

A Factor-Graph Approach to Joint OFDM Channel Estimation and Decoding in Impulsive Noise Channels

Marcel Nassar*, Phil Schniter†, and Brian L. Evans*

* Wireless Networking and Communications Group, The University of Texas at Austin

† Department of Electrical and Computer Engineering, The Ohio State University

Email: mnassar@utexas.edu, schniter@ece.ohio.edu, bevans@ece.utexas.edu

Abstract—We propose a factor-graph-based approach to joint channel, impulse, symbol and bit estimation (JCISB) of LDPC-coded orthogonal frequency division multiplexing (OFDM) systems in impulsive noise environments. Impulsive noise arises in many modern wireless and wireline communication systems, such as cellular LTE and powerline communications, due to uncoordinated interference that is much stronger than thermal noise. Our receiver merges prior knowledge of the impulsive noise models with the recently proposed “generalized approximate message passing” (GAMP) algorithm, and soft-input soft-output decoding through the sum-product framework. Unlike the prior work, we explicitly consider channel estimation in the problem formulation. For N subcarriers, the resulting receiver has a complexity of $\mathcal{O}(N \log N)$, comparable to a typical DFT receiver. Numerical results indicate that the proposed receiver outperforms all prior impulsive noise OFDM decoders with improvements that reach 13dB when compared to the commonly used DFT receiver.

I. INTRODUCTION

Extensive measurement campaigns done in frequency bands occupied by many modern cellular wireless standards and PLC standards show that the additive noise deviates from the well known “additive white Gaussian noise” (AWGN) scenario and is in fact *impulsive* [1], [2]. In addition to multipath propagation, this noise is the main impairment for reliable communication in such systems. In this paper, we restrict our attention to systems employing coded orthogonal frequency division multiplexing (OFDM) modulation [3], as used in many modern cellular wireless standards (e.g., IEEE802.11n and LTE) and PLC standards (e.g., PRIME and IEEE1901). Coded OFDM is advantageous in that it facilitates data communication across convolutive multipath channels with high spectral efficiency and low complexity.

In classical OFDM receivers, tone-by-tone demodulation is optimal with AWGN and perfect channel estimates [3], and is highly desirable from a complexity standpoint, since it leaves the DFT as the primary source of receiver complexity, i.e., consuming $\mathcal{O}(N \log N)$ multiplies for N tones. When the time-domain noise is impulsive, however, the corresponding frequency-domain noise samples will be highly correlated, and tone-by-tone demodulation is no longer optimal.

Prior strategies for OFDM receiver design in impulsive noise generally take a *decoupled* and *suboptimal* approach to the problem of channel and impulse noise estimation and data decoding. In one popular approach, the time-domain received signal is pre-processed via clipping or blanking techniques [4] or (nonlinear) MMSE estimation [5], and the result passed to a conventional DFT receiver for decoding. While agreeable from a complexity standpoint, these techniques perform relatively poorly because OFDM signal structure is not exploited for noise mitigation. In an attempt to improve performance, it has been suggested to iterate between temporal denoising and OFDM decoding, but the approaches suggested to date (e.g., [6]) have shown limited success, mainly because the adaptation of temporal denoising with each iteration is challenging and often done in an ad-hoc manner. Another popular approach models the time-domain impulsive noise sequence as a sparse vector and then uses sparse-reconstruction techniques to estimate this sequence from the observed OFDM null and pilot (i.e., known) tones (e.g., compressive-sensing in [7] and sparse Bayesian learning (SBL) [8]). The recovered impulse vector is then subtracted from the time-domain received signal, and the result is passed to a conventional DFT receiver for decoding. Even with these techniques, there is room for improvement because channel estimation was not performed jointly with noise estimation and symbol detection and because decoding was integrated in an ad-hoc manner.

In this paper, we propose a novel OFDM receiver that performs near-optimally in the presence of impulsive noise while maintaining the $\mathcal{O}(N \log N)$ complexity order of the conventional N -tone OFDM receiver. Our approach is based on *jointly* estimating the impulsive noise and the channel gains while decoding the bits. In doing so, we aim to exploit all available information about the null tones, the pilot tones, the finite-alphabet symbol constellation, the code structure, the noise distribution, and the channel distribution. To maintain a low complexity, we leverage recent work on “generalized approximate message passing” (GAMP) [9], its “turbo” extension to larger factor graphs [10], and belief-propagation-based soft-input/soft-output (SISO) decoding [11]. The receiver we propose can be categorized as a factor-graph-based receiver, and in particular an extension of [12], that explicitly addresses the presence of impulsive noise. The resulting receiver pro-

M. Nassar and B. L. Evans were supported by National Instruments and Semiconductor Research Corporation under SRC GRC Task 1836.063, and P. Schniter was supported in part by the National Science Foundation under Grant CCF-1018368.

vides a flexible performance vs. complexity tradeoff and can be parallelized, making it suitable for FPGA implementations.

II. SYSTEM MODEL

A. Coded OFDM Model

We consider an OFDM system with N tones partitioned into N_p pilot tones (indexed by the set \mathcal{P}), N_n null tones (indexed by the set \mathcal{N}), and N_d data tones (indexed by the set \mathcal{D}) each modulated by a scalar symbol from an 2^M -ary constellation \mathbb{S} . The data bits (which determine the data symbols) are generated by encoding M_i information bits using a rate- R coder, interleaving them, and allocating the resulting $M_c = M_i/R$ bits among an integer number $Q = \lceil M_c/N_d M \rceil$ of OFDM symbols.

