the fading channel have finite number of Markov states. The mapping of flat fading channels to Finite State Markov Channels (FSMC) has been addressed in [3]-[4] that is also briefly discussed in the following section.

II. CHANNEL MODEL AND MAP ESTIMATION

Here we present discrete-time signals, that are sampled at signaling intervals, as vectors denoted by $\mathbf{w}_n = [w_1, w_2, ..., w_n]$. Let $x(t)$, $y(t)$ and $z(t)$ be the continuous-time channel input, output and receiver noise respectively. We consider flat fading channel model that is of multiplicative nature [8]

$$y(t) = c(t)x(t) + z(t), \tag{1}$$

where $z(t)$ is complex zero-mean white Gaussian process with variance per dimension equal to $N_0 / 2$ and $c(t)$ is also a complex Gaussian process and its phase component has uniform distribution ($U \sim [\,0, 2\pi\,]$).

The correlation of the channel behavior at consecutive signaling intervals is determined by parameter λ that depends on the speed of mobile unit v and carrier frequency f_0 as well as signaling rate and is usually taken to be

$$\lambda = J_0\left(2\pi f_d \,|\, t_1 - t_2 \,|\right) = J_0\left(2\pi \frac{f_0 v}{c} \,|\, t_1 - t_2 \,|\right), \tag{2}$$

where f_d is the Doppler frequency shift and J_0 is the zero-order Bessel function of the first kind. The relationship between channel coherence time T_0 and f_d is approximated by [9]

$$T_0 = \sqrt{\frac{9}{16\pi f_d^2}} = \frac{0.423}{f_d}. \tag{3}$$

The product $f_d \,|\, t_1 - t_2 \,|$ or equivalently the ratio $T_0 / |\, t_1 - t_2 \,|$ is indicative of the channel stability.

In our simulations, we approximate the envelope component of the fading channel to be constant [5], [6]. This simplification does not affect the relative comparison of implicit and explicit training methods and is also justified by the fact that in phase modulation schemes, information is stored in the signal phase. In BPSK signaling, and in the extreme case when noise is negligible, ambiguities in the amplitude of the channel response do not affect receiver performance, whereas phase ambiguities do. Hence the simplified version of (1) is given by

$$y(t) = e^{j\theta(t)}x(t) + z(t). \tag{4}$$

The joint probability of channel phase at two consecutive signaling time intervals is given by [6]

$$\Pr(\theta_1, \theta_2) = \frac{1-\lambda^2}{4\pi^2} \frac{\sqrt{1-B^2} + B\left(\pi - \cos^{-1} B\right)}{\left(1-B^2\right)^{1.5}} \tag{5}$$

where $B = \lambda \cos(\theta_1 - \theta_2)$.

For BPSK signaling, there is nothing more important than to know whether the channel is inverting the signal phase or not. Therefore we propose a simple two-state Markov model for the mapping the phase response of fading channel at the receiver side and derive the FSM channel parameters. After channel phase is quantized and from receiver point of view, the channel is either in inverting state $S = -1$: $\theta \in [\pi/2 : 3\pi/2]$, or non-inverting state $S = 1$: $\theta \in [-\pi/2 : \pi/2]$. The state transition probabilities are computed as

$$P_{ij} = \Pr(S_n = j \,|\, S_{n-1} = i) = 2\int_{D_i}\int_{D_j} \Pr(\theta_1, \theta_2)\, d\theta_1 d\theta_2 \tag{6}$$

We have to emphasize here that FSM mapping of the fading channel is a model at the receiver side used merely for channel estimation purposes "that reduces infinite cardinality of possible channel phases to a finite number of phase states" [6]. However, the actual phase component of the channel response varies in time according to (5).

The system block diagram is shown in Fig. 1. We define the general system information rate as

$$R = \frac{knH(u)}{nN + T}, \tag{7}$$

where k / n is the coder rate, $H(u)$ is the information source entropy, kN is the number input bits in each block, and T is the total number of training bits in each block. In our simulations, we compare implicit and explicit training methods at the same information rate.

When the information bits are passed through the coder, there is no guaranty that the bit distribution at the coder output, $\Pr(x)$, is the same as the bit distribution at the coder input, $\Pr(u)$, especially when the bit distribution is not symmetrical. Convolutional coders add memory and symmetry to the coded bits. This is worst when the coder is recursive. In this case (parity) bits have infinite memory and equal bit probability, independent of the input bit distribution. This prevented us from using turbo codes in the system since non-symmetrical channel bit distribution is a key factor in channel estimation in implicit training. Instead, we use non-recursive non-systematic convolutional coders.

