Non-Parametric Impulsive Noise Mitigation in OFDM Systems Using Sparse Bayesian Learning

Jing Lin, Marcel Nassar and Brian L. Evans
The University of Texas at Austin, Austin, Texas 78712 USA
Email: {linj, nassar.marcel}@mail.utexas.edu, bevans@ece.utexas.edu

***Abstract*—Additive asynchronous impulsive noise limits communication performance in certain OFDM systems, such as powerline communications, cellular LTE and 802.11n systems. Under additive impulsive noise, the fast Fourier transform (FFT) in the OFDM receiver introduces time-dependence in the subcarrier noise statistics. As a result, complexity of optimal detection becomes exponential in the number of subcarriers. Many previous approaches assume a statistical model of the impulsive noise and use parametric methods in the receiver to mitigate impulsive noise. Parametric methods require the overhead of training and parameter estimation, and degrade due to model mismatch. In this paper, we apply sparse Bayesian learning techniques to estimate and mitigate impulsive noise in OFDM systems without the need for training. We propose two non-parametric iterative algorithms: (1) estimate impulsive noise by its projection onto null and pilot tones so that the OFDM symbol is recovered by subtracting out the impulsive noise estimate; and (2) jointly estimate the OFDM symbol and impulsive noise utilizing information on all tones. In our simulations, the estimators achieve 5dB and 10dB SNR gains in communication performance respectively, as compared to conventional OFDM receivers.**

I. Introduction

Additive noise in powerline communications (PLC) and wireless networks are more complicated than the general assumption of additive white Gaussian noise (AWGN). In particular, PLC receivers in the frequency range from several hundred kHz up to 20 MHz are primarily affected by impulsive noise, which consists of short bursts of noise with high power spectral density and random occurrence [1]. Similarly, the ever increasing demand for data drives wireless networks toward dense spatial reuse of available spectrum leading to severe co-channel interference which is impulsive in nature [2].

Orthogonal frequency-division multiplexing (OFDM) has been adopted in many modern wireless communication standards (e.g. IEEE802.11n and LTE) and recent PLC standards (e.g. PRIME and G3). It offers great advantages in combating multipath and increasing data rates. In addition, [3] showed that OFDM is more resilient to impulsive noise than single carrier systems. At the receiver, taking the discrete Fourier transform (DFT) of the received signal spreads out the impulsive energy across all tones (Fig. 1). As the number of tones becomes large, the resulting noise in each tone tends toward being Gaussian with increased variance due to the impulse energy (by central limit theorem), and hence conventional data detectors can be applied. If the impulsive noise power is moderate, this technique will provide significant gain over single carrier systems [4]. However, in highly impulsive environments with impulsive noise power at least 20dB above background noise, such as PLC and dense cellular networks, the conventional OFDM receivers will encounter dramatic performance degradation. Therefore, in this work we propose impulsive noise mitigation algorithms that can be integrated into a conventional OFDM receiver as shown in Fig. 2.

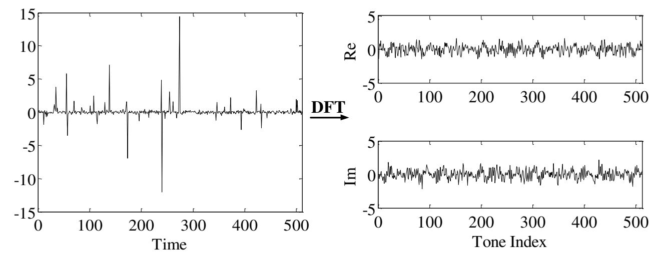

Fig. 1. DFT spreads out the time-domain impulsive energy over all tones.

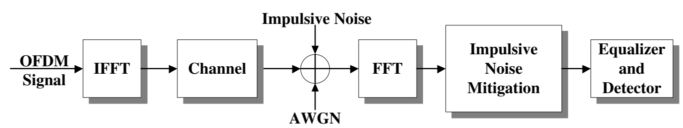

Fig. 2. The OFDM receiver that explicitly mitigates impulsive noise.

