Improved Modulation Classification using a Factor-Graph-based Iterative Receiver

Daniel Jakubisin and R. Michael Buehrer
Mobile and Portable Radio Research Group (MPRG), Wireless@VT,
Virginia Tech, Blacksburg, Virginia, USA. {jakubisin, buehrer}@vt.edu

***Abstract*—We bring together two research topics which have been the focus of significant research individually: modulation classification and iterative receiver design. In this work, these topics are joined within the framework of factor graphs which provide a unified approach to representing a variety of algorithms, especially iterative algorithms. Specifically, in this paper we present a factor graph which incorporates modulation classification into the iterative receiver structure. The proposed iterative receiver applies message passing on the factor graph to approximate the optimal solution to joint modulation classification, demodulation, and decoding. This results in a classifier which treats feedback from the decoder as *a priori* probabilities for the coded bits. We show that the proposed receiver is able to achieve significant performance gains over a receiver which performs maximum likelihood classification separately from demodulation and decoding.**

I. Introduction

Traditionally, receivers have been designed to complete tasks such as synchronization, channel estimation, demodulation, and decoding in a sequential fashion. Optimal maximum *a posteriori* (MAP) or maximum likelihood (ML) detection of the information bits requires that these tasks be completed jointly which is prohibitively complex. The term "iterative receiver" refers to a communication receiver in which the MAP or ML estimate of the transmitted message is approximated using an iterative algorithm (i.e., using the turbo principle [1]). In the context of turbo codes, empirical results have shown that iterative decoding provides performance near the Shannon capacity bound [2]. Iterative algorithms for joint synchronization and decoding have been shown to achieve performance near that of decoding with perfect synchronization [3]. Significant contributions have been made to iterative receiver design in the areas of channel estimation [4], [5], equalization [6], and synchronization [3], [7], [8]. Numerous concepts from these areas were joined together in the work by Wymeersch [9].

In this work, we propose incorporating automatic modulation classification into the iterative receiver structure. To the best of our knowledge, this concept has not been previously explored. Adaptive modulation systems are attractive because they provide a means of optimizing link throughput according to channel conditions. Generally, information about the selected modulation scheme must be conveyed to the receiver which wastes bandwidth and energy. In order to avoid this loss, it is desirable that the receiver have the ability to automatically classify the modulation. Automatic modulation classification has also been the focus of significant research effort [10].

The concept behind the proposed iterative receiver is to improve classification by exploiting extrinsic information obtained from soft decoding. The extrinsic information from the soft decoder is treated as *a priori* coded bit probabilities (or equivalently the symbol probabilities). This is in contrast to past research into ML classification in which the symbols are assumed to be equally likely *a priori*. ML modulation classification with equally likely symbols has been shown to have arbitrarily low probability of error when the number of symbols used by the classifier goes to infinity [11]. In this work we consider relatively short frame lengths and low signal to noise ratio (SNR). We demonstrate through simulation, that the iterative receiver achieves significant performance gains over ML classification performed separately from demodulation and decoding which we will term the "traditional receiver".

Joint modulation classification and decoding requires a receiver structure which supports each modulation scheme simultaneously. Factor graphs provide a useful tool in constructing such a system. The sum-product algorithm is a message passing algorithm for computing marginals on a factor graph [12]. We implement the sum-product algorithm on the factor graph to iteratively solve the joint problem.

In Section II we present background on the system model, factor graphs, and ML classification. In Section III we present the proposed iterative receiver. Simulation results are presented in Section IV and conclusions are given in Section V.

II. Background

A. System model

We consider a communication system which uses bit interleaved coded modulation (BICM) and may employ one of M amplitude-phase modulation schemes. The number of symbols K in a transmitted frame is fixed and the number of coded bits is allowed to vary according to the selected modulation. In order to handle the variable number of coded bits, the length of the channel encoder is set to provide K coded output bits and the bit interleaver is set to interleave K bits at a time. In this way a single encoder and a single bit interleaver are used for BPSK. In general, when the transmitter uses a modulation with a set of S constellation points, $Z = \log_2 S$ encoder/interleaver blocks are required. This structure for the encoder and interleaver makes it possible to design a receiver which can incorporate the de-interleaving and decoding of