The receiver applies joint MAP state estimation and detection. Except for a few changes, regarding modulation and coder structure, the algorithm is similar to the method in [6]. The receiver performs channel estimation and decoding on separate trellises. This allows the use of channel interleaver/deinterleaver. The branch metric used in the BCJR algorithm for the channel estimator (Θ) module is given by

$$\begin{aligned} \gamma_t^{\Theta}(s', s) &= \Pr\left(y_t, S_t = s \,|\, S_{t-1} = s'\right) \\ &\Pr\left(S_t = s \,|\, S_{t-1} = s'\right)\sum_x \Pr(x_t)\Pr\left(y_t \,|\, S_t = s, x_t\right) \end{aligned} \tag{8}$$

In this equation $\Pr(x)$ is simply channel symbol probability. For the first iteration, a priori knowledge about channel symbol distribution (or coder output distribution) is used that is agreed upon between the transmitter and receiver; for example in the case of explicit training $\Pr(x = 0) = \Pr(x = 1) = 1/2$ for information bits and set accordingly for training bits. In the implicit case, $\Pr(x)$ is empirically determined from coder output bits (that of course depends on the information source probability, $\Pr(u)$, and coder structure). This needs to be done only once and then $\Pr(x)$ for the first iteration is assumed to be known at the receiver throughout the simulation runs. In the next iterations, the $\Pr(x)$ is provided by the soft outputs from the decoder module as shown in the receiver loop in Fig. 1. The decoder, C module, receives deinterleaved version of the soft outputs from the channel state estimator module and computes the branch metric in BCJR algorithm as

$$\gamma_t^C(c',c) = \Pr(\mathbf{y}_t, C_t = c \mid C_{t-1} = c') =$$

$$\Pr(u_t; \angle c' \to c) \sum_s \Pr(s_t) \Pr(\mathbf{y}_t \mid \mathbf{S}_t = \mathbf{s}, \mathbf{x}_t) = \tag{9}$$

$$\Pr(u_t; \angle c' \to c) \prod_{i=1}^{n} \left(\sum_s \Pr(s_{t,i}) \Pr(y_{t,i} \mid S_{t,i} = s, x_{t,i}) \right)$$

where the last equation follows from the assumption of deep channel interleaving so that channel phase becomes independent at consecutive time intervals. The common term in both (8) and (9) is given by

$$\Pr(y \mid S = s, x) \propto \int_{\theta \in s} \exp\left(\frac{-(y_r - x\cos\theta)^2 - (y_i - x\sin\theta)^2}{N_0} \right) \tag{10}$$

where y_r and y_i are the real and imaginary parts of received signal in (4) respectively.

III. Simulation Results

The detail of the system parameters is as follows:

- A rate $k / n = 1/2$ non-recursive non-systematic convolutional coder with memory degree equal m = 6 (g_0 = 133, g_1 = 171) is used for all simulations.
- Modulation type is BPSK.
- Total channel input block length, $nN + T$, is fixed to be 3500. Therefore the overall delay in both training methods is the same. This results in larger N in the case of implicit training (T is smaller for implicit training). A channel block interleaver of size 50 by 70 is used.
- The only channel parameter is fading rate that is either $f_d |t_1 - t_2| = 0.01$ (slow fading) or $f_d |t_1 - t_2| = 0.05$ (medium-speed fading).
- Two system rates are considered. R = 0.15 bits/ channel use and R = 0.25 bits/channel use. Table I shows how coder input block length and information source entropy are chosen to achieve these rates.

Fig. 2 shows the receiver BER performance in slow fading channel at information rate of 0.15 bits/channel use. The number of iterations is 3. For reference, the BER performance when the receiver has perfect knowledge of the channel phase is shown for two different cases. In the first case, information source emits bits with probability $\Pr(u=0) = 0.92755$, as it is the case in implicit training, and in the second case information source is symmetrical. This explains why implicit training has the potential of outperforming explicit training: If the precision of channel phase estimations is the same for both symmetrical and non-symmetrical information sources (as it is the case when the receiver has perfect phase knowledge), non-symmetrical information source achieves better results due to stronger a priori knowledge about information bits.

In short, the key points observed from Fig. 2 are:

- Implicit training results in performance gain of around 4 dB as compared to explicit training.
- Compared to the hard estimation method in [5], the same receiver performance ($P_b = 3^{-5}$) occurs at 5 dB lower SNR. This is due the use of joint iterative soft channel estimation and decoding.

