Co-Channel Interference Mitigation in Multiuser Systems with Unknown Channels

Dongning Guo, Yan Zhu and Michael L. Honig
Department of Electrical Engineering and Computer Science
Northwestern University, Evanston, IL, USA 60208
{dguo, yan-zhu, mh}@northwestern.edu

Abstract

This paper studies signal detection in multiuser systems where the uncertainty is due to fading as well as a strong co-channel interference of the same form as that of the desired signal. In particular, unlike for the desired signal, no pilot for the interference signal is available for measuring its fading channel state. Still, the interference is a non-Gaussian process and treating it as Gaussian noise can lead to poor performance. We propose a joint channel estimation and interference mitigation scheme based on belief propagation. Simulation results show that the receiver performs significantly better compared to conventional receivers with linear channel estimation.

I. Introduction

Wireless networks are fundamentally limited by two factors, namely, uncertainties about the time-varying communication channel and interference from other terminals in the system. Although each of the two issues has been studied in depth assuming the absence of the other one, much less is understood about the situation where both factors are significant. This work considers the detection of one signal in a time varying fading channel, which is also subject to interference signals of the same form and possibly of similar strength. Such problems arise for example in cellular networks where co-channel interference is strong and dominated by the signals from 1 or 2 particular users. Another example is ad hoc networks where terminals inevitably interfere with each other.

For simplicity, we restrict our attention to a two-user, narrow-band system with BPSK modulation, where the fading channel of each user is modeled as a Gauss-Markov process. Besides practical usage on its own, the model is the elementary building block of orthogonal frequency division multiplexing (OFDM), which is the basis for next generation wireless systems. We assume knowledge of the fading statistics, and that the desired user's channel is to be measured using known pilot symbols interleaved with data symbols, while no pilot is available for the interferer. This is a typical scenario in wireless systems where the pilot pattern of the co-channel interference signal is unavailable, e.g., in a cellular system where the interference is from a different cell than that of the desired user.

The detection problem is challenging due to simultaneous uncertainties of the channels as well as the interference. Conventionally, the channel state is measured assuming absence of interference, and interference is mitigated assuming complete knowledge about the channel state. In particular, as is seen in [1], knowledge of interfering pilots is indispensable to the success of linear channel estimation, even with iterative Turbo processing. The reason is that linear estimators treat the interference process as white Gaussian noise. For example, if the interference is as strong as the desired signal, the signal-to-noise-and-interference ratio is no better than 0 dB, which leads to poor channel estimates and unacceptable error probability.

In this work, we propose a general approach to joint channel estimation and interference mitigation in small multiuser systems, which can fully exploit knowledge of the non-Gaussian statistics of the interference. Consider the detection of a sequence of symbols of the desired user with one strong interferer of the same signaling format, where the fading processes of both the desired user and the interference are Gauss-Markov. The problem is fundamentally a compound hypothesis testing problem (averaged over channel uncertainty). Unfortunately, the maximum likelihood sequence detector, which exhausts the exponential number of hypotheses, is computationally prohibitive.

The low-complexity scheme proposed in this paper is based on belief propagation (BP), which is an iterative message-passing algorithm for performing statistical inference on *graphical models* by propagating locally computed "beliefs" [2]. One of the most successful applications of BP is the decoding of low-density parity-check (LDPC) codes. Combined channel estimation and detection, based on BP, has been previously proposed for a single-user fading channel (e.g., [3], [4]). A factor graph approach to multiuser detection for CDMA is presented in [5] with ideal (nonfading) channels. In this paper, however, we consider fading as well as the presence of a strong interference.

The problem at hand is equivalent to statistical inference on a tree, for which BP is ultimately optimal. For implementation, BP is constrained both in the number of iterations and in the way the messages exchanged between vertices in the graph are described. Numerical results demonstrate that the resulting suboptimal scheme performs remarkably better than the usual techniques of linear channel estimation followed by detection of individual symbols.