Various parametric and non-parametric methods have been proposed to mitigate the effect of impulsive noise in OFDM systems. In parametric methods, the receiver assumes a particular statistical model for impulsive noise. Parametric methods estimate the parameters of the noise distribution through a training stage (i.e. they are non-blind). Middleton's Class A, Symmetric Alpha-Stable model and Gaussian mixture distributions are commonly used to characterize impulsive noise statistically [2], [5]. These models are based on the fact that impulsive noise is an aggregation of electromagnetic radiation events [6], [7]. Pre-filtering techniques based on the Alpha-Stable model has been proposed in [2]; and [8] derived two suboptimal maximum likelihood receivers for a 2×2 MIMO system, assuming Middleton's Class A model. While low in complexity, their performance can deteriorate significantly

This work was supported by the Global Research Collaboration Program of the Semiconductor Research Corporation under Task Id 1836.063.

as the constellation size increases. More promising results are proposed in [3], where two minimum mean-squared-error (MMSE) estimators were derived to recover OFDM symbols under Gaussian mixture modeled impulsive noise. Additionally, in [9] a turbo-based iterative decoding algorithm was introduced to cope with Middleton's Class A impulsive noise. The advantage of parametric methods is that they lead to performance gains by exploiting information of the noise model and its parameters. However, their performance can suffer when the noise model does not reflect the statistics of the impulsive noise or the parameters are inaccurate. On top of that, they require extra complexity due to the training sequences used for parameter estimation.

Non-parametric methods, on the other hand, require no additional training or parameter estimation since they do not assume any particular noise model. This makes them more robust to various environments and more resilient to model mismatches. In [10] the authors exploited the fact that the N-point IDFT is essentially an order-N polynomial code, which is redundant and error-tolerant given that the OFDM symbol contains unused subcarriers (a.k.a. null tones). The decoder by algebraic polynomial curve fitting can successfully recover the OFDM symbol if it is contaminated by only a few impulses. However, the presence of AWGN over all tones can cause severe decoding error. Taking advantage of the sparse nature of the time-domain impulsive noise, [11] developed an impulsive noise removal algorithm by applying recent advances in compressed sampling. The support of impulsive noise was determined by solving a compressed sampling problem. Based on that a least-squares (LS) or an MMSE problem was formulated to estimate the values of the impulsive noise. The algorithm was subject to a recovery sufficient condition that the number of impulses within an OFDM symbol does not exceed a threshold that is uniquely determined by the DFT size and the number of null tones. However, for common OFDM system settings, the threshold turns out to be too restrictive for many impulsive noise environments where an OFDM symbol is subjected to multiple impulses during its duration.

In this work, we assume the presence of impulsive noise, and seek to develop non-parametric estimation algorithms that are generally applicable to all impulsive noise models, especially in moderately and highly impulsive noise environments. Towards this end, we extend the compressed sensing algorithm proposed in [11] to a sparse Bayesian learning (SBL) approach [12] and propose two non-parametric algorithms that can mitigate the effects of impulsive noise in OFDM systems.

II. Impulsive Noise Enviroments

The main statistical-physical models for impulsive noise are the Gaussian mixture, the Middleton's Class A, and the Symmetric Alpha-Stable. These models have been shown to accurately model various impulsive noise scenarios both in PLC and wireless communications systems [1], [7]. Although the proposed methods are non-parametric to avoid expensive training and provide better generalization to different environments, these models are useful for simulating various impulsive noise scenarios to test our algorithms. Hence, we briefly discuss these models.

A random variable Z has a Gaussian mixture distribution if its probability density function (pdf) is a weighted summation of different Gaussian distributions, i.e.

$$f_Z(z) = \sum_{k=0}^{K} \pi_k \cdot \mathcal{N}(z; 0, \gamma_k), \tag{1}$$

where $\mathcal{N}(z; 0, \gamma_k)$ denotes a Gaussian pdf with zero mean and variance γ_k, and π_k is the mixing probability of the k-th Gaussian component. Middleton's Class A model, with an overlapping factor $A \in [10^{-2}, 1]$ and power ratio $\Gamma \in [10^{-6}, 1]$, can be considered as a special case of the Gaussian mixture distribution, with $\pi_k = e^{-A}\frac{A^k}{k!}$ and $\gamma_k = \frac{k/A+\Gamma}{1+\Gamma}$ as $K \to \infty$. In practice, only the first few significant terms are retained. On the other hand, a Symmetric Alpha-Stable (SαS) distribution is defined by its characteristic function

$$\Phi(\omega) = e^{j\delta\omega - \gamma|\omega|^\alpha}, \tag{2}$$

where $\alpha \in [0, 2]$ is the characteristic exponent measuring the thickness of the tail of the distribution, δ is the localization parameter which equals the mean when $1 \le \alpha \le 2$, and γ is the dispersion parameter that is analogous to variance. In general, SαS distribution does not have a closed form pdf.