In the sequel, we use $S^{(i)} \in \mathbb{S}$ for $i \in \{1, \dots, 2^M\}$ to denote the i th element of \mathbb{S} , and $\mathbf{c}^{(i)} = [c_1^{(i)}, \dots, c_M^{(i)}]^T$ to denote the corresponding bits as defined by the symbol mapping. Likewise, we use $S_k[q]$ to denote the scalar symbol transmitted on the k th tone of the q th OFDM symbol. Based on the tone partition, we note that: $S_k[q] = \mathbf{p}$ for all $k \in \mathcal{P}$, where \mathbf{p} is a known pilot symbol; $S_k[q] = 0$ for all $k \in \mathcal{N}$; and $S_k[q] = S^{(l)}$ for some l such that $\mathbf{c}_k[q] = \mathbf{c}^{(l)}$ for all $k \in \mathcal{D}$, where $\mathbf{c}_k[q] = [c_{k,1}[q], \dots, c_{k,M}[q]]^T$ denotes the coded/interleaved bits corresponding to $S_k[q]$. On the frame level, we use $\mathbf{c}[q]$ to denote the coded/interleaved bits allocated to the data tones of the q th OFDM symbol, and $\mathbf{c} = [\mathbf{c}[1], \dots, \mathbf{c}[Q]]$ to denote the entire codeword obtained from the information bits $\mathbf{b} = [b_1, \dots, b_{M_i}]^T$ by coding/interleaving. Similarly, we use $\mathbf{S}[q] = [S_0[q], \dots, S_{N-1}[q]]^T$ to denote the q th OFDM symbol's tone vector, including pilot, null, and data tones.

For modulation, an inverse of the unitary N -point discrete Fourier transform (IDFT) matrix \mathbf{F} is applied to the q th OFDM symbol's tone vector $\mathbf{S}[q]$, producing the time-domain sequence $\mathbf{F}^* \mathbf{S}[q] = \mathbf{s}[q] = [s_0[q], \dots, s_{N-1}[q]]^T$, to which a cyclic prefix is prepended. The resulting sequence propagates through an L -tap linear-time-invariant Rayleigh-fading channel with impulse response $\mathbf{h}[q] = [h_0[q], \dots, h_{L-1}[q]]^T$ before being corrupted by both AWGN and impulsive noise.

Assuming a cyclic prefix of length $\geq L - 1$, inter-symbol interference is avoided by simply discarding the cyclic prefix at the receiver, after which the remaining N samples are

$$\mathbf{y}[q] = \mathbf{H}[q]\mathbf{s}[q] + \mathbf{n}[q] = \mathbf{H}[q]\mathbf{F}^* \mathbf{S}[q] + \mathbf{n}[q] \quad (1)$$

where $\mathbf{n}[q]$ is the time-domain noise vector and $\mathbf{H}[q]$ is channel matrix. Applying a DFT, each transmitted tone $S_k[q]$ experiences a flat scalar subchannel, since

$$Y_k[q] = H_k[q]S_k[q] + N_k[q], \quad \forall k \in \{0, \dots, N-1\} \quad (2)$$

where $\mathbf{H}[q] = \sqrt{N}\mathbf{F}_{:,1:L}\mathbf{h}[q]$ is the frequency-domain channel vector, $\mathbf{N}[q] = \mathbf{F}\mathbf{n}[q]$ is the frequency-domain noise vector.

B. Impulsive Noise Models

The *Gaussian mixture* (GM) model is a commonly used model that captures the impulsiveness of interference in many wireless and power-line communication (PLC) systems [13], [14]. Since our message-passing receiver is inherently

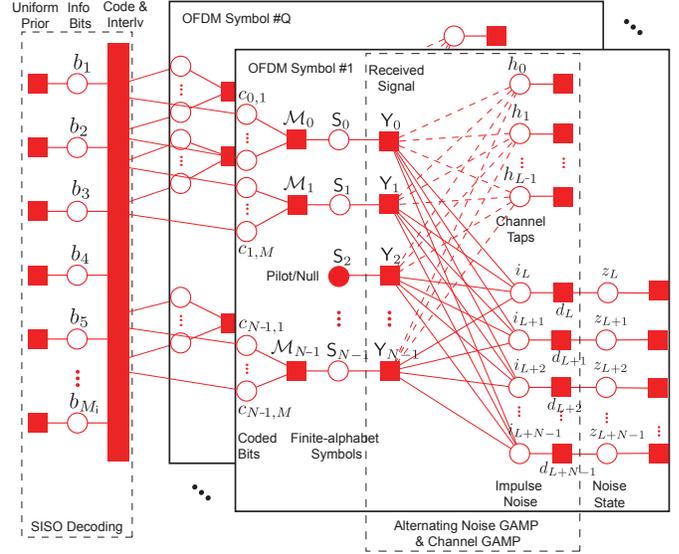


Fig. 1. Factor graph representation of a coded data frame spanning Q OFDM symbols. Round open circles denote random variables, round solid circles denote deterministic variables (e.g., known pilots or nulls), and solid squares denote pdf factors. The large rectangle on the left represents the coding-and-interleaving subgraph, whose details are immaterial. The time-domain impulse-noise quantities i_t and z_t start at time $t = L$ due to the use of an $L-1$ -length cyclic prefix.