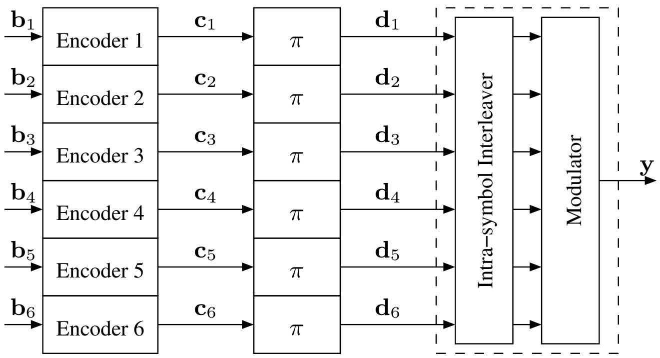

Fig. 1. Diagram of the transmitter structure for 64QAM

each modulation within a single structure. As an example, the diagram of the transmitter for 64QAM is shown in Fig. 1.

The message sequence is divided into Z blocks $\mathbf{b}_z$, each of which is the input to one of the encoders. The output of the zth encoder $\mathbf{c}_z$ is applied to the zth bit interleaver. The output of the bit interleaver $\mathbf{d}_z$ is a pseudo-random permutation (π) of the input bits. We denote the entire block of information, coded, and interleaved bits as $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_Z]$, $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_Z]$, and $\mathbf{d} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_Z]$, respectively.

The kth bit from the output of each bit interleaver is denoted $\mathbf{d}_{:,k} = [d_{1,k}, d_{2,k}, \ldots, d_{Z,k}]$ and is passed to the modulator in order to generate symbol y_k. Within the modulator the input bits are interleaved. We call this the "intra-symbol interleaver" and we include it so that the net result of the bit interleavers and the intra-symbol interleaver is to fully interleave the coded bits in keeping with the BICM design of Caire et. al. [13] while maintaining the separate encoder and bit interleaver blocks. The intra-symbol interleaver mapping is generated randomly for each symbol and is denoted $\psi_{m,k}$ where m refers to the modulation scheme. The modulator employs Gray labeling.

The modulator generates symbols y_k from a unit energy constellation $E[|y_k|^2] = 1$. We assume a system in which the combined response of the pulse shaping filter and receiver's matched filter satisfies the Nyquist condition for zero intersymbol interference. In order to test the concept of the proposed receiver, we simulate transmission of the signal over an additive white Gaussian noise (AWGN) channel. After matched filtering, the signal is sampled at the symbol rate to generate a vector of complex valued samples denoted by $\mathbf{r} = [r_1, r_2, \ldots, r_K]^T$. The complex samples are comprised of symbols and noise as given by $\mathbf{r} = \mathbf{y} + \mathbf{n}$ where $\mathbf{y} = [y_1, y_2, \ldots, y_K]$ and we assume that the noise $\mathbf{n}$ is a vector of independent and identically distributed (iid) complex Gaussian random variables. The noise power is denoted σ_N^2 and is assumed to be known.

B. Factor graphs and the sum-product algorithm

A factor graph is a graphical model which visually represents the factorization of a function. For our purposes the functions are joint probability density functions (f) or joint probability mass functions (p). These functions can be factored by exploiting the independence and/or conditional dependencies among the variables. Consider an example function $f(x_1, x_2, x_3, x_4, x_5)$ which can be factored as follows:

$$f(x_1, x_2, x_3, x_4, x_5) = f_1(x_1) f_2(x_1, x_2, x_3) \cdot f_3(x_3, x_4) f_4(x_4, x_5). \quad (1)$$

A Forney style factor graph [12] is constructed from nodes and edges where nodes represent the factors f_k ($k = 1, 2, 3, 4$) and edges represent the variables x_n ($n = 1, 2, 3, 4, 5$). A variable edge is connected to a factor node if the variable appears as an argument of the factor. The factor graph of (1) is shown in Fig. 2.

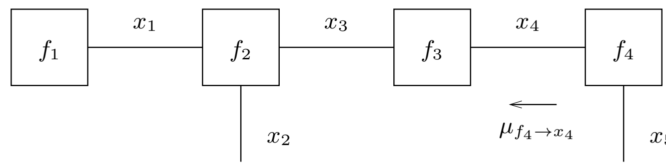

Fig. 2. Example factor graph

We are interested in computing marginals from a joint distribution. The marginal of the variable x_1 in our example is expressed as

$$f_{X_1}(x_1) = f_1(x_1) \left(\sum_{x_2, x_3} f_2(x_1, x_2, x_3) \cdot \left(\sum_{x_4} f_3(x_3, x_4) \left(\sum_{x_5} f_4(x_4, x_5) \right) \right) \right). \quad (2)$$