III. System Model and Problem Formulation

Consider a baseband OFDM system with $N/2$ subcarriers. We denote an OFDM symbol as $\mathbf{x} \in \mathbb{C}^N$. $\mathbf{x}$ is conjugate symmetric (i.e. $\mathbf{x}_n = \mathbf{x}^*_{N-n}, \forall n \in \{0, 1, ..., N-1\}$) so that the transmitted signal after taking the IDFT of $\mathbf{x}$ is real-valued. After removing the cyclic prefix, which is assumed to be longer than the channel delay spread, the OFDM receiver takes the DFT of received signal resulting in

$$\mathbf{y} = \mathbf{FHF}^*\mathbf{x} + \mathbf{Fe} + \mathbf{Fn} = \mathbf{\Lambda x} + \mathbf{Fe} + \mathbf{g} \tag{3}$$

where $\mathbf{F}$ is the N-point DFT matrix, $\mathbf{H} \in \mathbb{R}^{N\times N}$ is the circulant convolution matrix of the channel, $\mathbf{e}$ and $\mathbf{n} \in \mathbb{R}^N$ represent impulsive noise and AWGN respectively, and $(\cdot)^*$ denotes complex conjugate. Since $\mathbf{H}$ is circulant, $\mathbf{\Lambda} \stackrel{\triangle}{=} \mathbf{FHF}^* = \operatorname{diag}\{H_1, \cdots, H_N\}$ where $\{H_i\}_{i=1}^N$ are the DFT coefficients of the channel response. On the other hand, $\mathbf{g} \stackrel{\triangle}{=} \mathbf{Fn}$ is *i.i.d* complex Gaussian (since it is the DFT of AWGN). Vector $\mathbf{y}$ is conjugate symmetric, and hence (3) can be equivalently written in real numbers as

$$\begin{bmatrix} \Re\{\dot{\mathbf{y}}\} \\ \Im\{\dot{\mathbf{y}}\} \end{bmatrix} = \begin{bmatrix} \Re\left\{\dot{\mathbf{\Lambda}}\dot{\mathbf{x}}\right\} \\ \Im\left\{\dot{\mathbf{\Lambda}}\dot{\mathbf{x}}\right\} \end{bmatrix} + \begin{bmatrix} \Re\left\{\dot{\mathbf{F}}\right\} \\ \Im\left\{\dot{\mathbf{F}}\right\} \end{bmatrix} \mathbf{e} + \begin{bmatrix} \Re\{\dot{\mathbf{g}}\} \\ \Im\{\dot{\mathbf{g}}\} \end{bmatrix} \tag{4}$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts respectively, and $\dot{(\cdot)}$ denotes the first $N/2$ entries (or rows) of a vector (or a matrix). Overloading the notation for clarity, (4) can be written as an $N \times N$ linear system in the form of

$$\mathbf{y} = \mathbf{Fe} + \mathbf{\Lambda x} + \mathbf{g} = \mathbf{Fe} + \mathbf{v}, \tag{5}$$

where $\mathbf{e}$ is the input, $\mathbf{F}$ is the system matrix, $\mathbf{y}$ is the observation, and $\mathbf{v} \triangleq \mathbf{\Lambda x} + \mathbf{g}$ is considered as disturbance. Therefore, for a given $\mathbf{\Lambda x}$, the disturbance $\mathbf{v}$ is Gaussian distributed, i.e. $\mathbf{v} \sim \mathcal{N}\left(\mathbf{\Lambda x}, \sigma^2 \mathbf{I}\right)$. The goal is to estimate the impulsive noise vector $\mathbf{e}$ and use that estimate $\hat{\mathbf{e}}$ to improve the detection of the symbol vector $\mathbf{x}$. The impulsive noise estimate is subtracted out from the received symbol to form a new decision metric given by

$$\hat{\mathbf{y}} = \mathbf{y} - \mathbf{F}\hat{\mathbf{e}} = \mathbf{\Lambda x} + \mathbf{g} + \mathbf{F}\left(\mathbf{e} - \hat{\mathbf{e}}\right). \tag{6}$$

We assume that $\hat{\mathbf{e}} \approx \mathbf{e}$ and proceed as if only Gaussian noise were present and apply the standard OFDM decoder. The sparsity of $\mathbf{e}$ motivates the SBL approach discussed next.