Bayesian, the GM model provides us with natural priors on the impulsive noise. In particular, we decompose a given time-domain noise sample $n_t = g_t + i_t$ into a Gaussian background component $g_t \sim \mathfrak{N}(0, \gamma^{(0)})$ and a sparse impulsive component i_t with Bernoulli-GM pdf

$$p(i_t) = \pi^{(0)} \delta(i_t) + \sum_{k=1}^{K-1} \pi^{(k)} \mathfrak{N}(i_t; 0, \gamma^{(k)}) \quad (3)$$

where $\delta(\cdot)$ denotes the Dirac delta and $\sum_{k=0}^{K-1} \pi^{(k)} = 1$. Equivalently, we could model the (hidden) mixture state $z_t \in \{0, \dots, K-1\}$ of the impulsive component i_t as a random variable, giving rise to the hierarchical model

$$p(i_t | z_t = k) = \mathfrak{N}(i_t; 0, \gamma^{(k)}) \quad (4a)$$

$$P(z_t = k) = \pi^{(k)}. \quad (4b)$$

III. JOINT CHANNEL/NOISE ESTIMATION AND DECODING

MAP-optimal decoding of the information bits b_m involves marginalizing over the finite-alphabet symbols \mathbf{S} , coded bits \mathbf{c} , noise states \mathbf{z} , noise samples \mathbf{i} , channel taps \mathbf{h} , and other info bits $\mathbf{b}_{\setminus m}$. Clearly, direct evaluation of the posterior bit probabilities, represented by the factor graph in Fig. 1, is computationally intractable due to the high-dimensional integrals involved. We now propose a strategy to approximate the information bit posteriors by iterating (an approximation of) the SPA on the loopy factor graph in Fig. 1. In particular, we utilize an approximate version of SPA called *generalized approximate message passing* (GAMP) algorithm given in [9]. The GAMP algorithm addresses the estimation of a vector of independent possibly-non-Gaussian random variables \mathbf{x} that

are linearly mixed via a linear transform $\Phi \in \mathbb{C}^{M \times N}$ to form $\mathbf{z} = \Phi \mathbf{x} = [z_1 \cdots z_M]^T$ and subsequently observed as $\mathbf{y} = [y_1 \cdots y_M]^T$ according to the general likelihood function $p(\mathbf{y}|\mathbf{z}) = \prod_{i=1}^M p(y_i|z_i)$. We denote the application of GAMP to this problem as ‘‘GAMP($\mathbf{y}, \mathbf{z}, \Phi, \mathbf{x}$)’’.

To distinguish our approach from others in the literature, we will refer to it as ‘‘joint channel, impulse, symbol, and bit estimation’’ (JCISB). The following sections describe a particular message schedule of JCISB. We refer to a full pass over this schedule as a ‘‘turbo’’ iteration.

A. Bits to Symbols

Beliefs about the coded bits $\{c_{k,m}\}_{m=1}^M$ (for each data tone $k \in \mathcal{D}$) are first passed through the symbol mapping factor node \mathcal{M}_k . The SPA dictates that

$$\begin{aligned} \mu_{\mathcal{M}_k \rightarrow \mathcal{S}_k}(\mathcal{S}^{(i)}) &= \sum_{\mathbf{c}_k \in \{0,1\}^M} P(\mathcal{S}^{(i)}|\mathbf{c}_k) \prod_{m=1}^M \mu_{c_{k,m} \rightarrow \mathcal{M}_k}(c_{k,m}) \\ &= \prod_{m=1}^M \mu_{c_{k,m} \rightarrow \mathcal{M}_k}(c_m^{(i)}) \end{aligned} \quad (5)$$

where (5) follows from the deterministic symbol mapping $P(\mathcal{S}^{(i)}|\mathbf{c}^{(j)}) = \delta_{i-j}$. The resulting message is then copied forward through the \mathcal{S}_k node, i.e., $\mu_{\mathcal{S}_k \rightarrow \mathcal{Y}_k} = \mu_{\mathcal{M}_k \rightarrow \mathcal{S}_k}$, also according to the SPA. Note that, at the start of the first turbo iteration, we have no knowledge of the bits and thus we take $\mu_{c_{k,m} \rightarrow \mathcal{M}_k}(c)$ to be uniform across $c \in \{0,1\}$ for all m, k .

