Impulsive Noise Mitigation in Powerline Communications Using Sparse Bayesian Learning

Jing Lin, *Student Member, IEEE,* Marcel Nassar, *Student Member, IEEE,* and Brian L. Evans, *Fellow, IEEE*

***Abstract*—Asynchronous impulsive noise and periodic impulsive noises limit communication performance in OFDM powerline communication systems. Conventional OFDM receivers that assume additive white Gaussian noise experience degradation in communication performance in impulsive noise. Alternate designs assume a statistical noise model and use the model parameters in mitigating impulsive noise. These receivers require training overhead for parameter estimation, and degrade due to model and parameter mismatch. To mitigate asynchronous impulsive noise, we exploit its sparsity in the time domain, and apply sparse Bayesian learning methods to estimate and subtract the noise impulses. We propose three iterative algorithms with different complexity vs. performance trade-offs: (1) we utilize the noise projection onto null and pilot tones; (2) we add the information in the date tones to perform joint noise estimation and symbol detection; (3) we use decision feedback from the decoder to further enhance the accuracy of noise estimation. These algorithms are also embedded in a time-domain block interleaving OFDM system to mitigate periodic impulsive noise. Compared to conventional OFDM receivers, the proposed methods achieve SNR gains of up to 9 dB in coded and 10 dB in uncoded systems in asynchronous impulsive noise, and up to 6 dB in coded systems in periodic impulsive noise.**

***Index Terms*—Asynchronous impulsive noise, periodic impulsive noise, PLC, OFDM, sparse Bayesian learning.**

I. Introduction

A smart grid is an intelligent energy delivery network that overlays communication and computation networks on top of existing power grids for better control and monitoring. Typical examples of smart grid applications include automatic meter reading, grid status control and monitoring, real-time pricing and energy consumption profiling. It is generally agreed that smart grid communications will be supported by a heterogeneous set of communication technologies, ranging from wireless to powerline and to fiber-optic, since no single solution fits all scenarios [2], [3].

Due to the high penetration of power line infrastructures and hence low deployment costs, powerline communications (PLC) plays a prominent role in enabling a variety of smart grid applications. For example, PLC is currently the most adopted (60% market share) communication technology in smart meters [4]. Indoor broadband (BB) PLC provides home area networks that interconnect smart appliances with smart meters for energy consumption profiling and automatic control. On the other hand, outdoor narrowband (NB) PLC is used for "last mile" communications between smart meters and data concentrators, which are deployed by local utilties on medium-voltage (in the US) or low-voltage (in Europe) power lines. We distinguish these two categories of PLC in TABLE I. Both BB and NB PLC use multicarrier communications, such as orthogonal frequency division multiplexing (OFDM), since it offers great advantages in combating frequency-selective channel and spectrally-shaped noise.

Categories	Operating Bands	Data Rates	Standards
Broadband	3–500 kHz	up to 800 kbps	PRIME, G3, IEEE P1901.2, ITU-T G.hnem
Narrowband	1.8–250 MHz	up to 200 Mbps	IEEE P1901, ITU-T G.hn

TABLE I
BROADBAND VS. NARROWBAND PLC.

One of the major challenges for PLC is to overcome additive powerline noise. Such noise is generated by electrical devices connected to the power lines, and by external noise and interference coupled to the power grids via radiation or conduction [3]. Recent field measurements have identified two types of impulsive noise that are dominant in the 1.8–250 MHz band for BB PLC and in the 3–500 kHz band for NB PLC.

- *Asynchronous impulsive noise* is the primary noise component in BB PLC [5], [6]. This type of noise consists of short duration, high power impulses (up to 50 dB above background noise power [5]) with random arrivals. The impulses typically arise from switching transients caused by connection and disconnection of electrical devices. In addition, uncoordinated interference from non-interoperable neighboring PLC modems [3] is shown to be asynchronous impulsive noise in nature [6].
- *Periodic impulsive noise* (also termed "cyclostationary noise") is observed to be dominant in NB PLC [7], [8]. Compared to asynchronous impulsive noise, this type of noise contains longer noise bursts that occur periodically with half the AC cycle. Typical noise bursts cover $10\%-30\%$ of a period, which amounts to $833\ \mu s-2.5\ ms$ in the US. A single noise burst may corrupt multiple consecutive OFDM symbols. For example, the OFDM symbol duration in G3 operating in the CENELEC-A band from 3–95 kHz [9] is $695\ \mu s$, and a noise burst lasting for 30% of a period will contaminate up to 4 consecutive OFDM symbols. During the bursts, the noise power in certain frequency bands can reach 30–50 dB

J. Lin, M. Nassar and B. L. Evans are with the Department of Electrical and Computer Engineering, the University of Texas at Austin, Austin, TX, 78712 USA e-mail: linj@mail.utexas.edu, mnassar@utexas.edu, bevans@ece.utexas.edu.

This work was supported by gift funding and equipment donations from National Instruments, as well as grant funding from the Semiconductor Research Corporation under SRC GRC Task 1836.063 with liaisons Freescale Semiconductor, IBM and Texas Instruments. Part of the work was presented in 2011 IEEE International Global Communications Conference [1].

higher than in the rest of the period [7]. A primary source of periodic impulsive noise is switching mode power supplies (e.g. light dimmers and DC-DC converters) [8].

The presence of impulsive noise may lead to severe deterioration in communication performance at PLC receivers. On one hand, commercial PLC modems feature low power transmission (e.g. less than 18W as specified in [10]), which is likely to be overwhelmed by impulsive noise, not to mention the significant path loss over power lines [3]. On the other hand, statistics of impulsive noise deviates drastically from additive white Gaussian noise (AWGN). Conventional receivers which were optimized for certain performance criteria under the assumption of AWGN may no longer be optimal in non-Gaussian impulsive noise.

Various statistical properties of the impulsive noise can be exploited to improve the reliability and throughput of PLC systems. In particular, assuming a specific statistical noise model, one can design a noise whitening filter [11], minimum mean square error (MMSE) equalizer [12] or decoder [13], [14] to compensate the performance loss due to non-Gaussianity of the noise. Such approaches, however, entail training overhead for model parameter estimation, and may be vulnerable to parameter estimation errors.

In this paper we aim to mitigate asynchronous impulsive noise and periodic impulsive noise, respectively, at OFDM-based PLC receivers. Our work distinguishes from the above approaches in two perspectives: (1) we propose "non-parametric" algorithms that do not make any assumptions on statistical noise models and hence do not require extra training; and (2) our approach estimates and subtracts the impulsive noise from received signal and can be implemented as a denoising block prepended to conventional receivers.

For asynchronous impulsive noise, we develop three denoising algorithms based on sparse Bayesian learning (SBL) techniques [15]. We exploit the sparse structure of the noise in the time domain and estimate it using SBL by observing various subcarriers (a.k.a. tones) of received OFDM symbols. In coded systems, we also show that decision feedback from the convolutional decoder can be used as side information to further improve the denoising performance.

Unlike asynchronous impulsive noise, periodic impulsive noise occurs in bursts that generally span multiple OFDM symbols, and therefore denoising methods directly applied to individual OFDM symbols may not be successful. Instead, we rely on a time-domain block interleaving OFDM transceiver structure (as previously proposed in [16]), where the transmitted and received signals are interleaved and deinterleaved, respectively, in the time domain across multiple OFDM symbols. The deinterleaver effectively scatters the noise bursts into short impulses, which enables us to leverage the SBL-based algorithms for noise estimation.

The rest of the paper is organized as follows. In Section II we briefly describe statistical models for asynchronous and periodic impulsive noise. We review existing receiver algorithms for impulsive noise mitigation in Section III. Having established the system model in Section IV, we propose three non-parametric impulsive noise mitigation algorithms in Section V. Then we perform complexity analysis and present a low-complexity implementation of the first proposed algorithm in Section VI. To demonstrate the performance of our algorithms, simulation results are presented and discussed in Section VII.