IV. Sparse Bayesian Learning

SBL solves an underdetermined linear regression problem under sparsity constraints using Bayesian learning. This problem is given by $\mathbf{t} = \mathbf{\Phi w} + \mathbf{z}$, where $\mathbf{t}$ is an observation vector, $\mathbf{\Phi}$ is an overcomplete basis, $\mathbf{w}$ is a sparse weight vector, and $\mathbf{z}$ is zero-mean Gaussian noise. SBL imposes a parameterized sparsity promoting Gaussian prior on $\mathbf{w}$ [12]

$$p\left(\mathbf{w}; \boldsymbol{\gamma}\right) = \prod_{i=1}^{N} \mathcal{N}\left(w_i; 0, \gamma_i\right). \tag{7}$$

The hyperparameters $\{\gamma_i\}_{i=1}^N$ are estimated using evidence maximization, which can be computed iteratively by the expectation maximization (EM) algorithm. The sparse vector is estimated from the posterior density given the observations and the hyperparameters. SBL has been proven to be more robust compared to the Basis Pursuit [13] and FOCUSS [14] algorithms in terms of (1) guaranteeing the convergence to a sparse solution; and (2) reducing the number of local maxima when compared to other compressed sampling algorithms.

Starting from the system model in (5), we propose two SBL-based algorithms to estimate and mitigate impulsive noise in OFDM systems. Upon convergence, we have an estimate of $\mathbf{\Lambda x}$, as well as a sparse estimate of $\mathbf{e}$. The proposed algorithms, while guaranteeing robustness by the convergence properties of SBL, do not require prior information on the impulsive noise statistics.

V. Estimation Using Null Tones and Pilots

In general, OFDM symbols contain a set of known tones such as null tones used for spectral shaping and reducing inter-carrier interference and pilot tones to aid in channel estimation and synchronization. Let $\mathcal{I}$ denote the set of indices of the known tones in an OFDM symbol $\mathbf{x}$. The cardinality of $\mathcal{I}$ is $|\mathcal{I}| = M/2 < N/2$. Let $(\cdot)_{\mathcal{I}}$ denote the sub-matrix (or sub-vector) corresponding to the rows indexed $\mathcal{I}$. Without loss of generality, we assume that $\mathcal{I} = \{n : \mathbf{x}_n = 0, 0 \leq n \leq N/2 - 1\}$, i.e. the set $\mathcal{I}$ corresponds to the set of null tones' indices. This can be generalized to the case where we have pilot tones and channel information by subtracting out the known parts of $\mathbf{\Lambda x}$ from the received symbol $\mathbf{y}$. Substituting $\mathbf{x}_{\mathcal{I}} = \mathbf{0}$ into (5) and selecting the $\mathcal{I}$-tones from $\mathbf{y}$ we obtain the following $M \times N$ underdetermined linear system

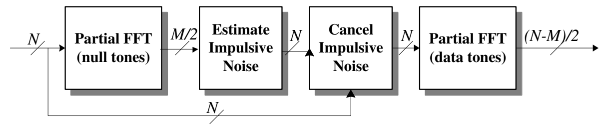

Fig. 3. Proposed non-parametric impulsive noise estimation in OFDM systems, using $M/2$ null tones. N is the FFT size.

$$\mathbf{y}_{\mathcal{I}} = \mathbf{F}_{\mathcal{I}}\mathbf{e} + \mathbf{v}_{\mathcal{I}}, \quad \mathbf{v}_{\mathcal{I}} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}). \tag{8}$$

Imposing the SBL parameterized prior from (7) on $\mathbf{e}$, the pdf of $\mathbf{y}_{\mathcal{I}}$ can be expressed as

$$p(\mathbf{y}_{\mathcal{I}}; \mathbf{\Gamma}, \sigma^2) = (2\pi)^{-\frac{N}{2}} |\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{I}}}|^{-\frac{1}{2}} \exp[-\frac{1}{2}(\mathbf{y}_{\mathcal{I}})^T \mathbf{\Sigma}_{\mathbf{y}_{\mathcal{I}}}^{-1} \mathbf{y}_{\mathcal{I}}], \tag{9}$$