B. Channel Estimation

The next step in our message-passing schedule is to pass messages between the factor nodes $\{\mathcal{Y}_k\}$ and the time-domain channel nodes $\{h_l\}$. According to the SPA, the message passed from \mathcal{Y}_k to h_l is

$$\begin{aligned} \mu_{\mathcal{Y}_k \rightarrow h_l}(h_l) &= \sum_{\mathcal{S}_k} \int_{\mathbf{i}, \mathbf{h} \setminus i} p(\mathcal{Y}_k|\mathcal{S}_k, \mathbf{h}, \mathbf{i}) \mu_{\mathcal{S}_k \rightarrow \mathcal{Y}_k}(\mathcal{S}_k) \\ &\quad \times \prod_{z \neq l} \mu_{h_z \rightarrow \mathcal{Y}_k}(h_z) \prod_j \mu_{i_j \rightarrow \mathcal{Y}_k}(i_j). \end{aligned} \quad (6)$$

Exact evaluation of (6) involves an intractable high-dimensionality integration. Thus, we instead approximate the message passing between the $\{\mathcal{Y}_k\}$ and $\{h_l\}$ nodes using GAMP. To do this, we temporarily treat the messages $\{\mu_{\mathcal{S}_k \rightarrow \mathcal{Y}_k}\}$ and $\{\mu_{i_j \rightarrow \mathcal{Y}_k}\}$ as fixed, allowing us to employ ‘‘GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$)’’.

The application of GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$) requires us to specify the likelihood $p(\mathcal{Y}_k|\mathbf{H}_k)$ relating the transform output \mathbf{H}_k to the corresponding observed output \mathcal{Y}_k . From Fig. 1, we see that there are two types of belief flowing into each \mathcal{Y}_k node (apart from beliefs about $\{h_l\}$) that determine this likelihood: beliefs about the symbols $\{\mathcal{S}_k\}$, which we parameterize as $\beta_k = [\beta_k^{(1)}, \dots, \beta_k^{(|\mathcal{S}|)}]$ with $\beta_k^{(i)} = \mu_{\mathcal{S}_k \rightarrow \mathcal{Y}_k}(\mathcal{S}^{(i)})$, and beliefs about the frequency-domain impulsive noise $\{\mathbf{l}_k\}$, which GAMP approximates as $\mathfrak{N}(\mathbf{l}_k; \hat{\mathbf{l}}_k, \gamma_k^{\mathbf{l}})$, where the values

$\{\hat{\mathbf{l}}_k, \gamma_k^{\mathbf{l}}\}$ were computed by GAMP($\mathbf{Y}, \mathbf{l}, \mathbf{F}, \mathbf{i}$) in the previous turbo iteration.¹ Here, $\mathbf{l} = \mathbf{F}\mathbf{i}$ refers to the impulsive component of the frequency-domain noise $\mathbf{N} = \mathbf{l} + \mathbf{G}$, with $\{\mathcal{G}_k\} \sim \text{i.i.d } \mathfrak{N}(0, \gamma^{(0)})$, so that

$$\mathbf{Y} = \mathbf{H} \circ \mathbf{S} + \mathbf{l} + \mathbf{G}. \quad (7)$$

From (2) and (7), the GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$) likelihood is

$$p(\mathcal{Y}_k|\mathbf{H}_k) = \begin{cases} \mathfrak{N}(\mathcal{Y}_k; \rho \mathbf{H}_k + \hat{\mathbf{l}}_k, \gamma_k^{\mathbf{l}} + \gamma^{(0)}) & k \in \mathcal{P} \\ \sum_{l=1}^{|\mathcal{S}|} \beta_k^{(l)} \mathfrak{N}(\mathcal{Y}_k; \mathcal{S}^{(l)} \mathbf{H}_k + \hat{\mathbf{l}}_k, \gamma_k^{\mathbf{l}} + \gamma^{(0)}) & k \in \mathcal{D} \end{cases} \quad (8)$$

with the corresponding ‘‘output MMSE estimation functions’’, required by GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$), $\mathbb{E}\{\mathbf{H}_k|\mathbf{Y}; \hat{\rho}, \gamma^{\rho}\}$ and $\mathbb{V}\{\mathbf{H}_k|\mathbf{Y}; \hat{\rho}, \gamma^{\rho}\}$ specified in Table I. (See [15] for derivations). GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$) also requires us to derive the ‘‘input MMSE estimation functions’’ $\mathbb{E}\{h_j|\mathbf{Y}, \hat{\rho}, \gamma^{\rho}\}$ and $\mathbb{V}\{h_j|\mathbf{Y}, \hat{\rho}, \gamma^{\rho}\}$. It is straightforward to show that the input MMSE estimation functions are $\mathbb{E}\{h_j|\mathbf{Y}, \hat{\rho}, \gamma^{\rho}\} = \nu_j \hat{\rho} / (\nu_j + \gamma^{\rho})$ and $\mathbb{V}\{h_j|\mathbf{Y}, \hat{\rho}, \gamma^{\rho}\} = \nu_j \gamma^{\rho} / (\nu_j + \gamma^{\rho})$ [15]. After GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$) is iterated to convergence, it outputs $\{\hat{\mathbf{H}}_k\}$ and $\{\gamma_k^{\mathbf{H}}\}$ that are close approximations to the marginal posterior mean and variance, respectively, of $\{\mathbf{H}_k\}$. These outputs will be used in the next step of the message-passing schedule, as described below. Similarly, GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$) also outputs $\{\hat{h}_l\}$ and $\{\gamma_l^h\}$ that are close approximations to the marginal posterior mean and variance, respectively, of $\{h_l\}$.