Fig. 3 shows the BER results for the medium-speed fading channel at information rate of 0.15 bits/channel use. The number of iterations is 5. From the figure we observe that

- Implicit training still outperforms explicit training especially at high BER region but as the SNR increases the gap becomes smaller and reaches 1.5 dB. It seems that BER saturation occurs in implicit training.
- The gap between the performance of two-state Markov model and perfect phase situation is wider in the faster fading condition. This indicates the need for more number of states in the FSM model. This is consistent with the results obtained in [4], where moving from 2-state to 4 and 8-state Markov channel results in noticeable capacity improvement.

Fig. 4 presents the simulation results for information rate of 0.25 bits/channel use in slow fading channel. From this figure we observe that

- Implicit training achieves a gain of more than 1.2 dB compared to explicit training.
- Explicit training BER tends to saturate at high SNR conditions.

IV. Conclusions

In this paper we presented and compared two different methods of receiver training in flat fading channels. We used joint MAP channel estimation and decoding and showed through computer simulations the superior performance of implicit training method. We also presented the receiver performance in the case of perfect phase knowledge; this revealed that if the same precision of channel estimation is obtained from both training methods, implicit training will result in overall superior receiver performance, because it still can benefit from non-symmetrical bit distribution of the information source in the decoding process. Implicit training method is a new topic and there are many issues still to be addressed such as comparison of the receiver performance at higher information rates. This is challenging since information source entropy is a nonlinear function of the bit distribution and also the coder affects the input bit distribution in a nonlinear fashion. Another issue to be addressed is the generalization of this method to realistic fading channels with amplitude distortions.

Up to now much effort has been spent on the information source compression and then adding redundancies either in coding or in adding training bits. Here we showed how redundancies in the information source could be used to perform superior channel estimation.

References

[1] L. R. Bahl, J. Coke, F. Jelinek, J. Raviv, "Optimal decoding of linear codes for minimizing symbol error", *IEEE Trans. Inform. Theory*, vol. 20, pp. 284-287, Mar. 1974.

[2] A. Goldsmith, P. Varaiaya, "Capacity, mutual information and coding for finite state Markov channels" ", *IEEE Trans. Inform. Theory*, vol. 42, no 3, pp. 868-886, May. 1996.
[3] H. Wang, N. Moayeri, "Finite state Markov channels-A useful model for radio communication channels", *IEEE Trans. Veh. Technol*, vol. 44, no 1, pp. 163-171, Feb. 1995.
[4] P. Sadeghi, P. B. Rapajic, "Capacity measure for finite state Markov modeling of the phase process in flat fading channels", in *Proc. Int. Conf. on Telecommunications (ICT)*, Papeete, 2003, pp. 572-575.
[5] P. Sadeghi, V. Trajkovic, P. B. Rapajic, "Implicit and explicit receiver training in flat fading channels modeled as finite state Markov processes", in *Proc. Int. Symposium on Information Theory (ISIT)*, Yokohoma, 2003, p 95.
[6] C. Komninakis, R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading", *IEEE J. Selected Areas in Commun.*, vol. 19, no 9, pp. 1706-1717, Sept. 2001 .
[7] H. Kong, E. shwedyk, "Sequence detection and channel estimation over finite state Markov channels ", *IEEE Trans. Veh. Technol*, vol. 48, no 3, pp. 833-839, May. 1999.
[8] J. G. Proakis, *Digital Communications*, 4th ed., New York: Mc Graw Hill, 2000, ch. 14.
[9] T. S. Rappaport, *Wireless Communications Principles and Practice*, 2nd ed., Upper Saddle River: Prentice Hall, 2002, pp. 203-204.
[10] P. Sadeghi, P. Rapajic, "Constrained information capacity of two-state binary Markov channels", submitted to 37th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, 2003.

TABLE I. SYSTEM INFORMATION RATE DETAILS.

	N	T	$H(u)$	$Ratio$
Explicit Rate = 0.15	525	2450	1	$\frac{525}{2*525+2450}$
Explicit Rate = 0.25	875	1750	1	$\frac{875}{2*875+1750}$
Implicit Rate = 0.15	1400	700	0.375	$\frac{1400*0.375}{2*1400+700}$
Implicit Rate = 0.25	1400	700	0.625	$\frac{1400*0.625}{2*1400+700}$