II. Statistical Modeling of Impulsive Noise in PLC

In this section, we briefly discuss existing statistical models for asynchronous impulsive noise and periodic impulsive noise in PLC. Although our proposed noise mitigation methods are non-parametric, these statistical models are useful for simulating various impulsive noise environments, under which the robustness of the proposed algorithms can be tested.

A. Asynchronous Impulsive Noise Modeling

Time-domain properties of the asynchronous impulsive noise have been empirically modeled in the literature. Many studies targeted indoor BB PLC, and took noise measurements in higher frequency bands from several hundred kHz to 20 MHz [17], where asynchronous impulsive noise is dominant. To describe the instantaneous amplitude statistics of the noise, various studies empirically fitted the noise data to Nakagami [18], Gaussian mixture and Middleton Class A distributions [19], [20]. By characterizing random emissions of interference events in a PLC network as a temporal Poisson point process, analytical derivation in [6] showed that interference seen at a receiver within a PLC network can be modeled by Gaussian mixture and Middleton Class A distributions. The three network scenarios and corresponding noise models in [6] are given in Fig. 1 and TABLE II.

For convenience of the discussion in later sections, we briefly describe the Gaussian mixture and the Middleton Class A models as follows.

1) Gaussian Mixture Model: 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.

2) Middleton Class A Model: The Middleton Class A model [21] is parameterized by an overlapping factor A and background-to-impulsive noise power ratio Ω ($A \in [10^{-2}, 1]$ and $\Omega \in [10^{-6}, 1]$ in general [22]). It can be considered as a special case of the Gaussian mixture distribution, with $\pi_k = e^{-A}\frac{A^k}{k!}$ and $\gamma_k = \frac{k/A+\Omega}{1+\Omega}$ as $K \to \infty$. In practice, only the first few significant terms are retained.

Time-domain noise traces simulated from Gaussian mixture and Middleton Class A models are depicted in Fig. 2 and 3, respectively.

B. Periodic Impulsive Noise Modeling

Studies in statistical modeling of periodic impulsive noise primarily targeted outdoor NB PLC in the 3–500 kHz band

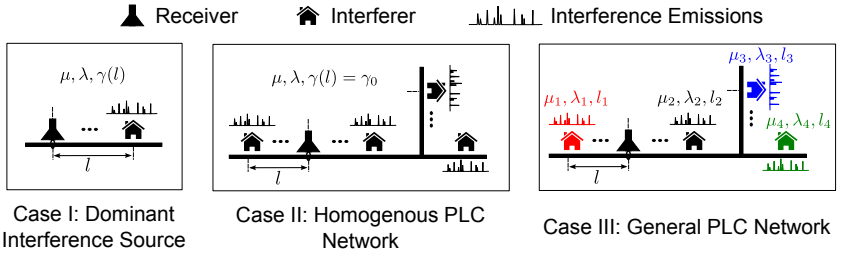

Fig. 1. Interference scenarios in PLC networks. Each interferer emits a random sequence of emissions onto the powe line, which add up at the receiver. An interferer is described statistically by a mean number of emission events μ, mean duration between emission events λ, and the pathloss to the receiver γ.

Scenario	Example Network	Statistical Model
Dominant Interferer	Rural Area Industrial Area	Middleton Class A: $A = \lambda\mu$, $\Omega = A\gamma E\left[h^2 B^2\right]/2$
Homogeneous PLC Network	Urban Area Residential Buildings	Middleton Class A: $A = M\lambda\mu$, $\Omega = A\gamma E\left[h^2 B^2\right]/2M$
General PLC Network	Dense Urban Area Commercial	Gaussian Mixture: π_k and γ_k given in [6]

TABLE II
STATISTICAL-PHYSICAL MODELS OF INTERFERENCE IN PLC NETWORKS CATEGORIZED BY NETWORK TYPES. PARAMETERS ARE GIVEN IN FIG. 1 AND M IS THE NUMBER OF INTERFERERS.

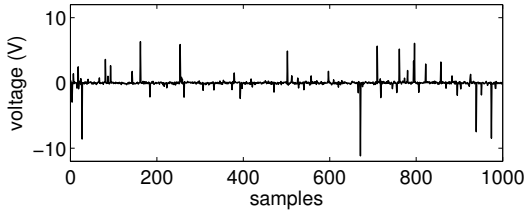

Fig. 2. Asynchronous impulsive noise simulated from a Gaussian mixture distribution with $\boldsymbol{\pi} = [0.9, 0.07, 0.03]$ and $\boldsymbol{\gamma} = [1, 100, 1000]$.

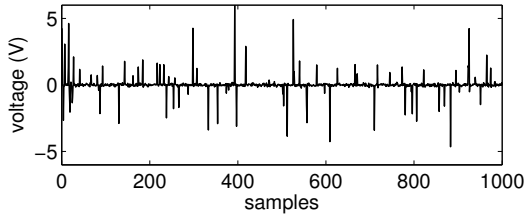

Fig. 3. Asynchronous impulsive noise simulated from a Middleton Class A distribution with $A = 0.1$, $\Omega = 0.01$, and the pdf truncated to the first 10 mixture components.

[7], [23]. In [23], the noise was expressed as a cyclostationary Gaussian process whose instantaneous variance is a periodic function of time. A linear time invariant (LTI) filter was used for shaping the noise spectrum. A more general linear periodically time varying (LPTV) system model was proposed in [7] and was accepted into the IEEE P1901.2 NB-PLC standard. The model was established on the approximation that each AC cycle can be partitioned into a number of intervals, within each the noise is a stationary Gaussian process characterized by a particular power spectral density. The periodic impulsive noise can therefore be generated by passing an AWGN input through a set of LTI filters and switching the output periodically among them. More specifically, suppose that an AC cycle is paritioned into M intervals $\{\mathcal{R}_i\}_{i=1}^{M}$, the noise samples n_k

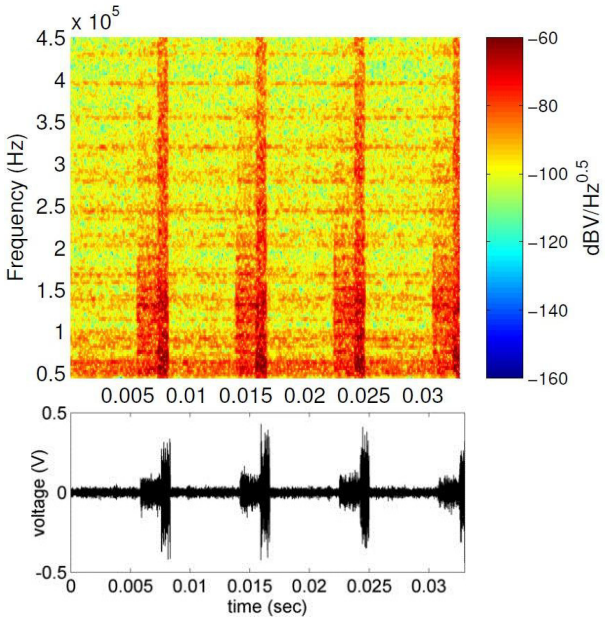

Fig. 4. A time-domain trace and spectrogram of the periodic impulsive noise synthesized from a linear periodically time varying system model.

can be expressed as

$$n_k = \sum_{i=1}^{M} \mathbf{1}_{k\in\mathcal{R}_i} \sum_{\tau} h_\tau^{(i)} v_{k-\tau}, \;\; v_k \sim \mathcal{N}(0,1). \qquad (2)$$

where $\mathbf{1}_{\mathcal{A}}$ is the indicator function, and $h_\tau^{(i)}$ denotes the impulse response of the LTI filter that the system switches to during the interval $\mathcal{R}_i$.

A time-domain trace and spectrogram of the periodic impulsive noise synthesized from an LPTV system model is shown in Fig. 4. One period of the noise is divided into three intervals, each assuming a different spectral shape. The spectral shapes are fitted to noise measurements collected at an outdoor low-

voltage site as shown in [7].