where $\mathbf{\Gamma} \triangleq \text{diag}\{\gamma_0, ..., \gamma_{N-1}\}$ consists of the variance parameters of the Gaussian prior. (9) represents a Gaussian distribution with zero mean and covariance matrix $\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{I}}} \triangleq \sigma^2\mathbf{I} + \mathbf{F}_{\mathcal{I}}\mathbf{\Gamma}\mathbf{F}_{\mathcal{I}}^T$. The hyperparameters $\boldsymbol{\theta} \triangleq (\mathbf{\Gamma}, \sigma^2)$ are estimated by maximizing the likelihood function (9) over $\mathbf{\Gamma}$ and σ^2 independently. Let us denote $\hat{(\cdot)}$ as an estimate of an unknown variable. Given $\hat{\mathbf{\Gamma}}$ and $\hat{\sigma}^2$, an MMSE estimate of $\mathbf{e}$ is the mean of the posterior density $p(\mathbf{e}|\mathbf{y}; \hat{\mathbf{\Gamma}}, \hat{\sigma}^2)$, which is Gaussian $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma}_{\mathbf{e}})$ and

$$\boldsymbol{\mu} = \frac{1}{\hat{\sigma}^2}\mathbf{\Sigma}_{\mathbf{e}}\mathbf{F}_{\mathcal{I}}^T\mathbf{y}_{\mathcal{I}}, \tag{10}$$

$$\mathbf{\Sigma}_{\mathbf{e}} = \hat{\mathbf{\Gamma}} - \hat{\mathbf{\Gamma}}\mathbf{F}_{\mathcal{I}}^T\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{I}}}^{-1}\mathbf{F}_{\mathcal{I}}\hat{\mathbf{\Gamma}}. \tag{11}$$

To solve the evidence maximization problem based on (9), we treat $\mathbf{e}$ as the latent variable and employ the EM algorithm. The $(k+1)$-th iteration is computed from the k-th estimates

$$\begin{aligned} \gamma_i^{(k+1)} &= \underset{\gamma_i \geq 0}{\text{argmax}}\ \mathbb{E}_{\mathbf{e}|\mathbf{y}_{\mathcal{I}};\boldsymbol{\theta}^{(k)}}[\log\ p(\mathbf{y}_{\mathcal{I}}, \mathbf{e}; \boldsymbol{\theta}^{(k)})] \\ &= \mathbb{E}_{\mathbf{e}|\mathbf{y}_{\mathcal{I}};\theta^{(k)}}[\mathbf{e}_i^2] \\ &= (\mathbf{\Sigma}_{\mathbf{e}}{}^{(k)})_{i,i} + (\boldsymbol{\mu}_i^{(k)})^2, \end{aligned} \tag{12}$$

$$\begin{aligned} (\sigma^2)^{(k+1)} &= \frac{1}{M}(||\mathbf{y}_{\mathcal{I}} - \mathbf{F}_{\mathcal{I}}\boldsymbol{\mu}^{(k)}||^2 + \\ &\quad (\sigma^2)^{(k)}\sum_{i=1}^{N}[1 - (\gamma_i^{(k)})^{-1}(\mathbf{\Sigma}_{\mathbf{e}}{}^{(k)})_{i,i}]). \end{aligned} \tag{13}$$

Due to the sparsity promoting prior, upon convergence most components of $\boldsymbol{\mu}$ are driven to zero, rendering a sparse estimate of $\mathbf{e}$. A corresponding receiver design based on (6) is shown in Fig.3. The noise cancellation step is followed by frequency domain equalization and a standard MAP detector.

VI. Joint Estimation Using All Tones

As the number of null tones decreases, performance of the estimator in Section V suffers. On the other hand, increasing the number of null tones reduces the throughput of an OFDM system. As a result, it is desirable to make use of all tones to

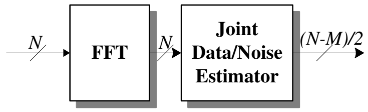

Fig. 4. A non-parametric SBL-based joint data and impulsive noise estimator in OFDM systems, using all tones. N is the FFT size, and M is twice the number of null tones.

Estimator	Operation	Complexity
Using Known Tones	Matrix Multiply	$\mathcal{O}(N^2M)$
	Matrix Inversion	$\mathcal{O}(M^3)$
Using All Tones	Matrix Multiply	$\mathcal{O}(N^3)$
	Matrix Inversion	$\mathcal{O}(N^3)$

TABLE I
COMPLEXITY PER ITERATION OF THE PROPOSED ALGORITHMS. N IS THE FFT SIZE, AND M IS TWICE THE NUMBER OF NULL TONES.

estimate the impulsive noise vector $\mathbf{e}$. However, the data tones are unknown to us. As a result, we seek to jointly estimate the impulsive noise $\mathbf{e}$ (as the input vector) and the OFDM symbol $\mathbf{x}$ (relaxed to be a continuous hyperparameter) using all tones. The modified system level diagram is given in Fig.4.

