Improved iterative joint detection and estimation through variational Bayesian inference

S.M.-S. Sadough *, M. Modarresi

Cognitive Telecommunication Research Group, Department of Electrical Engineering, Shahid Beheshti University G. C., Tehran, Iran

A R T I C L E   I N F O

Article history:
Received 6 March 2011
Accepted 18 September 2011

Keywords:
Variational Bayesian inference
Joint iterative detection and channel estimation
Channel estimation errors
OFDM

A B S T R A C T

We propose an improved variational Bayesian (VB) receiver for orthogonal frequency-division multiplexing (OFDM) systems over frequency-selective block-fading channels. Conventional VB receivers provide distribution-estimates for the channel and information symbols iteratively and jointly. The proposed scheme is different from conventional VB inference in that the VB iterative receiver also exploits the hard channel estimate extracted from previous iterations to update the channel and symbol distributions. In this way, we reduce the impact of channel uncertainty on the decoder performance by means of a modified formulation of the VB formalism. The adequacy of the proposed approach compared to classically used VB receivers is demonstrated by simulations.

© 2011 Elsevier GmbH. All rights reserved.

1. Introduction

Orthogonal frequency-division multiplexing (OFDM) is a spectrally efficient modulation technique adopted in many communication systems and standards. It is well known that obtaining an accurate estimate in highly mobile environments only through the use of pilots would require inserting multiple training symbols per frame, which can result in a significant reduction of the spectral efficiency. Data-aided channel estimation methods can enhance system performance by exploiting the unknown data symbols in addition to few pilots (usually one or two used for algorithm initialization) in the channel estimation process [1].

The variational Bayesian (VB) inference [2,3], which is closely related to mean-field methods in statistical physics, has been recently adopted as an effective method for tractable receiver design. The VB method is applied in [4] for joint data detection and channel tracking of an OFDM system. A similar contribution is provided in [5] for an OFDM system working under a time-varying multipath channel. In [6], variational inference is used for soft detection in multiple access channels. An instance of using VB inference for joint image restoration and segmentation is presented in [7].

Regardless of the deployed technique (pilot-only or data-aided), channel estimation is an imperfect process and the poor quality of channel estimates degrades the performance of decoding at the receiver.

Recently in [8], assuming pilot-only channel estimation, Sadough et al. showed that the a posteriori probability density function (pdf) of the channel conditioned on its estimate plays an important role in improving the receiver’s performance, especially when the channel is estimated by a small number of pilots.

In this letter, we focus on data-aided VB reception and propose a modified probabilistic formulation for deriving the metrics of the VB receiver. In our formulation, in addition to the channel distribution, a hard channel estimate is involved in the derivation of unknown symbol distribution. This enables us to derive an alternative approximation for the symbol distribution that makes the receiver robust to channel uncertainties at the receiver.

Notational conventions: (·) refers to expectation with respect to the random vector x, cN(m, Σ) denotes complex Gaussian distribution with mean m and covariance matrix Σ; |·| and |·| denote absolute value and vector norm, respectively, (·)T, (·)H and (·)∗ denote vector transpose, Hermitian transpose and conjugation, respectively.

2. System description and channel model

As an example of a widely used signal model, we consider a coded OFDM system with M subcarriers through a frequency-selective multipath fading channel. We assume a block-fading channel model where each frame of size Mframe symbols corresponds to Mblock independent fading blocks. Notice that Mblock = 1

1 Although here we have considered the widely used OFDM signal model, it is important to mention that the proposed approach can be extended to any transmission scenario.

returns to the quasi-static channel model whereas $M_{\text{block}} = M_{\text{frame}}$ returns to the fast-fading channel model. Since the channel is assumed to be block-fading, for estimating the kth complex channel frequency coefficient $H_k$, we dispose of $N = M_{\text{frame}}/M_{\text{block}}$ observations. At the receiver, after removing the cyclic prefix, the signal corresponding to the kth subcarrier in a given fading block reads

$$y_k = H_k s_k + z_k \quad \text{for} \quad k = 1, \ldots, M,$$

(1)

where the $(1 \times N)$ vector $y_k = [y_{1,k}, \ldots, y_{N,k}]$, the entries of the noise vector $z_k$ are assumed to be zero-mean circularly symmetric complex Gaussian (ZMSCG), and the definition of $s_k$ and $z_k$ follow that of $y_k$. For the sake of notational simplicity, we will not specify hereafter the subscript k in (1).