III. Prior Work

In this section, we briefly review prior work on receiver methods and transmission schemes to mitigate the effect of asynchronous impulsive noise and periodic impulsive noise on communication performance.

A. Asynchronous Impulsive Noise Mitigation

Asynchronous impulsive noise arises not only in BB PLC but also in wireless networks such as *ad hoc* and cellular networks [24]. Earlier approaches in mitigating asynchronous impulsive noise involve parametric methods, which assume a particular statistical noise model and typically estimate the parameters of the statistical model during a training stage. Examples of such algorithms include pre-filtering techniques [25], [26], nulling and clipping methods [27], MMSE symbol-by-symbol detectors [13], and iterative decoders [14], [28]. The advantage of parametric methods is that they lead to performance gains by exploiting information of the noise model and its parameters. However, they require extra training overhead and can suffer from performance degradation when the noise model or parameters mismatch the possibly time varying noise statistics.

Recently, there has been growing interest in developing non-parametric denoising methods that exploit the sparse structure of the asynchronous impulsive noise in the time domain. In particular, [29] applied the compressed sensing (CS) techniques to estimate the impulsive noise from the null tones (i.e., tones that do not carry data or pilots) of the received signal. The algorithm was subject to a sufficient recovery condition stating that the number of impulses within an OFDM symbol does not exceed a threshold that is uniquely determined by the discrete Fourier transform (DFT) size and the number of null tones. However, for common OFDM system settings in PLC, the threshold turns out to be too restrictive for many impulsive noise environments where an OFDM symbol is corrupted by multiple impulses. This CS-based approach was extended in [30] to a bursty impulsive noise detector that exploits the block-sparsity of the noise. The performance of the algorithm, however, is affected by parameters that should be ideally adapted to the number of noise bursts within an OFDM symbol, and the background noise level.

Our work seeks to develop non-parametric mitigation algorithms that are applicable to all asynchronous impulsive noise scenarios. Towards this end, we extend the CS based algorithm in [29] to a sparse Bayesian learning (SBL) approach [15] for improved performance and robustness.

B. Periodic Impulsive Noise Mitigation

In general, parameter estimation in periodic impulsive noise is even more difficult than that in asynchronous impulsive noise. This is because of the significant increase in the number of parameters, and hence the degrees of freedom, in order to capture the non-negligible time-domain correlation in periodic impulsive noise. Accurate estimation of these parameters generally requires a large amount of data, i.e., over multiple cycles, which entails not only significant training overhead, but also a large memory typically not present in current PLC modems. Furthermore, the increased degrees of freedom makes the estimation more vulnerable to outliers.

Despite of the difficulty in parameter estimation, parametric methods for cyclostationary noise mitigation, assuming perfect knowledge of the second-order statistics, have been explored in the literature. In [31] and [32], it was observed that the cyclic spectrum, i.e., the Fourier transform of the autocorrelation function, of a second-order cyclostationary process contains harmonic peaks, and therefore can be used for the detection and extraction of such process. In [12] and [33], a linear MMSE frequency domain equalizer for single-carrier OFDM systems was derived based on the second-order noise statistics.

Exploiting the strong correlation between time-domain noise samples, adaptive filtering algorithms were proposed to predict [34], [35] or whiten [11] the periodic impulsive noise at the NB PLC receivers. In particular, in [11], the noise was fitted to a periodically switching autoregressive (AR) process by nonparametric Bayesian learning. Based on the estimated AR model, a periodically switching moving average filter was adopted at the receiver for noise whitening. A common drawback of these filter-based methods [11], [34], [35] is the vulnerability to outliers, e.g. asynchronous impulsive noise simultaneously present in the higher frequency bands of NB PLC. The improvement in robustness against such outliers generally requires longer training sequences.

At the transmitter side, coding and interleaving schemes that are resilient to bursty impulsive noise have been investigated [16], [36], [37]. Since periodic impulsive noise occurs in bursts that typically span more than one OFDM symbols, joint processing across a large number of OFDM symbols can be beneficial. Such joint processing includes forward error correction (FEC) coding at the application layer [36] and time-domain interleaving [16], [37]. In particular, a time-domain block interleaving OFDM (TDI-OFDM) transceiver structure was proposed in [16] to cope with bursty impulsive noise. Unlike conventional frequency-domain interleaving OFDM (FDI-OFDM) systems where the interleaver is placed before the inverse discrete Fourier transform (IDFT) at the transmitter, the TDI-OFDM scheme interleaves and deinterleaves the signal in the time domain, i.e., post-IDFT at the transmitter and pre-DFT at the receiver. The purpose is to spread the samples that are corrupted by impulsive noise, and thus average the impact on bit error rates (BER) over a large number of OFDM symbols. It was shown in [16] that TDI-OFDM has superior BER improvement over FDI-OFDM at higher SNRs (e.g. above 10 dB or 20 dB, depending on noise scenarios). However, typical SNR values in NB PLC systems range from -5 dB to 10 dB [9], in which TDI-OFDM generally has diminishing gains over FDI-OFDM and even performs worse towards lower SNRs.

Although the TDI-OFDM scheme itself does not provide much benefit in NB PLC, we will show that by embedding our proposed SBL-based denoising algorithms into the TDI-OFDM framework, significant BER improvement over conventional FDI-OFDM systems can be achieved even at low

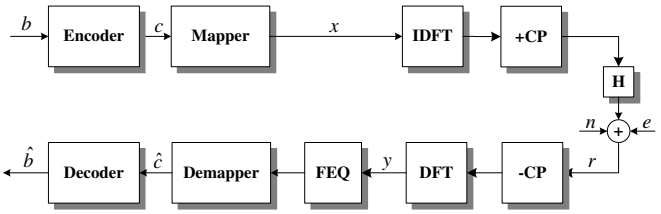

Fig. 5. A conventional baseband coded OFDM system.

SNR regimes. The idea is to exploit the sparse structure of the noise after the time-domain deinterleaver, and leverage the SBL-based denoising algorithms we have developed for asynchronous impulsive noise mitigation.

IV. System Model

For asynchronous impulsive noise mitigation in BB PLC, we consider a conventional coded OFDM system whose complex baseband equivalent representation is shown in Fig. 5. At the transmitter, a binary data packet $\mathbf{b}$ is encoded into a codeword $\mathbf{c}$. The codeword is then mapped to OFDM symbols, each with M non-data tones and $N-M$ data tones. The non-data tones are either null tones for spectral shaping and inter-carrier interference reduction, or pilots for channel estimation and synchronization. An OFDM symbol, denoted by $\mathbf{x}$, is converted to the time domain by IDFT. A cyclic prefix (CP), assumed to be longer than the channel delay spread, is inserted to the beginning of each OFDM symbol to prevent inter-symbol interference (ISI).

At the receiver, we remove the CP from the received OFDM symbols, resulting in

$$\mathbf{r} = \mathbf{H}\mathbf{F}^*\mathbf{x} + \mathbf{e} + \mathbf{n}, \tag{3}$$

where $\mathbf{F}$ is the N-point DFT matrix, $\mathbf{H} \in \mathbb{C}^{N\times N}$ is the convolutional matrix of the channel and is circulant due to the cyclic prefix insertion, and $\mathbf{e}, \mathbf{n} \in \mathbb{C}^N$ represent impulsive noise and AWGN, respectively. The OFDM demodulator takes the DFT of $\mathbf{r}$, leading to

$$\begin{aligned} \mathbf{y} &= \mathbf{F}\mathbf{H}\mathbf{F}^*\mathbf{x} + \mathbf{F}\mathbf{e} + \mathbf{F}\mathbf{n} \\ &= \mathbf{\Lambda}\mathbf{x} + \mathbf{F}\mathbf{e} + \mathbf{g}. \end{aligned} \tag{4}$$

Here $\mathbf{\Lambda} \triangleq \mathbf{F}\mathbf{H}\mathbf{F}^*$ is a diagonal matrix, with $\{H_i\}_{i=1}^N$ (the N-point DFT coefficients of the channel impulse response) on its diagonal, and $\mathbf{g} \triangleq \mathbf{F}\mathbf{n}$ is the DFT of $\mathbf{n}$ and is also AWGN since $\mathbf{F}$ is unitary.