The derivation of the joint estimator follows very similar steps as in Section V. The main differences are taking into account $\mathbf{x}$ as one of the hyperparameters, and performing evidence maximization over $\boldsymbol{\theta} \triangleq (\mathbf{x}, \boldsymbol{\Gamma}, \sigma^2)$. More specifically, a bias of $\boldsymbol{\Lambda}\mathbf{x}$ is subtracted from $\mathbf{y}$ in all relevant pdf's; and the EM algorithm maximizes the conditional expectation independently over all three hyperparameters. Assuming known channel information $\boldsymbol{\Lambda}$, the $(k+1)$-th iteration of the EM algorithm is summarized in the following.

$$\begin{aligned}
\gamma_i^{(k+1)} &= \underset{\gamma_i \geq 0}{\text{argmax}}\ \mathbb{E}_{\mathbf{e}|\mathbf{y};\boldsymbol{\theta}^{(k)}}[\log\ p(\mathbf{y},\mathbf{e};\boldsymbol{\theta}^{(k)})] \\
&= (\boldsymbol{\Sigma_e}^{(k)})_{i,i} + (\boldsymbol{\mu}_i^{(k)})^2, \qquad (14) \\
(\sigma^2)^{(k+1)} &= \frac{1}{N}(||\mathbf{y} - \boldsymbol{\Lambda}\mathbf{x}^{(k)} - \mathbf{F}\boldsymbol{\mu}^{(k)}||^2 + \\
&\quad (\sigma^2)^{(k)} \sum_{i=1}^{N} [1 - (\gamma_i^{(k)})^{-1}(\boldsymbol{\Sigma_e}^{(k)})_{i,i}]), \qquad (15) \\
\boldsymbol{\Lambda}\mathbf{x}^{(k+1)} &= \mathbf{y} - \mathbf{F}\mathbf{e}^{(k)}, \qquad (16) \\
\boldsymbol{\Sigma_e}^{(k)} &= \boldsymbol{\Gamma}^{(k)} - \boldsymbol{\Gamma}^{(k)}\mathbf{F}^T\Sigma_{\mathbf{y}}^{-1}\mathbf{F}\Gamma^{(k)}, \qquad (17) \\
\mathbf{e}^{(k)} &= \boldsymbol{\mu}^{(k)} = \frac{1}{(\sigma^2)^{(k)}}\boldsymbol{\Sigma_e}^{(k)}\mathbf{F}^T(\mathbf{y} - \boldsymbol{\Lambda}\mathbf{x}^{(k)}) \qquad (18)
\end{aligned}$$

$\mathbf{x}$ is relaxed to be continuous to insure convergence. The estimate of $\mathbf{x}$ is then sliced by a standard MAP detector.

VII. Complexity Analysis

The computational complexity of the two proposed algorithms is dominated by the matrix multiplication and inversion operations in (11) and (17). The complexities per iteration of the two estimators are summarized in Table I. Compared to the estimator using known tones, the joint estimator using all tones increases the complexity from $\mathcal{O}(N^2M)$ per iteration to $\mathcal{O}(N^3)$ per iteration, where the joint estimator trades the complexity for estimation accuracy (since all information available is utilized) and possibly higher throughput (since it does not require the existence of null tones and pilot tones).

VIII. Experimental Results

To quantify the performance of our proposed algorithms, we simulate a baseband OFDM system over a flat channel. We observe that the estimator using null tones is not affected by the channel response. For the estimator using all tones, $\boldsymbol{\Lambda}\mathbf{x}$ can be estimated as a whole from (15), (16) and (18). After that $\mathbf{x}$ can be recovered by a frequency domain equalizer, which requires the channel state information. Therefore the existence of non-flat channel has no effect on the quality of our estimators. In the OFDM simulator, the signal is 4-QAM modulated and mapped to 256 subcarriers, 100 out of which are unused[1]. The null tones are selected to locate at the low frequency side of the spectrum. For comparison, we implement the Compressed Sampling (by Basis Pursuit) and least-squares based algorithm (denoted as CS+LS in the following) proposed in [11]. We use the two MMSE detectors proposed in [3] as references, since both are parametric methods assuming perfect knowledge of the Gaussian mixture impulsive noise model, and one even assumes perfect channel state information (CSI), i.e. the noise variance at each time instance.