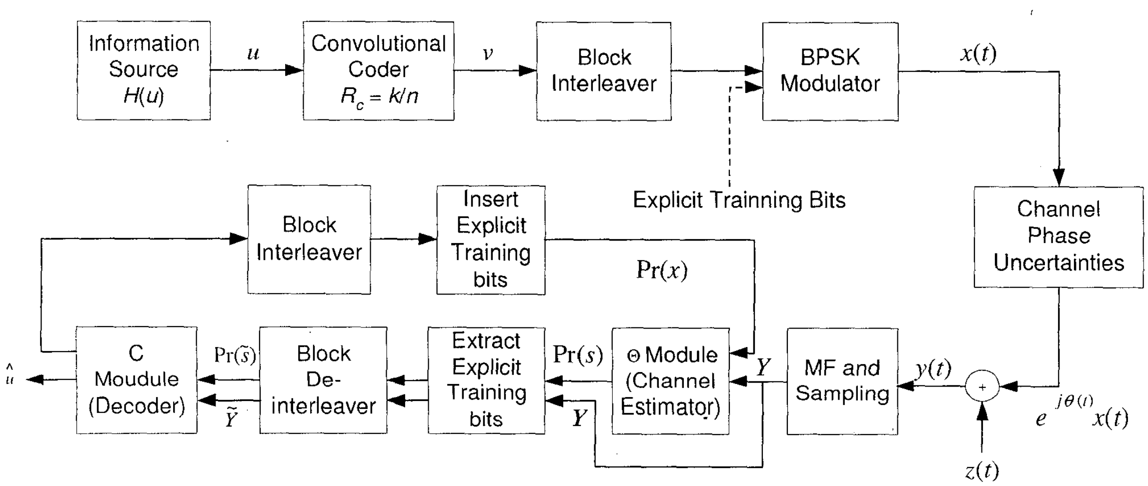

Figure 1. System block diagram.

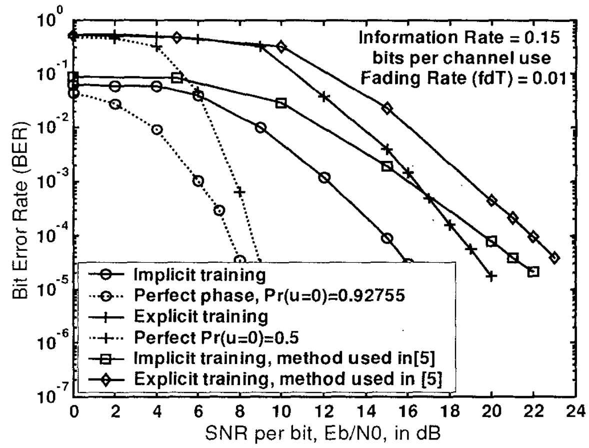

Figure 2. BER performance of implicit and explicit training. Slow fading channel, Information rate = 0.15 bits/channel use.

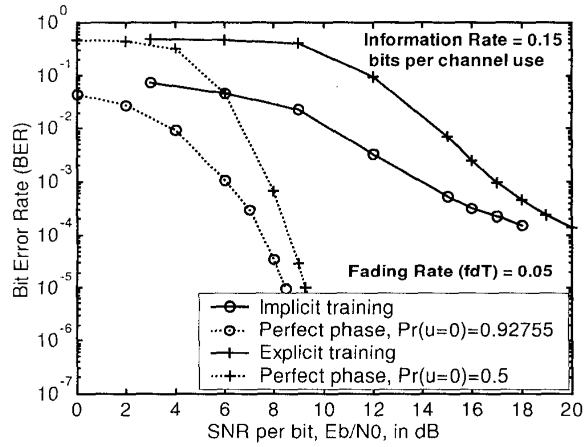

Figure 3. BER performance of implicit and explicit training. Medium-speed fading channel, Information rate = 0.15 bits/channel use.

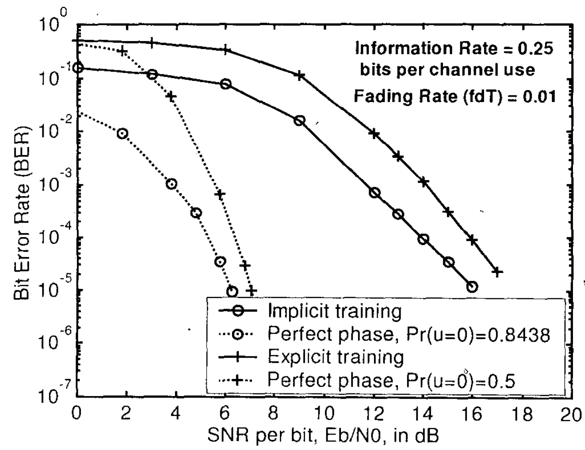

Figure 4. BER performance of implicit and explicit training. Slow fading channel, Information rate = 0.25 bits/channel use.