II. System model

Consider a narrow-band system where the received signal at time i in a frame of length n is expressed as

$$Y_i = H_i X_i + H'_i X'_i + N_i \qquad i = 1 \dots n \tag{1}$$

where X_i and X'_i denote the transmitted symbols of the desired user and interferer respectively, H_i and H'_i denote the corresponding channel gains, and $\{N_i\}$ is the sequence of i.i.d. circularly-symmetric complex Gaussian noise (CSCG) with zero mean and variance σ_N^2. For simplicity, we assume BPSK modulation, i.e., $X_i, X'_i = \pm 1$.

Assuming Rayleigh fading, $\{H_i\}$ and $\{H'_i\}$ are modeled as two independent Gauss-Markov processes, that is, they are generated by first-order auto-regressive systems (e.g., [6]):

$$H_i = \alpha H_{i-1} + \sqrt{1-\alpha^2}\, W_i, \qquad H'_i = \alpha H'_{i-1} + \sqrt{1-\alpha^2}\, W'_i \tag{2}$$

where $\{W_i\}$ and $\{W'_i\}$ are two independent white CSCG processes with variance σ_H^2 and $\sigma_H'^2$, respectively, and α is the correlation coefficient between adjacent fading coefficients.

Typically, pilots are inserted periodically between data symbols. Let Y_i^j denote the sequence $Y_i, \dots, Y_j$. The problem can be formulated as follows: Given the observations Y_1^n and the small subset of known pilots in X_1^n, we would like to detect the information symbols from the desired user, i.e., the remaining unknown symbols in X_1^n. Note that no knowledge of the channels and interfering symbols are available except for their statistics.

III. Graphical Model and BP Algorithm

A. Graphical Model

The fading coefficients $\{(Hi, Hi')\}_{i=1}^n$ form a Markov chain with state space in $\mathbb{C}^2$, conditioned on which the 3-tuple (X_i, X'_i, Y_i) at each time instant i is independent over time $i = 1, 2, \dots, n$. Therefore, the probability law of the system defined by relationships (1) and (2) can be fully described using a *factor graph* as is shown in Fig. 1.

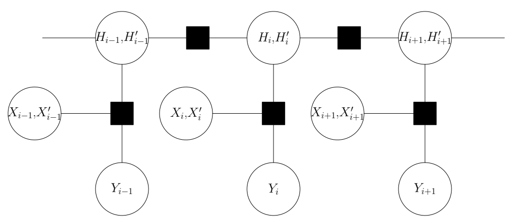


Fig. 1. A factor graph describing the two-user system.

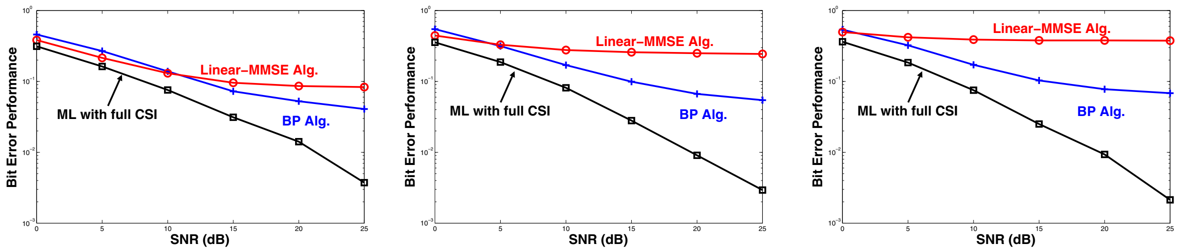
Generally, a factor graph is a *bipartite graph*, which consists of the variable nodes, denoted by circles in the graph, which represent the variables; and the factor nodes, denoted by squares which represent the constraints among the variable nodes connected to them. The factor node connecting nodes Y_i, (H_i, H'_i), and (X_i, X'_i) represents the probabilistic constraint among them, which is specified by (1) and is equivalent to the conditional Gaussian probability density $p(y_i|h_i, h'_i, x_i, x'_i)$. Similarly, the factor node between nodes (H_i, H'_i) and (H_{i-1}, H'_{i-1}) represents the relationship given by (2), i.e., $p(h_i, h'_i|h_{i-1}, h'_{i-1})$. The prior probability distribution of the data symbols is assigned as follows. All BPSK symbols X_i and X'_i are uniformly distributed on $\{-1, 1\}$ except for the subset of pilot symbols in X_1^n, for which we assume, without loss of generality, $P\{X_i = 1\} = 1$.