The sum-product algorithm is used to efficiently compute a marginal of a distribution by making local summations at the nodes of the factor graph and by passing these summations in the form of "messages" along the edges of the graph. Messages are labeled according to the node and edge they are associated with and are functions of the edge variable. For example, the message from node f_4 to variable x_4 is denoted $\mu_{f_4 \to x_4}(x_4)$ and is given by

$$\mu_{f_4 \to x_4}(x_4) = \sum_{x_5} f_4(x_4, x_5). \quad (3)$$

It can be seen from (2) that the summation in (3) contains all the information about x_5 and f_4 that is necessary for computing the marginal $f_{X_1}(x_1)$. The message $\mu_{f_4 \to x_4}(x_4)$ is used in the summation at node f_3 to generate the message $\mu_{f_3 \to x_3}(x_3)$ and this process is continued. Finally, the marginal $f_{X_1}(x_1)$ is given by the multiplication of the messages in both directions on edge x_1 expressed as $f_{X_1}(x_1) = \mu_{f_1 \to x_1}(x_1) \cdot \mu_{f_2 \to x_1}(x_1)$. For a complete introduction to factor graphs and the sum-product algorithm please refer to [12].

C. Maximum Likelihood Classification

Let $\mathcal{H}_m$ denote the hypothesis that the received frame employs the mth modulation scheme. For an observation of $\mathbf{r}$, the MAP modulation class is the class $\mathcal{H}_m$ which maximizes the probability $\mathrm{P}[\mathcal{H}_m|\mathbf{r}]$. When the modulations are equally likely *a priori*, the MAP and ML classifiers are equivalent as given by [14]

$$\mathrm{P}[\mathcal{H}_m|\mathbf{r}] = \frac{f(\mathbf{r}|\mathcal{H}_m)\mathrm{P}[\mathcal{H}_m]}{f(\mathbf{r})} \propto f(\mathbf{r}|\mathcal{H}_m) \tag{4}$$

where we are concerned with the maximum (not the actual probabilities) so multiplicative constants can be ignored.

It is assumed that the symbols are independent [11]. Due to the assumption of iid noise samples and independent symbols, the likelihood function $f(\mathbf{r}|\mathcal{H}_m)$ can be expressed by the product [11]

$$f(\mathbf{r}|\mathcal{H}_m) = \prod_{k=1}^{K} f(r_k|\mathcal{H}_m). \tag{5}$$

Let $s_{m,l}$ denote the lth constellation point of the mth modulation and let S_m denote the total number of constellation points in modulation m. The distribution $f(r_k|\mathcal{H}_m)$ is expressed as a marginalization over the joint distribution $f(r_k, s_{m,l}|\mathcal{H}_m)$ [11] which produces

$$\begin{aligned} f(\mathbf{r}|\mathcal{H}_m) &= \prod_{k=1}^{K} \sum_{l=1}^{S_m} f(r_k|s_{m,l}, \mathcal{H}_m)\mathrm{P}[s_{m,l}|\mathcal{H}_m] \\ &= \prod_{k=1}^{K} \sum_{l=1}^{S_m} \frac{1}{\pi\sigma_N^2} \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{m,l}\right|\right|^2\right)\frac{1}{S_m} \\ &\propto \prod_{k=1}^{K} \frac{1}{S_m} \sum_{l=1}^{S_m} \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{m,l}\right|\right|^2\right) \end{aligned} \tag{6}$$

where we assume that all symbols within a modulation are equally likely ($p(s_{m,l}|\mathcal{H}_m) = 1/S_m$). From (6) we can write the ML classifier as follows

$$\mathcal{H}_{\mathrm{ML}} = \arg\max_{\mathcal{H}_m} \prod_{k=1}^{K} \frac{1}{S_m} \sum_{l=1}^{S_m} \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{m,l}\right|\right|^2\right). \tag{7}$$

The result in (7) is used for the traditional receiver and will also be useful to verify the sum-product algorithm on the iterative receiver in Section III-B.