For periodic impulsive noise mitigation in NB PLC, we consider a TDI-OFDM system [16] as shown in Fig. 6. At the transmitter, multiple OFDM symbols are interleaved using a sample-level block interleaver after the IDFT and before CP insertion. Inserting the CP after the interleaver maintains the cyclic structure within each transmitted OFDM symbol, and hence the received signal after CP removal is the circular convolution of the transmitted signal with the multi-path channel. Similarly to conventional OFDM systems, such signal can be equalized by one-tap frequency-domain channel equalizer (FEQ). The equalized signal is then deinterleaved before converted to the frequency domain by DFT. Assuming perfect channel estimation, the demodulated OFDM signal $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \mathbf{x} + \mathbf{F}\mathbf{e}_\pi + \mathbf{F}\mathbf{n}_\pi = \mathbf{x} + \mathbf{F}\mathbf{e}_\pi + \mathbf{g}_\pi. \tag{5}$$

Here $\mathbf{e}_\pi$ and $\mathbf{n}_\pi$ denote the time-domain impulsive noise and additive Gaussian noise after deinterleaving, and $\mathbf{g}_\pi \triangleq \mathbf{F}\mathbf{n}_\pi$. Note that although the AWGN $\mathbf{n}$ is spectrally shaped by the FEQ, it becomes less correlated in the time domain after the block deinterleaver and hence $\mathbf{n}_\pi$, as well as $\mathbf{g}_\pi$, can be well approximated by AWGN.

Given a particular noise scenario, the size of the interleaver is an important design factor that determines the sparseness of $\mathbf{e}_\pi$ and therefore the performance of the SBL-based algorithms. The key is to maintain the sparseness of $\mathbf{e}_\pi$ (i.e., number of non-zero elements) below a certain level that allows accurate estimation by the SBL techniques. In NB PLC systems, the interleaving can be done over an entire packet, which contains up to 56 QPSK modulated OFDM symbols according to the G3 standard in the CENELEC-A band. This gives a maximum interleaver size of $38.92\ ms$, spanning about 2.3 AC cycles in the US, or equivalently 4.6 noise periods. The maximum packet duration will be doubled in BPSK modulation and even larger when repetition code is used. As such, we claim that the assumption of having a large interleaver with the size approximately equal to integer multiples of the noise period is realistic in NB PLC systems. Such interleavers will result in $\mathbf{e}_\pi$ with sparseness typically ranging from 10% to 30%, which can be accurately recovered by the SBL-based algorithms, as will be demonstrated by the simulation results.

With the assumption that the interleaver size is appropriately selected to render $\mathbf{e}_\pi$ a sparse vector, (5) can be considered as a special case of (4) with $\mathbf{\Lambda} = \mathbf{I}$. Therefore, the following algorithms derived from (4) can also be directly applied to (5).

Let $\mathcal{I}$ denote the index set of the null and pilot tones, where $|\mathcal{I}| = M < N$. Also, let $(\cdot)_\mathcal{I}$ denote the sub-matrix (or sub-vector) corresponding to the rows (or elements) indexed by the set $\mathcal{I}$. Assuming perfect channel estimation, i.e. complete knowledge of Λ, the impulsive noise can be observed from the null and pilot tones of the received OFDM symbol, since

$$\begin{aligned} \mathbf{z} &\triangleq \mathbf{y}_\mathcal{I} - (\mathbf{\Lambda}\mathbf{x})_\mathcal{I} \\ &= \mathbf{F}_\mathcal{I}\mathbf{e} + \mathbf{g}_\mathcal{I}, \\ \mathbf{g}_\mathcal{I} &\sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_M). \end{aligned} \tag{6}$$

The recovery of the length-N vector $\mathbf{e}$ from the noisy underdetermined $M \times N$ linear system is generally an ill-conditioned problem. However, exploiting the sparse nature of $\mathbf{e}$ (since impulsive noise has very few non-zero samples in the time domain), we can possibly get an accurate estimate of $\mathbf{e}$ by applying various compressed sensing techniques.

We would like to use the estimated impulsive noise to improve the detection of $\mathbf{x}$. More specifically, the impulsive noise estimate $\hat{\mathbf{e}}$ can be subtracted from the received symbol on the data tones to form a new decision metric

$$\begin{aligned} \hat{\mathbf{y}}_{\overline{\mathcal{I}}} &= \mathbf{y}_{\overline{\mathcal{I}}} - \mathbf{F}_{\overline{\mathcal{I}}}\hat{\mathbf{e}} \\ &= (\mathbf{\Lambda}\mathbf{x})_{\overline{\mathcal{I}}} + \mathbf{g}_{\overline{\mathcal{I}}} + \mathbf{F}_{\overline{\mathcal{I}}}(\mathbf{e} - \hat{\mathbf{e}}). \end{aligned} \tag{7}$$

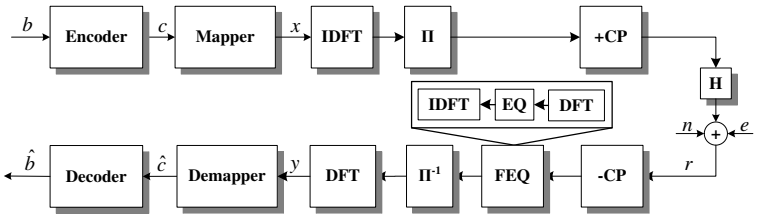

Fig. 6. A time-domain interleaved OFDM system. Π denotes the sample-level interleaver, and Π^{-1} the corresponding deinterleaver.

where $\overline{(\cdot)}$ indicates set complement and thus $\overline{\mathcal{I}}$ indicates the set of data tone indices. Assuming that $\hat{\mathbf{e}} \approx \mathbf{e}$, the receiver can then proceed as if only Gaussian noise were present and apply the conventional detection and decoding algorithms.

V. Non-parametric Impulsive Noise Estimation

The estimation of impulsive noise converts to solving an underdetermined linear regression problem in (6) under sparsity constraints. Among various compressed sensing algorithms, sparse Bayesian learning (SBL) has become increasingly attractive due to its improved robustness over deterministic approaches such as Basis Pursuit [38]. Furthermore, the Bayesian framework makes it convenient to fuse information not only in the null tones but also in the data tones of the received signal to enhance the accuracy of the noise estimation. Therefore we apply the SBL techniques and propose three non-parametric algorithms for impulsive noise estimation, with different complexity vs. performance trade-offs. Before introducing the algorithms, we briefly describe the SBL framework.

A. Sparse Bayesian Learning

SBL was first proposed by Tipping [39], and was introduced to sparse signal recovery by Wipf and Rao in [15]. Generally, SBL is a Bayesian learning approach for solving the linear regression problem

$$\mathbf{t} = \mathbf{\Phi}\mathbf{w} + \mathbf{v}, \quad \mathbf{v} \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_M), \tag{8}$$

where $\mathbf{t} \in \mathcal{C}^M$ is an observation vector, $\mathbf{\Phi} = \begin{bmatrix}\Phi_1 & \cdots & \Phi_N\end{bmatrix} \in \mathcal{C}^{M\times N}$ is an overcomplete basis (i.e. $M < N$), and $\mathbf{w} \in \mathcal{C}^N$ is a sparse weight vector to be estimated.