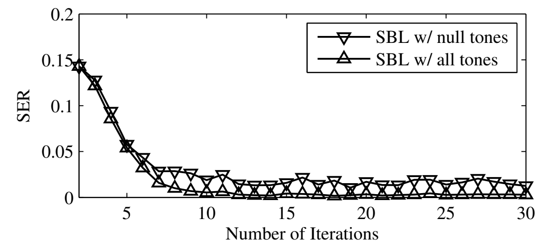

Fig. 6. Symbol error rate trend vs. the number of EM iterations for both proposed algorithms at SNR=0dB. The SER after 30 iterations reaches 1.25×10^{-2} and 2.9×10^{-3} respectively.

We simulate three noise scenarios: Gaussian mixture, Middleton's Class A and Symmetric Alpha-Stable distributions. The noise samples are assumed to be *i.i.d.*. We plot the symbol error rate (SER) versus signal-to-noise ratio (SNR) curves for comparison among various algorithms (Fig. 5). The EM algorithms and the Basis Pursuit algorithm are run for 50 iterations. Fig. 6 shows that both EM algorithms converge after about 15 iterations. Compared to conventional OFDM receivers, for both Gaussian mixture and Middleton's Class A noise models, our estimators achieve around 5dB SNR gain using 100 null tones, and around 10dB gain using all 256 tones. At SNR higher than 5dB (with 100 unused tones) and 0dB (with all tones), our estimators even outperform the reference MMSE detector which assumes known impulsive noise model and unknown CSI. For Gaussian mixture noise, our estimator using all tones gets as close as 4dB to the

[1]In modern wireless communication and PLC standards which employ OFDM, it is commonly suggested that about half of the spectrum be held by null tones. The PRIME PLC standard suggests 158 (out of 256) null tones.

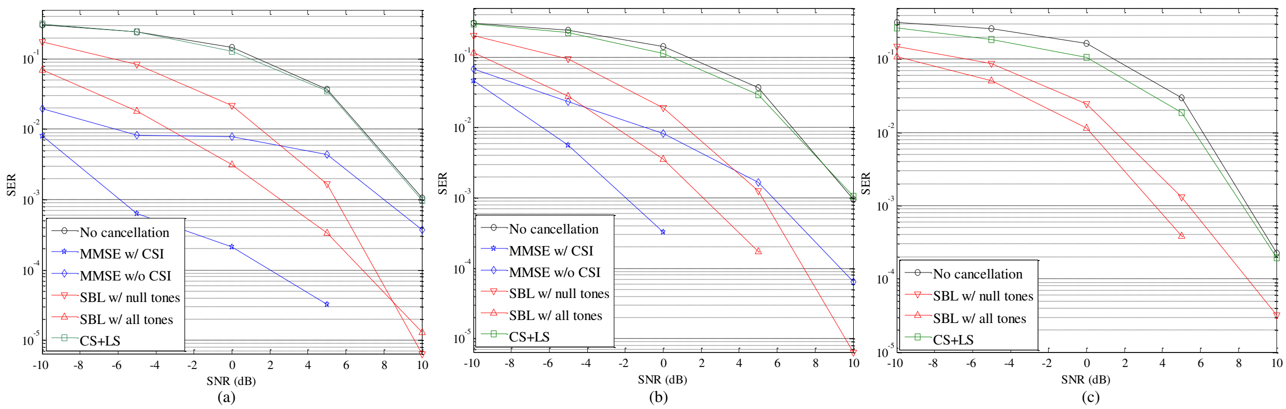

Fig. 5. Performance comparison among impulsive noise mitigation algorithms under three noise scenarios: (a) Gaussian mixture model with $\pi = [0.9, 0.07, 0.03]$ and $\gamma = [1, 100, 1000]$; (b) Middleton's Class A model with $A = 0.1$, $\Gamma = 0.01$ and the first 10 terms are kept for approximation; and (c) Symmetric Alpha-Stable model with $\alpha = 1.2$, $\delta = 0$ and $\gamma = 1$. Our algorithms are non-parametric and applicable to all impulsive noise models.