3. Conventional variational Bayesian iterative detection

The optimal estimate of the symbol vector $s$ in (1) by using the maximum a posteriori (MAP) rule is given by

$$\hat{s}_{\text{MAP}} = \arg\max_s p(s|y).$$

(2)

The objective function in (2) can be written as

$$p(s|y) = \int p(s, H) p(y|H) dH = \int p(s|H, y) p(H) dH,$$

(3)

where $H$ is regarded as a nuisance parameter. Here, we assume that the channel is not known prior to data detection and thus the optimal solution is infeasible to obtain. The central idea of VB approximation [3] is to approximate the exact but intractable joint distribution into a product of marginals probabilities. Referring to our model (1), the VB method tries to find a distribution denoted by $q(s, H) = q(s|H) q(H)$ which approximates the exact posterior $p(s, H|y)$. This is achieved by solving the following minimization problem

$$\{q^*(s), q^*(H)\} = \arg\max_{q(s), q(H)} \text{KL}[q(s)q(H)||p(s, H|y)],$$

(4)

subject to:

$$q(s, H) = q(s)q(H),$$

$$\int q(s) ds = 1, \quad q(s) \geq 0 \quad \forall s,$$

$$\int q(H) dH = 1, \quad q(H) \geq 0 \quad \forall H,$$

where

$$\text{KL}[q(s)q(H)||p(s, H|y)] \triangleq \int q(s)q(H) \ln \frac{q(s)q(H)}{p(s, H|y)} ds dH,$$

is the Kullback–Leibler divergence [9]. Substituting $p(s, H|y) \approx q(s)q(H)$ into (3) leads to $p(s|y) \approx q(s)$ and $p(H|y) \approx q(H)$ which is equivalent to assuming that $s$ and $H$ are independent conditioned on $y$. Thus, the complicated integral in (3) is avoided since the VB method directly provides an approximation of $p(s|y)$ to be used in the MAP estimation (2).

Proposition 3.1. Any solution of the optimization problem (4) is obtained by alternating between the VBE-step and the VBM-step as

$$\text{VBE} : q^{t-1}(s) \propto \exp\{(p(s|H, y)q^{t-(1)}(H)|y\),$$

(5)

$$\text{VBM} : q^{t-1}(H) \propto \exp\{(p(s|H, y)q^{t-1}(s)|y\),$$

(6)

where the superscript $(t)$ denotes the iteration index, due to the fact that the solution of (4) is not explicit since $q(s)$ and $q(H)$ depend on each other.

Proof. The proof is provided in the appendix. □

4. Improved variational Bayesian iterative detection

At the $t$th VBM iteration, we assume that the distribution $q^{t-1}(H)$ is available from the $(t-1)$th iteration and obeys a Gaussian distribution $q^{t-1}(H) = CN\{\mu^{t-(1)}, \beta^{t-(1)}\}$, where $\mu^{t-(1)}$ and $\beta^{t-(1)}$ will be calculated later in this letter. We further define $H^{t-(1)} \equiv H^{t-(1)}$. To improve the detection performance of the conventional VBM, we propose to use the modified metric $p(s|y)$ as an alternative to $p(s|y)$ for data detection. This is achieved by modifying the formulation of the VBE update rules (5) and (6) as

$$\text{VBE} : q^{t}(s) \propto \exp\{(\ln p(s|H, y)q^{t-(1)}(H)|y\),$$

(7)

$$\text{VBM} : q^{t}(H) \propto \exp\{(\ln p(s|H, y)q^{t-1}(s)|y\),$$

(8)

Proposition 4.1. The update rules (7) and (8) are equal to

$$\text{VBE} : q^{t}(s) \propto p(s) \exp\left\{-\frac{1}{\sigma^2} \langle |y - Hs|^2 \rangle^{q^{t-1}(H)}\right\},$$

(9)

$$\text{VBM} : q^{t}(H) = CN\{\mu^{t}, \beta^{t}\},$$

(10)

where $\beta^{t} = (\sigma^2 + \beta^{t-(1)})/\beta^{t-(1)}$ and $p(s) = \prod_{i=1}^{N} p(s_i)$ where $p(s_i)$ is the a priori probability on symbol $s_i$ coming from the soft-input soft-output (SISO) decoder; and $\langle \cdot \rangle^{q^{t-1}(s)}$ and $\langle \cdot \rangle^{q^{t-1}(H)}$ are evaluated at the detector by using the distribution $q^{t}(s)$.