III. Joint Decoding and Classification

The joint probability density function of the presented system model can be factored into several conditional and marginal distributions as given by

$$\begin{aligned} &f(\mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{r}, m) \\ &\quad = f(\mathbf{r}|\mathbf{y})p(\mathbf{y}|\mathbf{d}, m)p(m)p(\mathbf{d}|\mathbf{c})p(\mathbf{c}|\mathbf{b})p(\mathbf{b}). \end{aligned} \tag{8}$$

where $\mathbf{r}$ is the observed sample vector. The modulations are equally likely *a priori*: $p(m) = 1/M$. The likelihood function of the information bits can be determined (up to a scaling factor) as the marginalization of (8) over $\mathbf{c}$, $\mathbf{d}$, $\mathbf{y}$, and m. This marginalization is expressed as

$$\begin{aligned} &f(\mathbf{r}|\mathbf{b}) \propto f(\mathbf{r}, \mathbf{b}) \\ &= \sum_{\mathbf{c},\mathbf{d},\mathbf{y},m} f(\mathbf{r}|\mathbf{y})p(\mathbf{y}|\mathbf{d}, m)p(m)p(\mathbf{d}|\mathbf{c})p(\mathbf{c}|\mathbf{b})p(\mathbf{b}) \\ &= p(\mathbf{b}) \sum_{\mathbf{c}} p(\mathbf{c}|\mathbf{b}) \sum_{\mathbf{d}} p(\mathbf{d}|\mathbf{c}) \sum_{\mathbf{y},m} f(\mathbf{r}|\mathbf{y})p(\mathbf{y}|\mathbf{d}, m)p(m). \end{aligned} \tag{9}$$

From the likelihood function of (9) we can determine the ML estimate of the information bits. In a similar way, the likelihood function of the modulation can be determined from the marginalization of (8) over $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, and $\mathbf{y}$ as given by

$$\begin{aligned} &f(\mathbf{r}|m) \propto f(\mathbf{r}, m) \\ &= \sum_{\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{y}} f(\mathbf{r}|\mathbf{y})p(\mathbf{y}|\mathbf{d}, m)p(m)p(\mathbf{d}|\mathbf{c})p(\mathbf{c}|\mathbf{b})p(\mathbf{b}) \\ &= p(m) \sum_{\mathbf{y}} f(\mathbf{r}|\mathbf{y}) \sum_{\mathbf{d}} p(\mathbf{y}|\mathbf{d}, m) \sum_{\mathbf{c}} p(\mathbf{d}|\mathbf{c}) \sum_{\mathbf{b}} p(\mathbf{c}|\mathbf{b})p(\mathbf{b}). \end{aligned} \tag{10}$$

The expression for the likelihood function of the modulation (10) incorporates information about $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{d}$ and therefore is not restricted to the assumption that the symbols are equally likely and independent. The marginalization of (9) and (10) can be obtained simultaneously from the sum-product algorithm.

A. Factor graph

In the iterative receiver we require the ability to decode the frame of several modulation schemes. Thus, the number of decoder/de-interleaver blocks is determined by the highest order modulation ($Z = \max_m \log_2 S_m$). The factor graph representing the factorization in (8) is given in Fig. 3. Each factor of (8) is considered in the paragraphs that follow.

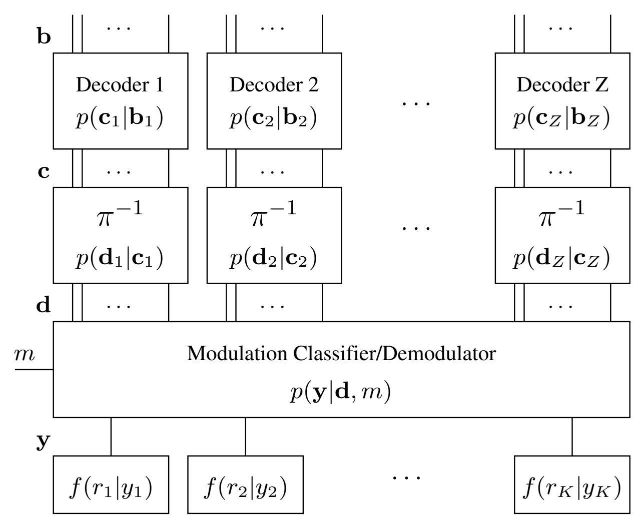

Fig. 3. Iterative receiver factor graph

The encoders and interleavers are deterministic functions of the input. Let the encoder be expressed by the function $g(\cdot)$