SBL imposes a parameterized Gaussian prior on $\mathbf{w}$

$$p(\mathbf{w};\mathbf{\Gamma}) = \mathcal{CN}(\mathbf{w};\mathbf{0},\mathbf{\Gamma}), \tag{9}$$

where $\mathbf{\Gamma} \triangleq \text{diag}\{\boldsymbol{\gamma}\}$, and $\boldsymbol{\gamma} \in \mathcal{R}^N$ whose i-th component γ_i is the variance of w_i. Given the prior, the likelihood of the observation can be expressed as

$$p(\mathbf{t};\mathbf{\Gamma},\sigma^2) = \mathcal{CN}(\mathbf{t};\mathbf{0},\mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^* + \sigma^2\mathbf{I}_M). \tag{10}$$

A maximum likelihood (ML) estimator solves the hyperparameters $\boldsymbol{\gamma}$ and σ^2 that maximize (10). The ML estimation is computed iteratively using expectation maximization (EM).

Given the observations and the estimated hyperparameters, the posterior density of $\mathbf{e}$ is also a Gaussian distribution

$$\begin{aligned} p(\mathbf{w}|\mathbf{t};\mathbf{\Gamma},\sigma^2) &= \mathcal{CN}(\mathbf{w};\boldsymbol{\mu}_{\boldsymbol{w}},\mathbf{\Sigma}_{\boldsymbol{w}}), \\ \boldsymbol{\mu}_{\boldsymbol{w}} &= \sigma^{-2}\mathbf{\Sigma}_{\boldsymbol{w}}\mathbf{\Phi}^*\mathbf{t}, \\ \mathbf{\Sigma}_{\boldsymbol{w}} &= (\sigma^{-2}\mathbf{\Phi}^*\mathbf{\Phi} + \mathbf{\Gamma}^{-1})^{-1}. \end{aligned} \tag{11}$$

The maximum *a posteriori* (MAP) estimate of $\mathbf{w}$ is the posterior mean $\boldsymbol{\mu}_{\boldsymbol{w}}$.

Due to the sparsity promoting property of the prior, upon convergence most components of $\boldsymbol{\gamma}$ and hence $\boldsymbol{\mu}_{\boldsymbol{w}}$ are driven to zero, rendering a sparse estimate of $\mathbf{w}$. It has been shown in [15] that SBL has improved robustness compared to other compressed sensing algorithms such as Basis Pursuit [38] and FOCUSS [40], since the global optimum is always the sparsest solution, all local optimal solutions are sparse, and the number of local optima is the smallest.

B. Estimation Using Null and Pilot Tones

The SBL technique can be directly applied to the impulsive noise estimation using null and pilot tones, since substituting $\mathbf{t} = \mathbf{z}$, $\mathbf{\Phi} = \mathbf{F}_{\mathcal{I}}$, $\mathbf{w} = \mathbf{e}$, and $\mathbf{v} = \mathbf{g}_{\mathcal{I}}$ into (8) gives exactly (6).

To obtain the ML estimates of the hyperparameters $\boldsymbol{\gamma}$ and σ^2, we treat $\mathbf{e}$ as the latent variable and apply the EM algorithm. The update of the hyperparameters in the k-th iteration is as follows, where we define $\boldsymbol{\theta} \triangleq (\boldsymbol{\gamma}, \sigma^2)$ for conciseness.

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

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

$$\boldsymbol{\mu}_{\boldsymbol{e}}{}^{(k)} = (\sigma^{-2})^{(k)}\mathbf{\Sigma}_{\boldsymbol{e}}{}^{(k)}\mathbf{F}_{\mathcal{I}}^*\mathbf{z}, \tag{14}$$

$$\mathbf{\Sigma}_{\boldsymbol{e}}{}^{(k+1)} = [(\sigma^{-2})^{(k)}\mathbf{F}_{\mathcal{I}}^*\mathbf{F}_{\mathcal{I}} + (\mathbf{\Gamma}^{(k)})^{-1}]^{-1}. \tag{15}$$

Upon termination of the EM algorithm, we obtain the MAP estimate of the time-domain impulsive noise $\hat{\mathbf{e}} = \boldsymbol{\mu}_{\boldsymbol{e}}$. We then transform $\hat{\mathbf{e}}$ to the frequency domain and subtract it from the received signal in the data tones according to (7).

C. Estimation Using All Tones

As will be demonstrated in the simulation results, performance of the estimator using null and pilot tones is improved as the number of these non-data tones increases. However having fewer data tones means reduced throughput. When the number of non-data tones is limited, it is desirable to exploit information available in all tones to estimate the impulsive noise. To do this, we define $\mathbf{u} \triangleq \mathbf{\Lambda}\mathbf{x} + \mathbf{g}$, and rewrite (4) as

$$\begin{aligned}\left[\begin{array}{c}\mathbf{z}\\ \mathbf{y}_{\overline{\mathcal{I}}}\end{array}\right] &= \mathbf{F}\mathbf{e}+\left[\begin{array}{c}\mathbf{u}_{\mathcal{I}}\\ \mathbf{u}_{\overline{\mathcal{I}}}\end{array}\right],\\ \mathbf{u}_{\mathcal{I}} &\sim \mathcal{CN}(\mathbf{0},\sigma^2\mathbf{I}_M),\\ \mathbf{u}_{\overline{\mathcal{I}}} &\sim \mathcal{CN}((\mathbf{\Lambda x})_{\overline{\mathcal{I}}},\sigma^2\mathbf{I}_{N-M}).\end{aligned} \tag{16}$$

Imposing the parameterized Gaussian prior on $\mathbf{e}$, i.e. $p(\mathbf{e};\mathbf{\Gamma})=\mathcal{CN}(\mathbf{0},\mathbf{\Gamma})$, the likelihood of $\mathbf{z}$ remains the same as (10), while the likelihood of $\mathbf{y}_{\overline{\mathcal{I}}}$, the received signal on the data tones, is

$$\begin{aligned}p\big(\mathbf{y}_{\overline{\mathcal{I}}};(\mathbf{\Lambda x})_{\overline{\mathcal{I}}},\mathbf{\Gamma},\sigma^2\big) &= \mathcal{CN}\big(\mathbf{y}_{\overline{\mathcal{I}}};(\mathbf{\Lambda x})_{\overline{\mathcal{I}}},\mathbf{\Sigma}_{\mathbf{y}_{\overline{\mathcal{I}}}}\big),\\ \mathbf{\Sigma}_{\mathbf{y}_{\overline{\mathcal{I}}}} &= \boldsymbol{F}_{\overline{\mathcal{I}}}\mathbf{\Gamma}\boldsymbol{F}^*_{\overline{\mathcal{I}}}+\sigma^2\mathbf{I}_{N-M},\end{aligned} \tag{17}$$

with the unknown transmitted signal $\mathbf{x}_{\overline{\mathcal{I}}}$ as a third hyperparameter in addition to $\mathbf{\Gamma}$ and σ^2. Although $\mathbf{x}_{\overline{\mathcal{I}}}$ consists of constellation points, which are discrete, we temporarily relax it to be continuous when treated as a hyperparameter in the EM algorithm.

The iterative updates in the EM algorithm now involve all three hyperparameters. Since each hyperparameter is updated while keeping the others fixed, the update equations for $\boldsymbol{\gamma}$ and σ^2 are in the same forms as (12) and (13). We treat $(\mathbf{\Lambda x})_{\overline{\mathcal{I}}}$ in a whole as a hyperparameter and update it as

$$\begin{aligned}(\mathbf{\Lambda x})_{\overline{\mathcal{I}}}^{(k+1)} &= \underset{(\mathbf{\Lambda x})_{\overline{\mathcal{I}}}}{\operatorname{argmax}}\ \mathbb{E}_{\mathbf{e}|\mathbf{y};\boldsymbol{\theta}^{(k)}}[\log\ p(\mathbf{y},\mathbf{e};\boldsymbol{\theta}^{(k)})]\\ &= \underset{(\mathbf{\Lambda x})_{\overline{\mathcal{I}}}}{\operatorname{argmin}}\ |\mathbf{y}_{\overline{\mathcal{I}}}-(\mathbf{\Lambda x})_{\overline{\mathcal{I}}}-\mathbf{F}_{\overline{\mathcal{I}}}{\boldsymbol{\mu}_{\boldsymbol{e}}}^{(k)}|^2\\ &= \mathbf{y}_{\overline{\mathcal{I}}}-\mathbf{F}_{\overline{\mathcal{I}}}{\boldsymbol{\mu}_{\boldsymbol{e}}}^{(k)}.\end{aligned} \tag{18}$$