B. Exact Inference Via Message Passing

The detection problem is equivalent to statistical inference on the graph. The goal is to obtain the marginal posterior probability $p(x_i|y_1^n)$ for each i which corresponds to an unknown symbol. BP is an efficient algorithm for computing the posteriors by iteratively passing messages between neighboring nodes on the graph. The result of message passing with sufficient iterations gives the exact *a posteriori* probability of each unknown random variable if the factor graph is a tree (i.e., free of cycles) as is the case in the problem at hand.

For notational convenience, let $G_i = [H_i, H'_i]^{\mathrm{T}}$ and $Z_i = [X_i, X'_i]^{\mathrm{T}}$. The goal is to compute for each $i = 1, \dots, n$: $p(x_i|y_1^n) = \sum_{x'_i} \int p(z_i, g_i|y_1^n)\, \mathrm{d}g_i \propto \sum_{x'_i} \int p(z_i, g_i, y_1^n)\, \mathrm{d}g_i$ where we use the "proportion" notation $\propto$. This is because $Y_1^n = y_1^n$ is observed, and so that its probability is irrelevant to the inference problem and can be treated as a constant. By the Markovian property, Z_i, Y_i, Y_1^{i-1} and Y_{i+1}^n are mutually independent given G_i. Therefore,

$$p(x_i|y_1^n) \propto \sum_{x'_i} \int p(y_i, z_i|g_i) p(y_1^{i-1}|g_i) p(y_{i+1}^n|g_i) p(g_i)\, \mathrm{d}g_i \propto \sum_{x'_i} \int p(y_i, z_i|g_i) p(g_i|y_1^{i-1}) p(g_i|y_{i+1}^n)/p(g_i)\, \mathrm{d}g_i\,.$$



(a) (b) (c)

Fig. 2. BER performance of the BP algorithm. (a) The power of the interference is 10 dB weaker than that of the desired user. (b)The power of the interference is 3 dB weaker. (c)The power of the interference is identical to that of the desired user.

Since Z_i and G_i are independent, we finally have

$$p(x_i|y_1^n) \propto \sum_{x_i'} p(z_i) \int p(y_i|g_i, z_i) p(g_i|y_1^{i-1}) p(g_i|y_{i+1}^n)/p(g_i) \mathrm{d}g_i. \tag{3}$$

Note that $p(y_i|g_i, z_i)$ can be obtained by (1) and $p(z_i) = p(x_i)p(x_i')$ since symbols from the two users are independent. In order to compute (3), it suffices to compute $p(g_i|y_1^{i-1})$ and $p(g_i|y_{i+1}^n)$ separately.

We give a brief derivation of $p(g_i|y_{i+1}^n)$ in below, whereas computation of $p(g_i|y_1^{i-1})$ is similar by symmetry. The technique is to develop a recursion for the posterior probability. First, we have $p(g_i|y_{i+1}^n) = \int p(g_i|g_{i+1}) p(g_{i+1}|y_{i+1}^n) \mathrm{d}g_{i+1}$ because given G_{i+1}, G_i and Y_{i+1}^n are independent. By the Markovian property, Y_{i+2}^n and Y_{i+1} are independent given G_{i+1}. Therefore, $p(g_i|y_{i+1}^n) \propto \int p(g_i|g_{i+1}) p(y_{i+1}|g_{i+1}) p(g_{i+1}|y_{i+2}^n) \mathrm{d}g_{i+1}$. Since G_{i+1} and Z_{i+1} are independent, $p(y_{i+1}|g_{i+1}) = \sum_{z_{i+1}} p(y_{i+1}, z_{i+1}|g_{i+1}) = \sum_{z_{i+1}} p(y_{i+1}|z_{i+1}, g_{i+1}) p(z_{i+1})$. Therefore, we have

$$p(g_i|y_{i+1}^n) \propto \sum_{z_i} \int p(g_i|g_{i+1}) p(g_{i+1}|y_{i+2}^n) p(y_{i+1}|g_{i+1}, z_{i+1}) p(z_i) \, \mathrm{d}g_{i+1}. \tag{4}$$