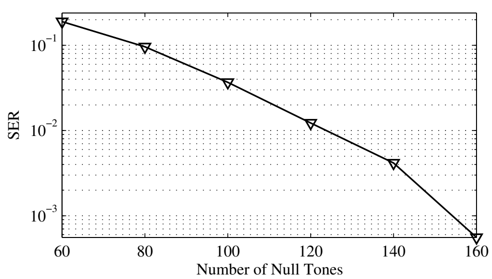

Fig. 7. Symbol error rate trend vs. the number of null tones for the first proposed algorithm at SNR=0dB.

parametric MMSE detector which assumes perfect knowledge of the noise model as well as perfect CSI. Under Symmetric Alpha-Stable modeled noise, the two MMSE estimators cannot be applied since they assume Gaussian mixture noise model. Our estimators achieve 4 to 6dB SNR gain using 100 null tones, and using all tones further improves the gain by 1dB. The CS+LS algorithm works poorly in all these scenarios, since the number of impulses per OFDM symbol is more than the threshold set by the recovery sufficient condition.

In high SNR regimes, the joint estimator's gain from utilizing more information may be overwhelmed by the inaccuracy from relaxing data symbols to be continuous, resulting in worse performance compared to the first proposed estimator (Fig. 5(a)). On the other hand, SER degradation of the first proposed estimator is near-exponential as the number of null tones decreases (Fig. 7), which makes the estimator using all tones more attractive in systems with very few known tones.

IX. Conclusion

This paper proposes two methods for improving communication performance of OFDM systems in the presence of impulsive noise. The methods apply sparse Bayesian learning techniques to estimate the impulsive noise from the received signal by observing information either on the known subcarriers or on all subcarriers. Both methods are non-parametric, i.e. do not require prior knowledge on the statistical noise model or model parameters. We validate the proposed algorithms based on Gaussian mixture, Middleton Class A and Symmetric Alpha-Stable modeled impulsive noise. In future work, we plan to analyze the proposed algorithms in the absence of impulsive noise.

References

[1] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Electromag. Compat.*, vol. 44, no. 1, pp. 249–258, 2002.

[2] M. Nassar, K. Gulati, M. DeYoung, B. Evans, and K. Tinsley, "Mitigating near-field interference in laptop embedded wireless transceivers," *Journal of Signal Proc. Sys.*, pp. 1–12, 2009.

[3] J. Haring, *Error Tolerant Communication over the Compound Channel*. Shaker-Verlag, Aachen, 2002.

[4] Y. Ma, P. So, and E. Gunawan, "Performance analysis of OFDM systems for broadband power line communications under impulsive noise and multipath effects," *IEEE Trans. Power Delivery*, vol. 20, no. 2, pp. 674–682, 2005.

[5] S. Ambike, J. Ilow, and D. Hatzinakos, "Detection for binary transmission in a mixture of Gaussian noise and impulsive noise modeled as an alpha-stable process," *IEEE Signal Proc. Letters*, vol. 1, no. 3, pp. 55–57, 1994.

[6] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: new methods an results for Class a and Class b noise models," *IEEE Trans. on Info. Theory*, vol. 45, no. 4, pp. 1129–1149, 1999.

[7] K. Gulati, B. Evans, J. Andrews, and K. Tinsley, "Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers," *IEEE Trans. on Signal Proc.*, vol. 58, no. 12, pp. 6207–6222, 2010.

[8] K. Gulati, A. Chopra, R. Heath, B. Evans, K. Tinsley, and X. Lin, "MIMO receiver design in the presence of radio frequency interference," in *Proc. IEEE Int. Global Comm. Conf.*, 2008, pp. 1–5.

[9] J. Haring and A. Vinck, "Iterative decoding of codes over complex numbers for impulsive noise channels," *IEEE Trans. Info. Theory*, vol. 49, no. 5, pp. 1251–1260, 2003.

[10] J. Wolf, "Redundancy, the discrete Fourier transform, and impulse noise cancellation," *IEEE Trans. Comm.*, vol. 31, no. 3, pp. 458–461, 2002.

[11] G. Caire, T. Al-Naffouri, and A. Narayanan, "Impulse noise cancellation in OFDM: an application of compressed sensing," in *Proc. IEEE Int. Sym. on Info. Theory*, 2008, pp. 1293–1297.

[12] D. Wipf and B. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Proc.*, vol. 52, no. 8, pp. 2153–2164, 2004.

[13] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.

[14] I. Gorodnitsky and B. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Proc.*, vol. 45, no. 3, pp. 600–616, 2002.