The entire EM algorithm is summarized as follows.

$$\begin{aligned}\gamma_i^{(k+1)} &= \underset{\gamma_i\geq 0}{\operatorname{argmax}}\ \mathbb{E}_{\mathbf{e}|\mathbf{y};\boldsymbol{\theta}^{(k)}}[\log\ p(\mathbf{y},\mathbf{e};\boldsymbol{\theta}^{(k)})]\\ &= \mathbf{\Sigma}_{\mathbf{e},ii}^{(k)}+(\boldsymbol{\mu}_{\boldsymbol{e},i}^{(k)})^2,\end{aligned} \tag{19}$$

$$\begin{aligned}(\sigma^2)^{(k+1)} &= \frac{1}{N}(||\mathbf{y}-\mathbf{\Lambda x}^{(k)}-\mathbf{F}\boldsymbol{\mu}^{(k)}||^2+\\ &\quad (\sigma^2)^{(k)}\sum_{i=1}^{N}[1-(\gamma_i^{(k)})^{-1}\mathbf{\Sigma}_{\mathbf{e},ii}^{(k)}]),\end{aligned} \tag{20}$$

$$(\mathbf{\Lambda x})_{\overline{\mathcal{I}}}^{(k+1)} = \mathbf{y}_{\overline{\mathcal{I}}}-\mathbf{F}_{\overline{\mathcal{I}}}\mathbf{e}^{(k)}, \tag{21}$$

$$\mathbf{\Sigma_e}^{(k)} = \mathbf{\Gamma}^{(k)}-\mathbf{\Gamma}^{(k)}\mathbf{F}^*\Sigma_{\mathbf{y}}^{-1}\mathbf{F}\Gamma^{(k)}, \tag{22}$$

$$\hat{\mathbf{e}}^{(k)} = \boldsymbol{\mu_e}^{(k)} = \frac{1}{(\sigma^2)^{(k)}}\mathbf{\Sigma_e}^{(k)}\mathbf{F}^*(\mathbf{y}-\mathbf{\Lambda x}^{(k)}) \tag{23}$$

D. Decision Feedback Estimation

The two estimators described above impose a parameterized Gaussian prior on the time-domain impulsive noise, i.e. $\mathbf{e}\sim\mathcal{CN}(0,\mathbf{\Gamma})$. Prior information on $\mathbf{\Gamma}$, or equivalently on the precision matrix $\mathcal{T}\triangleq\mathbf{\Gamma}^{-1}$, can be introduced by the conjugate prior distribution on these hyperparameters. Let $\boldsymbol{\tau}\triangleq[\tau_1,\cdots,\tau_N]^T$ denote the diagonal of $\mathcal{T}$. The conjugate prior on $\boldsymbol{\tau}$ is a Gamma distribution

$$P(\boldsymbol{\tau};\mathbf{a},\mathbf{b})=\prod_{i=1}^{N}\mathrm{Ga}(\tau_i;a_i,b_i). \tag{24}$$

where $\mathrm{Ga}(\cdot;a,b)$ is the Gamma distribution with parameters a and b. When $a_i=0,b_i=0,\forall i$, (24) reduces to a uniform distribution, which is an non-informative prior that is implicitly imposed in the previously described SBL framework. Non-zero values of a_i and b_i contain prior information that is integrated into the likelihood function in (10), resulting in

$$\begin{aligned}p(\mathbf{z};\mathcal{T},\sigma^2,\mathbf{a},\mathbf{b}) &= \mathcal{CN}(\mathbf{z};\mathbf{0},\mathbf{F}_{\mathcal{I}}\mathcal{T}^{-1}\mathbf{F}^*_{\mathcal{I}}+\sigma^2\mathbf{I}_M)\times\\ &\quad \mathrm{Ga}(\mathbf{P};\mathbf{a},\mathbf{b}).\end{aligned} \tag{25}$$

Maximizing (25) over $\boldsymbol{\tau}$, we obtain the maximum likelihood (ML) estimate of each element τ_i as

$$\tau_i=\gamma_i^{-1}=\frac{1+2a_i}{\boldsymbol{\mu}_{\boldsymbol{e},i}^2+\mathbf{\Sigma}_{\boldsymbol{e},ii}+2b_i}. \tag{26}$$

Comparing (26) to (12), we can see the prior information contained in a_i and b_i does affect the ML estimates of $\boldsymbol{\gamma}$. Since (26) is the conjugate prior on $\boldsymbol{\tau}$, the posterior probability of $\boldsymbol{\tau}$ given $\mathbf{e},\mathbf{a}$ and $\mathbf{b}$ is also Gamma distributed, i.e.

$$P(\boldsymbol{\tau}|\mathbf{e};\mathbf{a},\mathbf{b})=\prod_{i=1}^{N}\mathrm{Ga}(\tau_i;\tilde{a}_i,\tilde{b}_i) \tag{27}$$

with the updated parameters

$$\begin{aligned}\tilde{a}_i &= a_i+\frac{1}{2},\\ \tilde{b}_i &= b_i+\frac{|e_i|^2}{2}.\end{aligned} \tag{28}$$

Suppose that in addition to the MAP estimate $\hat{\mathbf{e}}$ given by the estimator using non-data tones, a second estimate of $\mathbf{e}$, denoted by $\hat{\mathbf{e}}'$, is available based on certain side information. The side information contained in $\hat{\mathbf{e}}'$ can be fused into $\hat{\mathbf{e}}$ via the posterior distribution of $\boldsymbol{\tau}$ given $\hat{\mathbf{e}}'$. More specifically, given $\hat{\mathbf{e}}'$, we update $\mathbf{a}$ and $\mathbf{b}$ according to (28), and then solve the ML estimate of $\boldsymbol{\tau}$ (26) with the updated $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$.

In coded OFDM systems, the redundancy in the coded data tones can be exploited as the side information to provide a second estimate of $\hat{\mathbf{e}}'$. More specifically, the decoder takes the OFDM symbols after impulsive noise mitigation as the input, and produces hard decisions on the uncoded and coded bits, $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$, respectively. Using $\hat{\mathbf{c}}$ we can recover the data tones of the OFDM symbols by appropriate constellation mapping. This gives an estimate of $\hat{\mathbf{x}}_{\overline{\mathcal{I}}}$, which is multiplied by the channel frequency response $\mathbf{\Lambda}$, transformed to the time domain and subtracted from the received signal $\mathbf{r}$ to generate the estimate $\hat{\mathbf{e}}'$. Then we use $\hat{\mathbf{e}}'$ to update $\mathbf{a}$ and $\mathbf{b}$, through which the information extracted from the coding redundancy is transferred back to the impulsive noise estimator. As such, we form a decision feedback estimator that transfers information back-and-forth between the impulsive noise estimator using non-data tones and the decoder using data tones (Fig. 7). Compared to the estimator using all tones in Section V-C, the decision feedback estimator is expected to have better performance by exploiting the redundant information (i.e. coding structure) on the data tones.

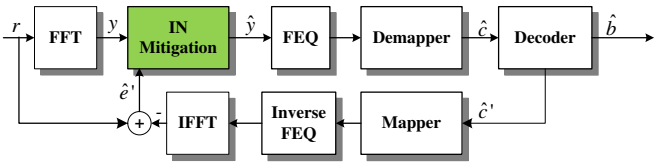

Fig. 7. A decision feedback impulsive noise estimator.

VI. Low-Complexity Implementation

The core SBL algorithm in (11) involves an $M \times M$ matrix inversion (M is the number of null and pilot tones) which might be practically infeasible on a hardware platform. An accelerated version of SBL that utilizes the properties of the marginal likelihood has been proposed in [41]. Using these properties, the accelerated SBL algorithm performs a sequential addition and deletion of candidate basis functions given by the columns of $\mathbf{\Phi}$ in (8), while keeping the same reconstruction performance [41].