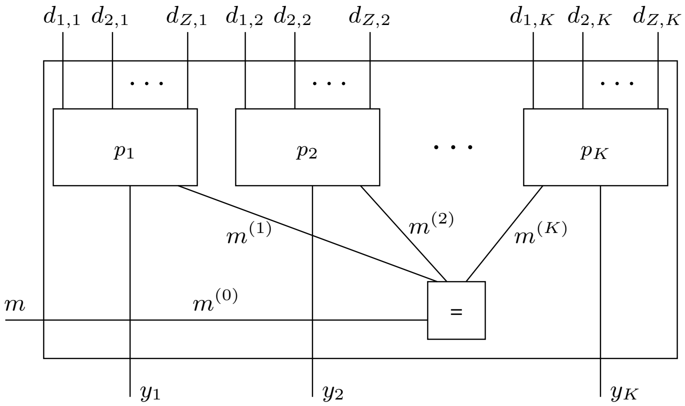

Fig. 4. Factor graph of the joint classifier and demodulator node

such that $\mathbf{c}_z = g(\mathbf{b}_z)$. This allows the conditional distribution $p(\mathbf{c}|\mathbf{b})$ to be expressed

$$p(\mathbf{c}|\mathbf{b}) = \prod_{z=1}^{Z} \mathrm{I}[\mathbf{c}_z = g(\mathbf{b}_z)] \tag{11}$$

where $\mathrm{I}[\cdot]$ is the indicator function with value 1 if the expression is true and value 0 otherwise. For additional details on the structure of the decoder and the corresponding sum-product algorithm the reader is referred to [9]. In a similar way, the conditional distribution $p(\mathbf{d}|\mathbf{c})$ may be expressed

$$p(\mathbf{d}|\mathbf{c}) = \prod_{z=1}^{Z} \prod_{k=1}^{K} \mathrm{I}[c_{z,k} = d_{z,\pi(k)}]. \tag{12}$$

Using factor graphs, we are able to combine modulation classification and demodulation in a very natural way. The conditional distribution $p(\mathbf{y}|\mathbf{d}, m)$ can be factored into K terms as follows:

$$p(\mathbf{y}|\mathbf{d}, m) = \prod_{k=1}^{K} p_k(y_k|\mathbf{d}_{:,k}, m) \tag{13}$$

A detailed view of the classifier/demodulator node is shown in Fig. 4 where the notation $m^{(k)}$ is used to denote multiple instances of the same modulation variable which are constrained to be equal through the "equals node" [9]. This factorization follows the transmitter's structure which takes an output from each bit interleaver to construct a symbol. The intra-symbol interleaver is not shown in this structure because it is handled within each factor node p_k. Let the function h_m provide a constellation point from the mth modulation given a vector of input bits. This leads to the following function

$$\begin{aligned} p_k(y_k|\mathbf{d}_{:,k}, m) = &\, \mathrm{I}\left[y_k \in \{s_{m,1}, \ldots, s_{m,S_m}\}\right] \\ &\cdot \mathrm{I}\left[h_m\left(\psi_{m,k}\left(\mathbf{d}_{:,k}\right)\right) = y_k\right]. \end{aligned} \tag{14}$$

The function enforces the structure of the intra-symbol interleaver and modulator for the modulation specified by m.

Finally, the distribution $f(\mathbf{r}|\mathbf{y})$ can be factored into K terms due to the conditional independence of the samples

$$f(\mathbf{r}|\mathbf{y}) = \prod_{k=1}^{K} f(r_k|y_k) = \prod_{k=1}^{K} \exp\left(-\frac{1}{\sigma_N^2} ||r_k - y_k||^2\right). \tag{15}$$

B. Message passing

The high level schedule for the sum-product algorithm on the factor graph of Fig. 3 begins with computing messages $\mu_{f(r_k|y_k)\to y_k}$. These messages are input to the sum-product algorithm on the classifier/demodulator which will be described in this section. The result of the sum-product algorithm on this block is messages $\mu_{p(\mathbf{y}|\mathbf{d},m)\to m}$ and $\mu_{p(\mathbf{y}|\mathbf{d},m)\to \mathbf{d}}$. The messages $\mu_{p(\mathbf{y}|\mathbf{d},m)\to \mathbf{d}}$ are passed to the de-interleaver which produces $\mu_{p(\mathbf{d}|\mathbf{c})\to \mathbf{c}}$. This provides input to the decoder which provides messages $\mu_{p(\mathbf{c}|\mathbf{b})\to \mathbf{c}}$ and $\mu_{p(\mathbf{c}|\mathbf{b})\to \mathbf{b}}$. The messages from the decoder to the coded bits are passed through the de-interleaver producing $\mu_{p(\mathbf{d}|\mathbf{c})\to \mathbf{d}}$. This provides local *a priori* input to the classifier/demodulator and completes one iteration. Using the new information about the messages $\mu_{p(\mathbf{d}|\mathbf{c})\to \mathbf{d}}$, successive iterations are performed.