C. Noise Estimation

The next step in our schedule is to pass messages between the factor nodes $\{\mathcal{Y}_k\}$, the time-domain impulse-noise nodes $\{i_t\}$, and the noise-state nodes $\{z_t\}$. According to the SPA, the message passed from \mathcal{Y}_k to i_t is

$$\begin{aligned} \mu_{\mathcal{Y}_k \rightarrow i_t}(i_t) &= \sum_{\mathcal{S}_k} \int_{\mathbf{i} \setminus i, \mathbf{h}} p(\mathcal{Y}_k|\mathcal{S}_k, \mathbf{h}, \mathbf{i}) \mu_{\mathcal{S}_k \rightarrow \mathcal{Y}_k}(\mathcal{S}_k) \\ &\quad \times \prod_l \mu_{h_l \rightarrow \mathcal{Y}_k}(h_l) \prod_{j \neq t} \mu_{i_j \rightarrow \mathcal{Y}_k}(i_j) \end{aligned} \quad (9)$$

which is intractable. By temporarily treating the messages $\{\mu_{d_t \rightarrow i_t}\}$, $\{\mu_{\mathcal{S}_k \rightarrow \mathcal{Y}_k}\}$, and $\{\mu_{h_l \rightarrow \mathcal{Y}_k}\}$ as fixed, we can apply GAMP($\mathbf{Y}, \mathbf{l}, \mathbf{F}, \mathbf{i}$) under the likelihood model

$$p(\mathcal{Y}_k|\mathbf{l}_k) = \begin{cases} \mathfrak{N}(\mathcal{Y}_k; \mathbf{l}_k, \gamma^{(0)}) & \text{if } k \in \mathcal{N} \\ \mathfrak{N}(\mathcal{Y}_k; \rho \hat{\mathbf{H}}_k + \mathbf{l}_k, \rho \gamma_k^{\mathbf{H}} + \gamma^{(0)}) & \text{if } k \in \mathcal{P} \\ \sum_{l=1}^{|\mathcal{S}|} \beta_k^{(l)} \mathfrak{N}(\mathcal{Y}_k; \mathcal{S}^{(l)} \hat{\mathbf{H}}_k + \mathbf{l}_k, \rho \gamma_k^{\mathbf{H}} + \gamma^{(0)}) & \text{if } k \in \mathcal{D} \end{cases} \quad (10)$$

implied by (2) and (7), and the coefficient prior

$$p(i_t) = \pi_t^{(0)} \delta(i_t) + \sum_{k=1}^{K-1} \pi_t^{(k)} \mathfrak{N}(i_t; 0, \gamma^{(k)}) \quad (11)$$

¹During the first turbo iteration, we use $\hat{\mathbf{l}}_k = 0$ and $\gamma_k^{\mathbf{l}} = \gamma^{\mathbf{l}} \forall k$.