In the context of impulsive noise estimation, an impulse at noise sample i is represented as a non-zero entry in vector $\mathbf{e}$ at index i. As a result, it contributes $e_i\mathbf{f}_i$ to the observed output $\mathbf{y}$, where we denote the i-th column in $\mathbf{F}_{\mathcal{I}}$ as $\mathbf{f}_i$. The accelerated algorithm will sequentially add, remove, or update basis $\mathbf{f}_i$ until convergence. At convergence, the basis that remain in the model will indicate the support of vector $\mathbf{e}$ and thereby the locations of the impulses. Algorithm 1 presents a high-level description of the sequential algorithm. Due to the space constraints, we refer interested readers to [41] for the mathematical details.

Algorithm 1 Sequential SBL algorithm [41]

1: Initialize background noise variance σ^2
2: Select $\mathbf{f}_0$ with largest projection $\|\mathbf{f}_0\mathbf{y}\|^2$ on the observed vector $\mathbf{y}$
3: Compute $\alpha_0 = \left(\|\mathbf{f}_0\mathbf{y}\|^2 - \sigma^2\right)^{-1}$, all other $\alpha_j = \infty$ (exclude from model)
4: Compute $\mathbf{\Sigma}$ and $\boldsymbol{\mu}$ as given by [41]
5: **while** $\Delta\alpha_j \leq$ threshold , $\forall j$ **do**
6: Select a candidate basis $\mathbf{f}_i$ from columns of F
7: Compute $\theta_i = \|q_i\|^2 - s_i$
8: **if** $\theta_i > 0$ and $\alpha_i < \infty$ **then**
9: $\mathbf{f}_i$ is in model, re-estimate α_i as given in [41]
10: **else if** $\theta_i > 0$ and $\alpha_i = \infty$ **then**
11: $\mathbf{f}_i$ not in model, add $\mathbf{f}_i$ and update parameters
12: **else**
13: $\mathbf{f}_i$ is in model, remove $\mathbf{f}_i$ and update parameters
14: **end if**
15: Update noise variance σ^2 as given in [41]
16: **end while**

Significant computational savings can be obtained if the background noise power σ^2 is known. This will allow for efficient calculations in steps 9, 11, 13 in algorithm 1 without any matrix inversions as given in [41]. TABLE III compares the complexity per iteration of the original SBL-based algorithms and the sequential implementation of it. The computational complexity of the original algorithms is dominated by the matrix multiplication and inversion operations in (13) and (20).

Estimator	Operation	Complexity
Using null and pilot Tones	Matrix multiply	$\mathcal{O}(N^2M)$
	Matrix inversion	$\mathcal{O}(M^3)$
Using all tones	Matrix multiply	$\mathcal{O}(N^3)$
	Matrix inversion	$\mathcal{O}(N^3)$
Sequential SBL w/ unknown background noise power	Matrix multiply	$\mathcal{O}(N^2K)$
	Matrix inversion	$\mathcal{O}(K^3)$
Sequential SBL w/ known background noise power	Matrix multiply	$\mathcal{O}(N^2K)$

TABLE III
Complexity per iteration of the proposed algorithms. N is the FFT size, M is the number of null and pilot tones, and K is the number of model basis in the current iteration.

Parameters	Simulation	G3 in CENELEC-A
FFT Length	128	256
Modulation	QPSK	DQPSK
# of Tones	128	128
# of Data Tones	72	36
# of Null Tones	56	92
FEC code	Rate-1/2 Convolutional	Rate-1/2 Convolutional
Interleaver	TDI or FDI	FDI
Interleave Size	0.5–1 AC cycles	up to 2.3 AC cycles

TABLE IV
Parameters of the simulated complex basedband OFDM system and the real passband OFDM system in the G3 standard operating in the CENELEC-A band. Interleaving is only simulated in periodic impulsive noise.

Compared to the estimator using non-data tones, the estimator using all tones increases the complexity from $\mathcal{O}(N^2M)$ per iteration to $\mathcal{O}(N^3)$ per iteration, where N is the DFT size. On the other hand, each iteration of the sequential SBL involves matrix multiplications and inversions that have complexities of $\mathcal{O}(N^2K)$ and $\mathcal{O}(K^3)$, respectively, where K is the number of model basis in that particular iteration. Furthermore, it can make use of the knowledge of background noise power to eliminate any matrix inversion operations.

VII. Simulation Results

To evaluate the performance of our proposed algorithms, we simulate a complex baseband OFDM system over a flat channel. The system parameters are listed in Table IV and compared with those in the G3 standard operating in the CENELEC-A band. In all simulations, we use the SBL algorithm in its full-complexity version, since the sequential SBL algorithm has the same reconstruction performance [41].

A. BER Performance in Asynchronous Impulsive Noise

We generate asynchronous impulsive noise from two different statistical models: a 3-component Gaussian mixture (GM) distribution with $\boldsymbol{\pi} = [0.9, 0.07, 0.03]$ and $\boldsymbol{\gamma} = [1, 100, 1000]$, and a Middleton Class A (MCA) distribution with $A = 0.1$, $\Omega = 0.01$, and the pdf truncated to the first 10 mixture components. The values of the model parameters are selected so that the impulsive-to-background noise power ratio is up to 30 dB in the GM noise, and 20 dB in the MCA noise, which reflect typical noise scenarios in the field measurement targeting BB PLC [5]. The noise samples are assumed to be independent and identically distributed (i.i.d.). Traces of