We now consider the sum-product algorithm in detail for the classifier/demodulator node. Message passing for the joint classifier and demodulator is outlined in the following steps:

1) Initialize input messages from $\mathbf{y}$ and $\mathbf{d}$ to nodes p_k and from $m^{(0)}$ to the equals node.
2) Compute messages $\mu_{p_k\to m^{(k)}}$ for $k = 1, 2, \ldots, K$.
3) Update the equals node outward messages $\mu_{m^{(k)}\to p_k}$.
4) Compute messages $\mu_{p_k\to d_{z,k}}$ for $z = 1, 2, \ldots, Z$ and $k = 1, 2, \ldots, K$.

In the factor graph framework, the message for a variable out of an equals node is the product of the messages from all other variables connected to the node. Thus, the messages for the modulation variable can be computed as

$$\mu_{eq\to m^{(0)}}(m) = \prod_{k=1}^{K} \mu_{p_k\to m^{(k)}}(m) \tag{16}$$

where the marginal is given by

$$p(m|\mathbf{r}) = \mu_{eq\to m^{(0)}} \cdot \mu_{m^{(0)}\to eq} \tag{17}$$

and the message $\mu_{m^{(0)}\to eq}$ contains the *a priori* probabilities of the modulation schemes $p(m)$.

In the first iteration of the sum-product algorithm, no prior information about the interleaved bits $\mathbf{d}$ is available. The messages $\mu_{d_{z,k}\to p_k}$ are initialized as uniform distributions. The messages on the edges of the symbol variable y_k contain probabilities for the constellation points in all modulations as given by

$$\mu_{y_k \to p_k}(y_k) = \begin{bmatrix} \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{1,1}\right|\right|^2\right) \\ \vdots \\ \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{m,l}\right|\right|^2\right) \\ \vdots \\ \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{M,S_M}\right|\right|^2\right) \end{bmatrix} \quad (18)$$

The message computation for $m^{(k)}$ at each factor p_k is computed according to the sum-product algorithm as follows:

$$\begin{aligned} &\mu_{p_k \to m^{(k)}}(m) \\ &= \sum_{y_k, \mathbf{d}_{:,k}} p_k(y_k|\mathbf{d}_{:,k}, m) \cdot \mu_{y_k \to p_k}(y_k) \\ &\quad \cdot \mu_{d_{1,k} \to p_k}(d_{1,k}) \cdot \mu_{d_{2,k} \to p_k}(d_{2,k}) \dots \mu_{d_{Z,k} \to p_k}(d_{Z,k}) \\ &= \frac{1}{2^Z} \sum_{y_k, \mathbf{d}_{:,k}} \mathrm{I}\left[y_k \in \{s_{m,1}, \dots, s_{m,S_m}\}\right] \\ &\quad \cdot \mathrm{I}\left[g_m\left(\psi_{m,k}\left(\mathbf{d}_{:,k}\right)\right) = y_k\right] \cdot \mu_{y_k \to p_k}(y_k). \end{aligned} \quad (19)$$

For BPSK ($m = 1$) the expression in (19) can be further reduced to

$$\begin{aligned} &\mu_{p_k \to m^{(k)}}(1) \\ &= \frac{1}{2} \sum_{y_k, d_{1,k}} \mathrm{I}\left[y_k \in \{s_{1,1}, s_{1,2}\}\right] \\ &\qquad \cdot \mathrm{I}\left[g_1\left(d_{1,k}\right) = y_k\right] \cdot \mu_{y_k \to p_k}(y_k) \\ &= \frac{1}{2}\Bigg(\exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{1,1}\right|\right|^2\right) \\ &\qquad + \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{1,2}\right|\right|^2\right)\Bigg). \end{aligned} \quad (20)$$

Equivalent results are obtained for other modulations. Substituting a general form of these results into (16) produces

$$\mu_{eq \to m^{(0)}}(m) = \prod_{k=1}^{K} \frac{1}{S_m} \sum_{l=1}^{S_m} \exp\left(-\frac{1}{\sigma_N^2}\left|\left|r_k - s_{m,l}\right|\right|^2\right) \quad (21)$$

which is identical to (7) for ML classification and provides a verification of the factor graph. In successive iterations the messages $\mu_{p(\mathbf{d}|\mathbf{c}) \to \mathbf{d}}$ provide feedback from the decoder and are not necessarily uniform.