Proof. The proof is provided in the appendix. □

Since the elements of $s$ are independent, we have $q^{t}(s) = \prod_{i=1}^{N} q^{t}(s_i)$ where comparing our proposed VBM formalism in (9) with the conventional VBM reveals that

$$q^{t}(s) = p(s) \exp\left\{-\frac{1}{\sigma^2} \langle |y - Hs|^2 \rangle^{q^{t-1}(H)}\right\},$$

(11)

is the VB approximation for the exact posterior $p(s|y, H^{t-(1)})$ which is written as

$$p(s_i|y_i, H^{t-(1)}) \propto p(s_i)p(y_i|s_i, H^{t-(1)})$$

$$= p(s_i) \int p(y_i|s_i, H^{t-(1)}) dH,$$

$$= p(s_i) \int p(y_i|s_i, H^{t-(1)}) p(H|H^{t-(1)}) dH,$$

$$= p(s_i)p(y_i|s_i, H^{t-(1)}).$$

(12)

Since in (12) we do not dispose of the posterior distribution $p(H|H^{t-(1)})$, we propose to approximate it by $q^{t}(H)$. By doing so, we can derive an alternative approximate expression for $p(s_i|y_i, H^{t-(1)})$ in (12) that we denote by $q^{t}(s_i)$. We have

$$q^{t}(s) \propto p(s_i) p(y_i|s_i, H^{t-(1)}).$$

(13)

By using Theorem 1.2 in the appendix of [8] and standard Gaussian calculus, we obtain

$$q^{t}(s) \propto \frac{p(s_i)}{\pi(\sigma^2 + \beta^{t-(1)})} \exp\left\{-\frac{1}{\sigma^2 + \beta^{t-(1)}} \langle |s_i|^2 \rangle^{q^{t-1}(s_i)}\right\},$$

(14)

Notice that (14) is actually another approximation for the exact posterior $p(s_i|y_i, H^{t-(1)})$, and provides an alternative expression for distribution (11). We conclude that the proposed scheme leads to a modification of the metric used in conventional VBM. Thus, implementing the improved VBM will not increase the decoder.
complexity. Note also that in the proposed improved receiver, \( \hat{q}(s_i) \) is evaluated at the detector (also called soft demapper). In addition, the bit metrics are fed from the detector to the SISO decoder in our BICM reception scheme are derived from marginalizing the distribution \( \hat{q}(s_i) \). Due to space limitations, the author is urged to see [8] for more details on the interaction between the detector and the decoder in a BICM receiver implementation.

5. Simulation results

The parameters used throughout the simulations are as follows: one OFDM symbol is composed of \( M = 100 \) subcarriers. For channel coding, we consider the rate 1/3 recursive systematic convolutional (RSC) code of constraint length 3 defined in octal form by [5, 7, 7]. The interleaver is pseudo-random and operates over the entire frame that contains 15 OFDM symbols. Data symbols belong to 16-QAM constellation with set-partition labeling. We perform one SISO decoding iteration for each VBEM iteration.

Fig. 1 shows the BER performance versus \( E_b/N_0 \) in the case of a block-fading channel with \( N = 3 \) and 15. For comparison, we have also provided the BER obtained with pilot-only (one pilot symbol) channel estimation, the expectation–maximization (EM) algorithm as well as the BER obtained with perfect channel state information (CSI). For instance, when \( N = 3 \), it can be observed that the \( E_b/N_0 \) to attain a BER of \( 10^{-4} \) is reduced by about 1 dB by using the proposed VB receiver, compared to the conventional VB receiver. A similar improvement is observed for the case of a quasi-static fading channel where \( N = 15 \).

Fig. 2 depicts the BER versus the number of VBEM iterations, at a fixed \( E_b/N_0 \) of 8 dB, for \( N = 15 \). This allows to evaluate the number of iterations necessary to attain a target BER. From Fig. 2 we observe that the improved VB detector requires 5 iterations to achieve a BER of \( 3 \times 10^{-4} \) while the conventional VB detector does not attain this BER even after 8 iterations. Noting that each iteration involves the complicated forward–backward [10] decoding algorithm in addition to the VBEM computations, makes the proposed detector particularly useful for reducing the complexity at the receiver.

6. Conclusion

The problem of signal detection based on the theory of variational Bayesian inference was investigated. It was shown that our improved variational detector formulation leads to a modification of the metrics used in the conventional VB iterative receiver.