TABLE I
GAMP OUTPUT MMSE ESTIMATION FUNCTIONS USED IN JCISB

Tone Type	GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$)	
	$\mathbb{E}\{\mathbf{H}_k \mathbf{Y}_k; \hat{\mathbf{p}}, \gamma^p\}$	$\mathbb{V}\{\mathbf{H}_k \mathbf{Y}_k; \hat{\mathbf{p}}, \gamma^p\}$
Pilot: $k \in \mathcal{P}$	$\hat{\mathbf{p}} + \gamma^p \mathbf{p}^*(\mathbf{Y}_k - \hat{\mathbf{I}}_k - \mathbf{p}\hat{\mathbf{p}})/(\gamma^{(0)} + \gamma_k^1 + \rho_p \gamma^p)$	$\gamma^p(\gamma^{(0)} + \gamma_k^1)/(\gamma^{(0)} + \gamma_k^1 + \rho_p \gamma^p)$
Data: $k \in \mathcal{D}$	$\hat{\mathbf{p}} + \sum_{l=1}^{ \mathcal{S} } \lambda_k^{(l)} \frac{\gamma^p \mathbf{S}^{*(l)}(\mathbf{Y}_k - \hat{\mathbf{I}}_k - \hat{\mathbf{p}}\mathbf{S}^{(l)})}{(\gamma^{(0)} + \gamma_k^1 + \mathcal{S}^{(l)} ^2 \gamma^p)}$ where $\lambda_k^{(l)} = p(\mathbf{Y}_k \mathcal{S}^{(l)})\beta_k^{(l)}/\sum_j p(\mathbf{Y}_k \mathcal{S}^{(j)})\beta_k^{(j)}$ and $p(\mathbf{Y}_k \mathcal{S}^{(l)}) = \mathfrak{N}(\mathbf{Y}_k; \hat{\mathbf{I}}_k + \hat{\mathbf{p}}\mathbf{S}^{(l)}, \gamma^{(0)} + \gamma_k^1 + \mathcal{S}^{(l)} ^2 \gamma^p)$	$\sum_{l=1}^{ \mathcal{S} } \lambda_k^{(l)} \left[\frac{\gamma^p(\gamma^{(0)} + \gamma_k^1)}{\gamma^{(0)} + \gamma_k^1 + \mathcal{S}^{(l)} ^2 \gamma^p} + \left \hat{\mathbf{p}} + \frac{\gamma^p \mathbf{S}^{*(l)}(\mathbf{Y}_k - \hat{\mathbf{I}}_k - \hat{\mathbf{p}}\mathbf{S}^{(l)})}{(\gamma^{(0)} + \gamma_k^1 + \mathcal{S}^{(l)} ^2 \gamma^p)} \right ^2 \right] - \mathbb{E}\{\mathbf{H}_k \mathbf{Y}_k; \hat{\mathbf{p}}, \gamma^p\} ^2$
Tone Type	GAMP($\mathbf{Y}, \mathbf{I}, \mathbf{F}, \mathbf{i}$)	
	$\mathbb{E}\{\mathbf{I}_k \mathbf{Y}_k; \hat{\mathbf{p}}, \gamma^p\}$	$\mathbb{V}\{\mathbf{I}_k \mathbf{Y}_k; \hat{\mathbf{p}}, \gamma^p\}$
Null: $k \in \mathcal{N}$	$(\gamma^p \mathbf{Y}_k - \gamma^{(0)} \hat{\mathbf{p}})/(\gamma^{(0)} + \gamma^p)$	$\gamma^p \gamma^{(0)}/(\gamma^{(0)} + \gamma^p)$
Pilot: $k \in \mathcal{P}$	$\hat{\mathbf{p}} + \gamma^p(\mathbf{Y}_k - \hat{\mathbf{p}} - \hat{\mathbf{H}}_k \mathbf{p})/(\gamma^{(0)} + \gamma^p + \rho_p \gamma_k^H)$	$\gamma^p(\gamma^{(0)} + \rho_p \gamma_k^H)/(\gamma^{(0)} + \gamma^p + \rho_p \gamma_k^H)$
Data: $k \in \mathcal{D}$	$\hat{\mathbf{p}} + \sum_{l=1}^{ \mathcal{S} } \lambda_k^{(l)} \frac{\gamma^p(\mathbf{Y}_k - \hat{\mathbf{p}} - \hat{\mathbf{H}}_k \mathbf{S}^{(l)})}{(\gamma^{(0)} + \gamma^p + \mathcal{S}^{(l)} ^2 \gamma_k^H)}$ where $\lambda_k^{(l)} = p(\mathbf{Y}_k \mathcal{S}^{(l)})\beta_k^{(l)}/\sum_j p(\mathbf{Y}_k \mathcal{S}^{(j)})\beta_k^{(j)}$ and $p(\mathbf{Y}_k \mathcal{S}^{(l)}) = \mathfrak{N}(\mathbf{Y}_k; \hat{\mathbf{p}} + \hat{\mathbf{H}}_k \mathbf{S}^{(l)}, \gamma^{(0)} + \gamma^p + \mathcal{S}^{(l)} ^2 \gamma_k^H)$	$\sum_{l=1}^{ \mathcal{S} } \lambda_k^{(l)} \left[\frac{\gamma^p(\gamma^{(0)} + \mathcal{S}^{(l)} ^2 \gamma_k^H)}{(\gamma^{(0)} + \gamma^p + \mathcal{S}^{(l)} ^2 \gamma_k^H)} + \left \hat{\mathbf{p}} + \frac{\gamma^p(\mathbf{Y}_k - \hat{\mathbf{p}} - \hat{\mathbf{H}}_k \mathbf{S}^{(l)})}{(\gamma^{(0)} + \gamma^p + \mathcal{S}^{(l)} ^2 \gamma_k^H)} \right ^2 \right] - \mathbb{E}\{\mathbf{I}_k \mathbf{Y}_k; \hat{\mathbf{p}}, \gamma^p\} ^2$

implied by (3). In (10), $\beta_k^{(i)} = \mu_{S_k \rightarrow Y_k}(\mathcal{S}^{(i)})$ are the symbol beliefs coming from the $\{S_k\}$ nodes and $\{\hat{H}_k, \gamma_k^H\}$ are the frequency-domain channel estimates previously calculated by GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$). Meanwhile, in (11), $\{\pi_t^{(k)}\}_{k=0}^{K-1}$ represents the pmf on the noise state z_t that is set as $\pi_t^{(k)} = \mu_{z_t \rightarrow d_t}(k)/(\sum_{l=0}^{K-1} \mu_{z_t \rightarrow d_t}(l))$. The resulting output MMSE estimation functions, derived in [15], are listed in TABLE I, and the input MMSE estimation functions are

$$\mathbb{E}\{i_t|\mathbf{Y}, \hat{\mathbf{r}}, \gamma^r\} = \sum_{k=0}^{K-1} \alpha_t^{(k)} \frac{\gamma^{(k)} \hat{\mathbf{r}}}{\gamma^{(k)} + \gamma^r} \quad (12)$$

$$\mathbb{V}\{i_t|\mathbf{Y}, \hat{\mathbf{r}}, \gamma^r\} = \sum_{k=0}^{K-1} \frac{\alpha_t^{(k)}}{\gamma^r + \gamma^{(k)}} \left(\gamma^r \gamma^{(k)} + \frac{|\gamma^{(k)} \hat{\mathbf{r}}|^2}{\gamma^r + \gamma^{(k)}} \right) - |\mathbb{E}\{i_t|\mathbf{Y}, \hat{\mathbf{r}}, \gamma^r\}|^2. \quad (13)$$

Here, $\{\alpha_t^{(k)}\}_{k=0}^{K-1}$ is the posterior pmf for noise-state z_t , with

$$\alpha_t^{(k)} = P(z_t = k|\hat{\mathbf{r}}) = \frac{p(\hat{\mathbf{r}}|z_t = k)\pi_t^{(k)}}{\sum_{l=0}^{K-1} p(\hat{\mathbf{r}}|z_t = l)\pi_t^{(l)}} \quad (14)$$

where $p(\hat{\mathbf{r}}|z_t = k) = \mathfrak{N}(\hat{\mathbf{r}}; 0, \gamma^r + \gamma^{(k)})$ is the noise state likelihood.