IV. Simulation Results

In order to demonstrate the iterative receiver we simulate the design for $M = 4$ with modulations BPSK, QPSK, 16QAM, and 64QAM. The highest order modulation (64QAM) requires $Z = 6$ blocks to be incorporated into the receiver structure. The encoder employs a $1/2$-rate convolutional code with a constraint length of $k_c = 7$ (64 state). The octal generators of the code are (133, 171) and provide the maximum free distance [15]. A zero tail is used to convert the convolutional code into a block code. The receiver's decoder utilizes a sum-product implementation of the BCJR algorithm. We consider frame lengths of 50, 100, and 150 symbols in order to determine the influence that the frame size has on the classifier's performance. The receiver performs 10 iterations of the sum-product algorithm. Fig. 5 displays the probability of correct modulation classification averaged over all modulations. The iterative receiver demonstrates a gain of 5 to 8 dB over ML classification in the traditional receiver.

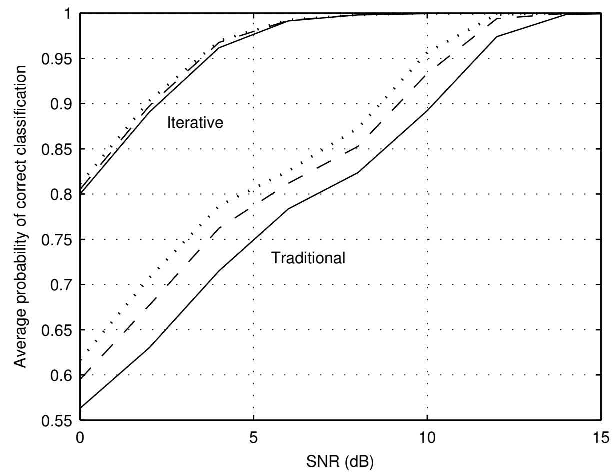

Fig. 5. Average classification performance of the iterative and traditional receivers with frame lengths of $K = 50$ (solid), $K = 100$ (dash), and $K = 150$ (dot) symbols.

At low SNR, the most significant gains in performance are for 16QAM which has the worst performance in the traditional receiver. In Table I, the confusion tables for the iterative and traditional receivers are given at an SNR of 2 dB. ML classification has a significant bias towards classifying 16QAM as 64QAM. The iterative receiver nearly eliminates this bias while improving overall classification.

TABLE I
CONFUSION TABLE FOR SNR = 2 dB, $K = 100$, AND 10 ITERATIONS

Iterative Receiver				
True\Hypothesis	BPSK	QPSK	16QAM	64QAM
BPSK	100	0	0	0
QPSK	0	99.07	0.47	0.46
16QAM	0	4.26	78.35	17.39
64QAM	0	4.23	14.17	81.59
Traditional Receiver				
True\Hypothesis	BPSK	QPSK	16QAM	64QAM
BPSK	100	0	0	0
QPSK	0	81.10	14.70	4.20
16QAM	0	19.49	36.57	43.94
64QAM	0	14.31	32.50	53.20

The end goal is to correctly detect the transmitted message. Improving the classification performance is useful if this translates to a reduction in the frame error rate (FER) or bit error rate (BER). For our purposes frame errors include both modulation classification errors and frame detection errors. In order to provide a reference point to the performance of the iterative and traditional receivers, we also simulate the

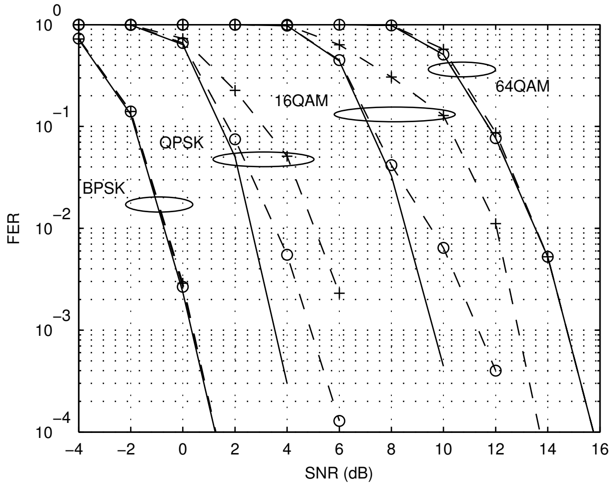

Fig. 6. FER performance by modulation scheme for known modulation (solid), the iterative receiver (dash, circle), and the traditional receiver (dash, plus) for a frame length of $K = 100$.