Therefore, our approach does not increase the complexity at the receiver. Our numerical results confirmed the adequacy of the improved VBEM detector in reducing the impact of channel estimation errors on the BER performance. Although in this letter we considered the widely used OFDM signal model, our proposed receiver design methodology holds for any transmission scenario.

Acknowledgements

The authors are grateful to Shahid Beheshti University for supporting this work under the grant number 600/1944.

Appendix A.

Derivation of conventional VBEM update rules (5) and (6)

To solve the minimization problem (4), we first take the functional derivative of \( KL[q(s)q(H)||p(s, H|y)] \) with respect to \( p(s) \), and equate to zero. We have

\[
\frac{\partial}{\partial q(s)} KL[q(s)q(H)||p(s, H|y)] = \int dHq(H) \left( \frac{\partial}{\partial q(s)} \int ds q(s) \ln \frac{q(s)}{p(s, H|y)} \right) = \int dHq(H) [\ln q(s) - \ln p(s, H|y) + 1] = 0, \quad (15)
\]

where we have omitted the terms not depending on \( q(s) \) and used the fact that \( q(H) \) and \( q(s) \) are pdfs. After straightforward simplifications, we get

\[
\ln q(s) = \int dHq(H) \ln p(s, H|y) + c = \int dHq(H) \ln p(s, H|y) + \ln p(y) - \ln p(y)] + c = \int dHq(H) \ln p(s, H, y) + c', \quad (16)
\]

where \( c \) and \( c' \) gather constant terms with respect to \( q(s) \) coming from enforcing the normalization of \( q(s) \) in (4). Taking the exponential from both sides of (16) and adopting an iterative implementation leads to (5).
By equating to zero the functional derivative of $\text{KLI}(q(s)q(H)|p(s, H|y))$ with respect to $q(H)$ and following similar steps as above, one can easily derive (6).

**Derivation of improved VBEM update rules (9) and (10)**

Let us first derive $q^{(t)}(s)$ in (9). Starting from equation (7) we have

$$q^{(t)}(s) \propto \exp\{\ln p(s, H, y|H^{(t-1)})q(H^{(t-1)})\}$$

$$\propto \exp\{\ln[p(y|H, s)p(H|H^{(t-1)})p(s)]q(H^{(t-1)})\}$$

$$\propto p(s)\exp\{\ln[p(y|H, s)]q(H^{(t-1)})\}$$

where we have used the independence between $s$ and $H$ and $H^{(t-1)}$, and omitted all terms that do not depend on $s$. Noting that the noise in (1) has a Gaussian distribution, it is straightforward to see that (17) is equivalent to (9).

Let us now calculate $q^{(t)}(H)$. Starting from (8) we have

$$q^{(t)}(H) \propto \exp\{\ln[p(s, H, y|H^{(t-1)})]q^{(t)}(s)\}$$

$$\propto \exp\{\ln[p(y|H, s)p(H|H^{(t-1)})p(s)]q^{(t)}(s)\}$$

$$\propto p(H|H^{(t-1)})\exp\{\ln[p(y|H, s)]q^{(t)}(s)\}$$

$$\propto p(H|H^{(t-1)})\exp\left\{-\frac{\|y-Hs\|^2}{\sigma^2}q^{(t)}(s)\right\},$$

where we have omitted all terms that do not depend on $H$. Assume that $p(H|H^{(t-1)}) = C N(\mu^{(t-1)}, \beta^{(t-1)})$, i.e., it obeys a Gaussian distribution. After simple calculus, Eq. (18) can be rewritten as

$$q^{(t)}(H) \propto \exp\left\{-\frac{\|y-Hs\|^2}{\sigma^2}q^{(t)}(s)\right\}$$

$$\propto \exp\left\{\frac{\|y-Hs\|^2}{\sigma^2}q^{(t)}(s)\right\}$$

$$+ 2\text{Re}\left[\frac{H^*y|s\rangle q^{(t)}(s)}{\sigma^2} + \frac{H^*\mu^{(t-1)}}{\beta^{(t-1)}}\right]$$

After some algebraic manipulations, we get

$$q^{(t)}(H) = \frac{1}{\pi\beta(t)} \exp\left\{-\frac{|H - \mu(t)|^2}{\beta(t)}\right\}$$

with

$$\beta(t) = \frac{\sigma^2}{\beta^{(t-1)} + \sigma^2}$$

$$\mu(t) = \frac{\beta^{(t-1)}y|s\rangle q^{(t)}(s)}{\sigma^2}$$

As a result, we obtain Eq. (10).

**References**