Using these input and output MMSE estimation functions, GAMP($\mathbf{Y}, \mathbf{I}, \mathbf{F}, \mathbf{i}$) is iterated until convergence, generating $\{\hat{I}_k, \gamma_k^I\}$ that are close approximations to the marginal posterior means and variances of $\{I_k\}$ and that will be used by GAMP($\mathbf{Y}, \mathbf{H}, \sqrt{N}\mathbf{F}, \mathbf{h}$) in the next turbo iteration. In addition, for each data tone $k \in \mathcal{D}$, GAMP($\mathbf{Y}, \mathbf{I}, \mathbf{F}, \mathbf{i}$) yields the leftward flowing soft symbol beliefs

$$\mu_{Y_k \rightarrow S_k}(\mathcal{S}) = \mathfrak{N}(Y_k; \hat{S}\hat{H}_k + \hat{I}_k, |\mathcal{S}|^2 \gamma_k^H + \gamma_k^I + \gamma^{(0)}) \quad (15)$$

that are subsequently used for decoding (as described below). Here, $\{\hat{H}_k, \gamma_k^H\}$ and $\{\hat{I}_k, \gamma_k^I\}$ play the role of soft frequency-domain channel and impulse-noise estimates, respectively.

D. Symbols to Bits

The SPA dictates that the messages flowing leftward through the symbol nodes $\{S_k\}$ come out unchanged, i.e., $\mu_{S_k \rightarrow \mathcal{M}_k} =$

$\mu_{Y_k \rightarrow S_k}$. Moreover, it dictates that the message flowing leftward out of the symbol-mapping node \mathcal{M}_k and into the coded-bit node $c_{k,m}$ takes the form

$$\begin{aligned} \mu_{\mathcal{M}_k \rightarrow c_{k,m}}(c) &= \sum_{l=1}^{|\mathcal{S}|} \sum_{\mathbf{c}_k \setminus c_{k,m}} P(\mathcal{S}^{(l)}|\mathbf{c}_k) \mu_{S_k \rightarrow \mathcal{M}_k}(\mathcal{S}^{(l)}) \\ &\times \prod_{m' \neq m} \mu_{c_{k,m'} \rightarrow \mathcal{M}_k}(c_{k,m'}) \quad (16) \\ &= \frac{\sum_{l: c_{k,m} = c} \mu_{S_k \rightarrow \mathcal{M}_k}(\mathcal{S}^{(l)}) \mu_{\mathcal{M}_k \rightarrow S_k}(\mathcal{S}^{(l)})}{\mu_{c_{k,m} \rightarrow \mathcal{M}_k}(c)} \quad (17) \end{aligned}$$

where the last step was is derived in [12].

Finally, the computed coded-bit beliefs are passed to the coding/interleaving factor node. This can be viewed as passing (extrinsic) soft information into a soft-input/soft-output (SISO) decoder, where it is treated as prior information for decoding according to the ‘‘turbo’’ principle. SISO decoding has been studied extensively and we refer the interested reader to [11] for a detailed account. After SISO decoding terminates, it will produce extrinsic soft information, in the form of beliefs $\{\mu_{c_{k,m} \rightarrow \mathcal{M}_k}\}$, that will be passed rightward to the symbol-mapping nodes at the start of the next turbo iteration. The turbo iterations are terminated after either the decoder detects no bit errors, the beliefs $\{\mu_{c_{k,m} \rightarrow \mathcal{M}_k}\}$ have converged, or a maximum number of turbo iterations has elapsed.

IV. NUMERICAL RESULTS

We consider two N -tone OFDM systems: (1) $N = 256$ (80 nulls and 15 pilots) with 4-QAM modulation under a 5-tap Rayleigh channel (PLC), and (2) $N = 1024$ (150 pilots) with 16-QAM under a 10-tap Rayleigh fading channel. This system is corrupted by iid GM noise having two impulsive components with powers 20dB and 30dB above the background noise occurring 7% and 3% of the time, respectively. The signal to noise ratio (SNR) refers to the ratio of the signal power to the second order moment of the noise. We fix the number

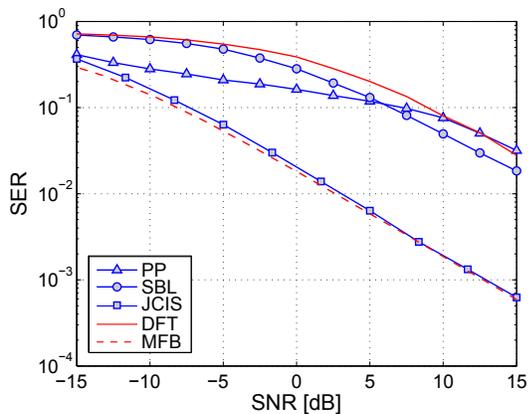


Fig. 2. Uncoded SER of system configuration (1)