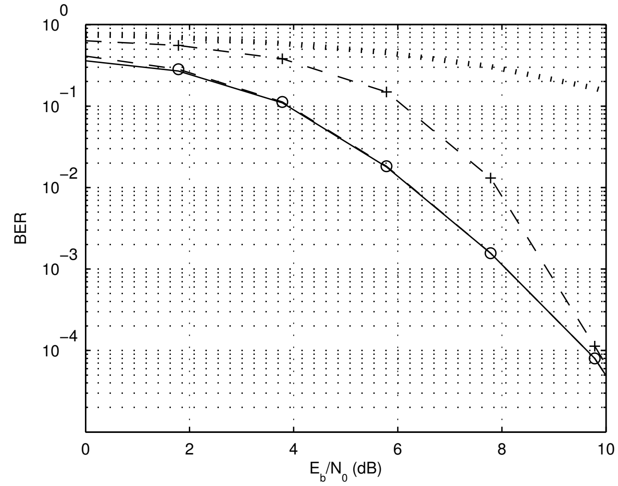

Fig. 7. BER performance of 64QAM for known modulation (solid), the iterative receiver (dash, circle), the traditional receiver (dash, plus), and uncoded transmission (dot) for a frame length of $K = 100$.

performance with known modulation. In all cases iterative demodulation and decoding is performed. Fig. 6 shows the FER of the iterative and traditional receivers. Improved classification performance in the iterative receiver does in fact lead to improvements in the FER of the system. For BPSK and 64QAM, the iterative receiver achieves ideal performance (that which is achieved with known modulation). For QPSK and 16QAM the system is limited by the classifier, but still improves upon the traditional receiver by as much as 3 dB.

When considering FER, the performance of the traditional receiver employing 64QAM is already very near the ideal performance. Therefore a comparison of BER performance versus E_b/N_0 is given in Fig. 7 where $E_b/N_0 = SNR \cdot 1/\log_2 S_m \cdot K/(K/2-6)$. For the purposes of simulating the BER we compare the transmitted sequence with the appropriate bits at the receiver. In terms of BER, the iterative receiver achieves a 2 dB gain over the traditional receiver for $E_b/N_0 < 8$ dB.

V. Conclusion

An iterative receiver for solving the problem of joint modulation classification, demodulation, and decoding is proposed. The iterative receiver is based on factor graphs and the associated sum-product algorithm. The performance of the iterative receiver is quantified with respect to the performance of a receiver which completes the task of modulation classification separately from demodulation and decoding. The iterative receiver is shown to provide greater than 5 dB gain in terms of modulation classification. By improving the reliability of automatic modulation classification, the proposed iterative receiver enables efficient use of the spectrum at low SNR.

References

[1] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proc. Int. Symp. Turbo Codes and Related Topics*, 1997, pp. 1–11.

[2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Commun., Geneva, Switzerland*, vol. 2, May 1993, pp. 1064–1070.

[3] H. Wymeersch, H. Steendam, H. Bruneel, and M. Moeneclaey, "Code-aided frame synchronization and phase ambiguity resolution," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2747–2757, July 2006.

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

[5] G. Caire, A. Tulino, and E. Biglieri, "Iterative multiuser joint detection and parameter estimation: a factor-graph approach," in *Proc. IEEE Inform. Theory Workshop*, 2001, pp. 36–38.

[6] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.

[7] J. Dauwels and H. A. Loeliger, "Phase estimation by message passing," in *Proc. IEEE Int. Conf. Commun.*, vol. 1, June 2004, pp. 523–527.

[8] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "A theoretical framework for soft-information-based synchronization in iterative (turbo) receivers," *EURASIP J. Wireless Commun. Netw.*, vol. 2005, no. 2, pp. 117–129, Apr. 2005.

[9] H. Wymeersch, *Iterative Receiver Design*. Cambridge Univ. Press, 2007.

[10] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.

[11] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 189–193, Feb. 2000.

[12] H. A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.

[13] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," in *Proc. IEEE Int. Conf. Commun., Montreal, Canada*, vol. 3, June 1997, pp. 1463–1467.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice-Hall, 1998.

[15] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. Boston:McGraw-Hill, 2008.