Note that $p(g_i|g_{i+1})$ can be obtained from (2). Therefore, (4) gives a recursion for computing $p(g_i|y_{i+1}^n)$ for each $i = 1, \dots, n$, which is the essence of the message passing algorithm. The joint channel estimation and interference mitigation algorithm is summarized as the following.

Initialization: $P(X_i' = 1) = P(X_i' = -1) = 1/2$ for all i. The same probabilities are also assigned to $P(X_i)$ for all i except for pilots, for which $P(X_i = 1) = 1$. For all i, $p(h_i)$ and $p(h_i')$ are zero mean Gaussian with variance σ_H^2 and $\sigma_{H'}^2$ respectively.

for $i = 1$ to n **do**

  Compute $p(g_i|y_{i+1}^n)$ from (4); then compute $p(g_i|y_1^{i-1})$.

end for

for $i = 1$ to n **do**

  Compute $p(x_i|y_1^n)$ from (3)

end for

Basically, the message from a factor node to a variable node is a summary of knowledge about the random variable(s) represented by the variable node based on all observations connected directly or indirectly to the factor node. For example, the message received by node (H_i, H_i') from the factor node on its right side summarizes all the information about (H_i, H_i') based on the observations $Y_{i+1}, \dots, Y_n$, which is proportional to $p(h_i, h_i'|y_{i+1}^n)$. The message from a variable node to a factor node is a summary of information about the variable node based on the observations connected to it. For example, the message passed by node (H_i, H_i') to the factor node on its left is the inference about (H_i, H_i') based on the observations $Y_i, \dots, Y_n$, i.e., $p(h_i, h_i'|y_i^n)$.

C. Implementation Issues

The above algorithm cannot be implemented directly using a digital computer because the messages are continuous probability density functions (PDF). An observation is that the random variables in Fig. 1 are either Gaussian or discrete. According to (4), it can be shown by induction that the density functions, $p(g_i|y_{i+1}^n)$ and $p(g_i|y_1^{i-1})$ are mixture Gaussian functions. The mixture Gaussian function is completely characterized by the amplitudes, means and variances of its components. Therefore, we can compute and pass these parameters only. Let $\mathcal{CN}(\mathbf{x}, \mathbf{m}, \mathbf{K})$ denote complex Gaussian density with mean $m_{l\times 1}$ and covariance $K_{l\times l}$. Then we can write $p(g_i|g_{i+1}) = \mathcal{CN}(g_i, \alpha g_{i+1}, B)$, $p(g_i)=\mathcal{CN}(g_i, 0, Q)$ and $p(y_i|g_i, z_i)=\mathcal{CN}(y_i, z_i^{\mathrm{T}} g_i, \sigma_N^2)$, where $Q=diag(\sigma_H^2, \sigma_{H'}^2)$ and $B = (1-\alpha^2)Q$.