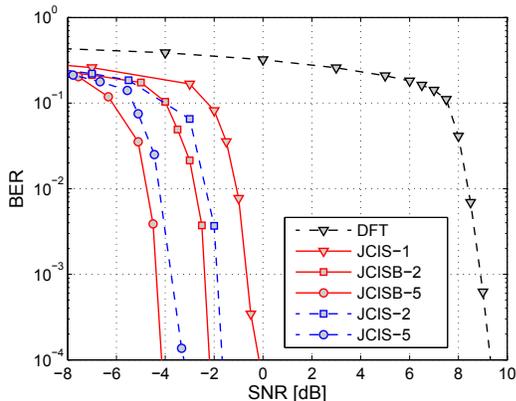


Fig. 3. LDPC coded BER for system configuration (2).

of iterations to 5 turbo iterations, 15 GAMP iterations, and a maximum of 50 LDPC iterations.

A. Comparison With Other Schemes

Fig. 2 shows the uncoded symbol error rate (SER) comparison of our proposed JCISB framework with the conventional “DFT” [3], “PP” [5], and “SBL” [8] receivers for system configuration (1). Since this is an uncoded transmission, we focus on symbol detection and JCIS reduces to “Joint channel, impulse, and symbol detection” (JCIS). By utilizing all the tones and explicitly accounting for channel estimation, the proposed JCIS receiver outperforms the DFT receiver by 15dB, the SBL receiver by 11dB, and the MMSE receiver by 7dB in the low SNR regime and by 15dB in the high SNR regime. Furthermore, the proposed JCIS receiver is within 1dB from the matched filter “MFB” lower bound.

B. Coded Systems

Fig. 3 shows the bit error rate (BER) performance of system configuration (2). The information bits are coded using an LDPC code with code-word length $\sim 60\,000$ and rate $1/2$. The label “alg-#” refers to the algorithm used by the receiver followed by the number of turbo iterations performed. We consider three types of receivers: the DFT receiver followed

by LDPC decoding; the proposed JCISB; and JCIS followed by LDPC decoding (note JCISB-1 is equivalent to JCIS-1). With only 1 turbo iteration, JCIS (also JCISB) provides an additional 9dB over the coded DFT receiver. An additional turbo iteration yields an extra 2dB while 5 turbo iterations in total yield a 13dB improvement over the DFT receiver. Furthermore, by decoding the LDPC code in each turbo iteration, JCISB provides an additional 1dB over JCIS.

V. CONCLUSION

In this paper, we presented a factor-graph approach to joint channel/noise estimation and data decoding in impulsive noise channels. Our approach merges recent work on modeling impulsive noise in communication systems [13] with recent advances in approximate message passing algorithms [9] and SISO decoding [11]. The presented receiver has a complexity comparable to the typical DFT receiver while providing tens of dBs in performance gain (1dB from a lower bound).

REFERENCES

- [1] K. Blackard, T. Rappaport, and C. Bostian, “Measurements and models of radio frequency impulsive noise for indoor wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 11, no. 7, pp. 991–1001, 1993.
- [2] M. Zimmermann and K. Dostert, “Analysis and modeling of impulsive noise in broad-band powerline communications,” *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 1, pp. 249–258, 2002.
- [3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [4] S. Zhidkov, “Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers,” *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 5–9, 2008.
- [5] J. Haring, *Error Tolerant Communication over the Compound Channel*. Shaker-Verlag, 2002.
- [6] J. Haring and A. Vinck, “Iterative decoding of codes over complex numbers for impulsive noise channels,” *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1251–1260, 2003.
- [7] G. Caire, T. Al-Naffouri, and A. Narayanan, “Impulse noise cancellation in OFDM: an application of compressed sensing,” in *IEEE Int. Symp. on Inf. Theory*, 2008, pp. 1293–1297.
- [8] J. Lin, M. Nassar, and B. L. Evans, “Impulsive noise mitigation in powerline communications using sparse Bayesian learning,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, Jul. 2013.
- [9] S. Rangan, “Generalized approximate message passing for estimation with random linear mixing,” in *IEEE Int. Symp. on Inf. Theory*, 2011, pp. 2174–2178, (See also longer version in *Arxiv:1010.5141*).
- [10] P. Schniter, “Turbo reconstruction of structured sparse signals,” in *Proc. Conf. on Info. Sciences and Systems*, 2010, pp. 1–6.
- [11] D. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [12] P. Schniter, “A message-passing receiver for BICM-OFDM over unknown clustered-sparse channels,” *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1462–1474, 2011.
- [13] K. Gulati, B. L. Evans, J. Andrews, and K. Tinsley, “Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers,” *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6207–6222, 2010.
- [14] M. Nassar, K. Gulati, Y. Mortazavi, and B. L. Evans, “Statistical modeling of asynchronous impulsive noise in powerline communication networks,” *IEEE Global Commun. Conf.*, pp. 1–6, 2011.
- [15] M. Nassar, P. Schniter, and B. Evans, “A factor graph approach to joint OFDM channel estimation and decoding in impulsive noise environments,” *IEEE Trans. Signal Process.*, 2014, to appear.