Note that the message passing starts with a Gaussian density function. According to (4), as the message is passed from node to node, it becomes a mixture of more and more Gaussian densities. Therefore, without loss of generality, we assume that $p(g_{i+1}|y_{i+2}^n) = \sum_j \rho_j \mathcal{CN}(g_{i+1}, m_j, K_j)$. Substituting into (4), we have after some manipulations $p(g_i|y_{i+1}^n) \propto \sum_j \sum_{z_{i+1}} \rho_j p(z_{i+1}) L(j, z_{i+1}) C(j, z_{i+1})$ where $L(j, z_{i+1})$ and $C(j, z_{i+1})$ are Gaussian densities, the mean and covariance of which can be evaluated from m_j and K_j. Basically, an explicit recursion can be established for the amplitude, mean and variance of each Gaussian component in message $p(g_i|y_{i+1}^n)$. Similar computations apply to $p(g_i|y_1^{i-1})$. Therefore, the BP algorithm requires message-passing backward and forward once each direction, which is similar to the BCJR algorithm [7]. The key difference between our algorithm and the BCJR is that the Markov chain here has a continuous state space. The preceding Gaussian mixture representation can also be used to compute $p(g_i|y_{i+1}^n)$, $p(g_i|y_1^{i-1})$ and $p(g_i|y_1^n)$.

The number of Gaussian components increases exponentially in the recursive, which is intractable in practice. One way to deal with this issue is to merge some of the components in order to keep a fixed total number of components. In this work, we simply keep the components with the largest amplitudes. One can also trade complexity with performance, in the sense that if the messages are unrestricted, then the algorithm implements the maximum *a posteriori* detector. By restricting the number of components in each message, the complexity of the algorithm is linear in the block length n, i.e., it is a constant per symbol. While the complexity is typically higher than the conventional linear estimator, the performance gain may easily justify the additional complexity. As the constellation size increases, while the complexity of the algorithm increases quickly, the advantage over linear channel estimation vanishes because the interference appears to be more Gaussian-like. Thus the algorithm is particularly suitable for BPSK and QPSK modulations.

IV. Simulation Results

The performance of the BP algorithm is plotted versus signal-to-noise ratio $SNR = \sigma_H^2/\sigma_N^2$. We set $\alpha = .99$, limit the maximum number of Gaussian components to 16 and set the block length $n = 200$. Within each block, there is one pilot in every 4 symbols. The BP algorithm with Gaussian mixture messages is simulated. We also simulate two other receivers for comparison. One is denoted by "MMSE", which estimates the desired channel by taking a linear combination of adjacent received value. This MMSE estimator treats the interference as white Gaussian noise. The other is denoted by "ML with full CSI", which performs maximum likelihood (ML) detection symbol by symbol assuming that the realization of the fading processes is revealed to the detector by a genie. It is a genie-aided receiver and serves as a benchmark or lower bound for all receivers. Figs. 2(a)-2(c) show the BER performance where the power of the interference is 10 dB weaker, 3 dB weaker and equal to that of the desired user, respectively. The BP algorithm generally gives a significant performance gain over the MMSE algorithm, especially in the high SNR region. Note that thermal noise dominates when the interference is weak. Therefore, relatively small gain over the MMSE algorithm is observed in Fig. 2(a). The trend of the BP curves shows that the BP algorithm effectively mitigates or partially cancels the interference, as opposed to suppressing it.

V. Conclusion

A joint channel estimation and interference mitigation scheme based on belief propagation has been studied. Unlike conventional linear channel estimation followed by symbol-wise detection, the proposed scheme exploits the non-Gaussian statistics of the interference. The BP algorithm has significant performance gain over the traditional interference suppression schemes, even though the overall complexity is limited to a constant per symbol.

References

[1] K. Sil, M. Agarwal, D. Guo, and M. L. Honig, "Performance of Turbo decision-feedback detection and decoding in downlink OFDM," in *IEEE Wireless Commun. & Networking Conf.*, 2007.

[2] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[3] A. P. Worthen and W. Stark, "Unified design of iterative receivers using factor graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 843–849, Feb. 2001.

[4] X. Jin, A. W. Eckford, and T. E. Fuja, "LDPC codes for non-coherent block fading channels with correlation: Analysis and design," *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 70–80, Jan. 2008.

[5] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1772–1793, Jul. 2002.

[6] I. Abou-Faycal, M. Médard, and U. Madhow, "Binary adaptive coded pilot symbol assisted modulation over Rayleigh fading channels without feedback," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 1036–1046, June 2005.

[7] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, 1974.