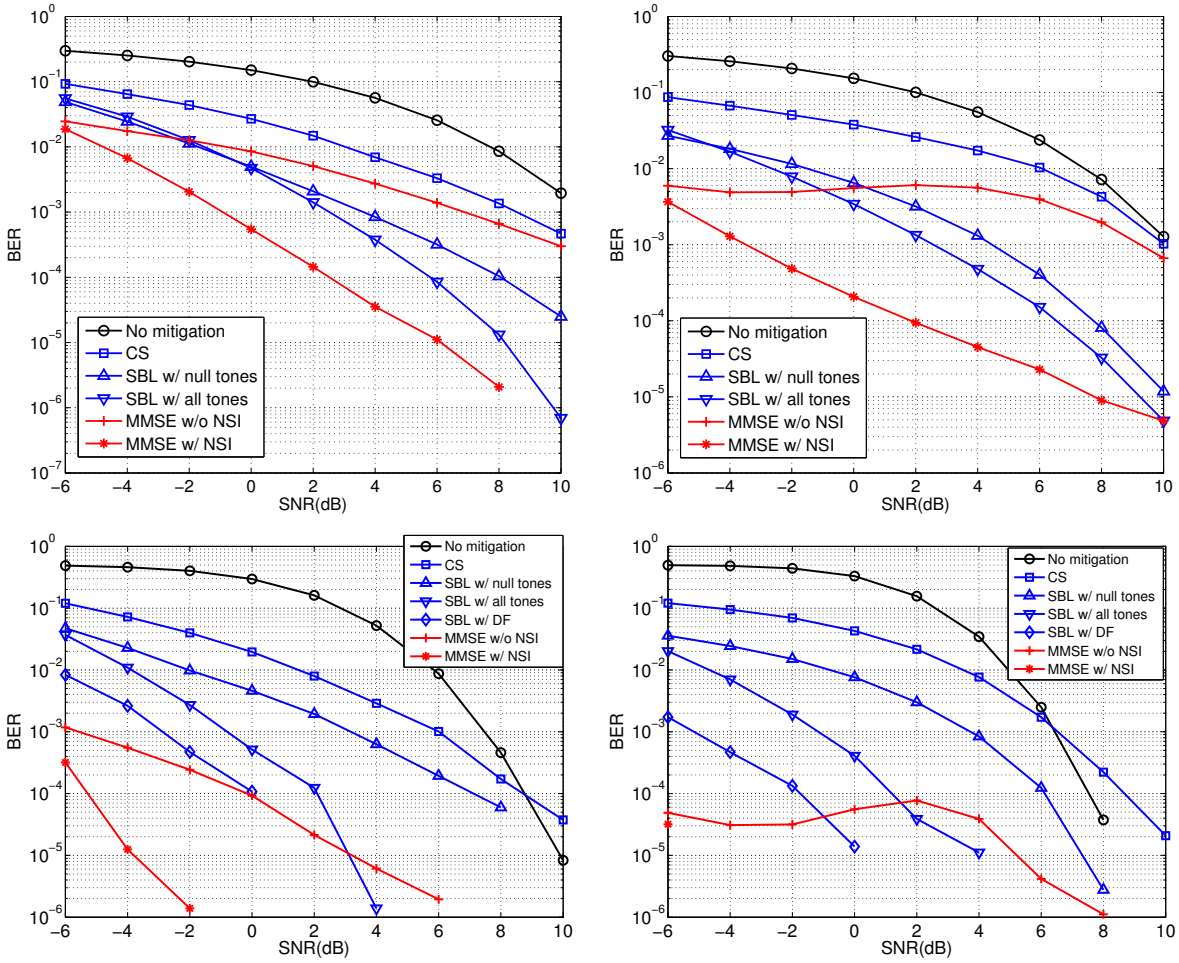

Fig. 8. Uncoded (top row) and coded (bottom row) BER performance of the proposed algorithms in Gaussian mixture (left column) and Middleton Class A (right column) modeled asynchronous impulsive noise, in comparison with the conventional OFDM system without noise mitigation, the compressed sensing based algorithm [29], and two parametric MMSE detectors with and without noise state information [13].

the simulated asynchronous impulsive noise have been shown earlier in Fig. 2 and 3.

In asynchronous impulsive noise, we test our proposed algorithms in the full-complexity version, and compare their BER performance with the compressed sensing based algorithm in [29]. To compare the performance of these non-parametric methods to the parametric ones, we also implement the two MMSE detectors in [13], since they are optimal in the MMSE sense among other parametric methods such as nulling and clipping [27]. Both MMSE detectors assume perfect knowledge of the GM model parameters (with the truncated MCA as a special case), and one even assumes complete noise state information (NSI), i.e., noise variance at each time instance. All algorithms (except for the SBL with decision feedback from the convolutional decoder) are simulated with and without the convolutional code, respectively.

The BER performance of all algorithms in uncoded and coded systems in different asynchronous impulsive noise scenarios are plotted in Fig. 8. For conciseness purposes, we denote the conventional OFDM system without noise mitigation as "No mitigation", the three proposed algorithms as "SBL w/ null tones", "SBL w/ all tones" and "SBL w/ DF", the compressed sensing based algorithm as "CS", and the two MMSE detectors as "MMSE w/o NSI" and "MMSE w/ NSI", respectively.

In the uncoded system, our proposed estimator using null tones achieves 6–8 dB SNR gain over conventional OFDM receivers. We can obtain additional 1–2 dB gain in a relatively wide SNR region by using all tones. The marginal performance loss of the estimator using all tones at lower SNRs is due to the error introduced by the continuous relaxation of constellation points $\mathbf{x}$ (see Section V-C). However, such error quickly becomes negligible as the SNR increases. All our proposed estimators outperform the MMSE detector without NSI in moderate to high SNR regimes. This does not even take into account the potential performance degradation of the MMSE detector due to parameter estimation errors. Moreover, our estimator using all tones reduces the SNR gap to the MMSE detector with NSI to as close as 1 dB at high SNRs. Note that the MMSE detector with NSI is practically infeasible since the NSI is unavailable at the receiver and cannot be estimated by training.

In the coded system, the proposed estimator using null tones can achieve up to 10 dB SNR gain over conventional OFDM

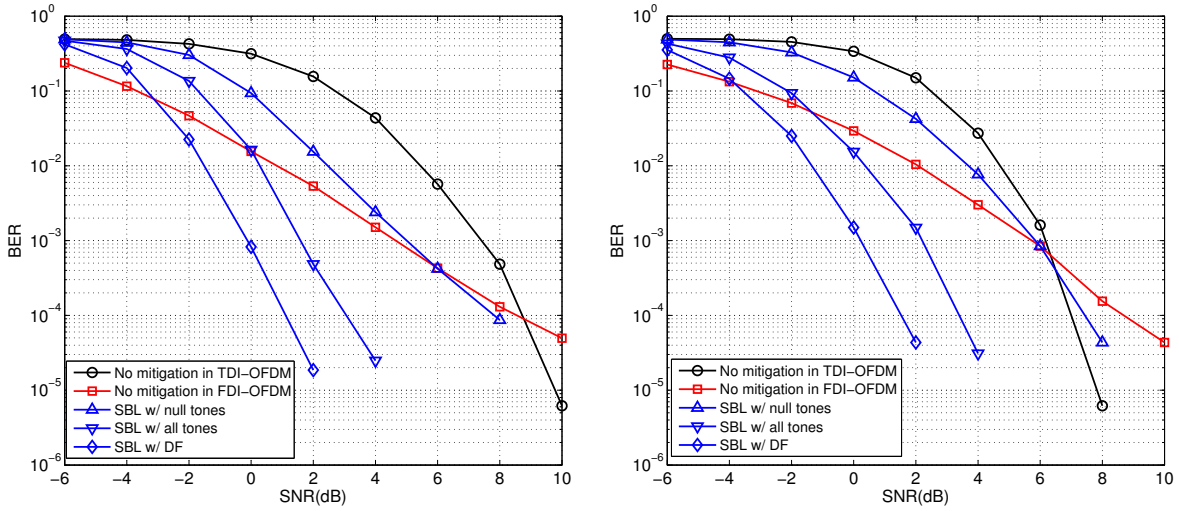

Fig. 9. Coded BER performance of the proposed algorithms in periodic impulsive noise, in comparison with the TDI-OFDM and FDI-OFDM systems without noise mitigation. The interleaving is done over an entire AC cycle. The burst interval varies from 10% (left) to 30% (right) of a period.

receivers. The estimator using all tones provides an additional 2–5 dB gains. Furthermore, using decision feedback from the convolutional decoder, we obtain an extra 2 dB gain. Again, the proposed estimators using all tones and decision feedback outperform the MMSE detector without NSI at moderate to high SNRs.

In all experiments, the compressed sensing based algorithm performs worse than our proposed estimators. As mentioned previously in Section III, this is because the compressed sensing algorithm can only recover the impulsive noise with high sparsity, i.e. typically less than 5 impulses per OFDM symbol in our system settings.

B. BER Performance in Periodic Impulsive Noise

We generate periodic impulsive noise using the LPTV system model in [7]. We divide one period of the noise into three intervals, each assuming an individual spectral shape (Fig. 4). The spectral shapes are fitted to noise measurement collected at an outdoor low-voltage site as shown in [7]. We vary the duration of noise bursts (i.e., the total duration of the second and the third intervals) from 10% to 30% of a period.

In periodic impulsive noise, we simulate our proposed algorithms in a coded TDI-OFDM system, and compare their BER performance with both TDI-OFDM and FDI-OFDM systems without noise mitigation. The parametric MMSE detectors in [13] cannot be applied in this case since the noise, either before or after the deinterleaver, does not follow Gaussian mixture distributions. In both TDI and FDI OFDM systems, we use two interleaver sizes, one spanning approximately half an AC cycle (i.e., one period of the noise), and the other about an entire AC cycle. Both interleaver sizes are smaller than the maximum interleaver size in G3, which according to Section IV spans 2.3 AC cycles.

With the interleaver size fixed at approximately an AC cycle, we increase the noise burst duration from 10% to 30% of a period. The BER performance of all algorithms are plotted in Fig. 9. Without any noise mitigation, the TDI-OFDM system

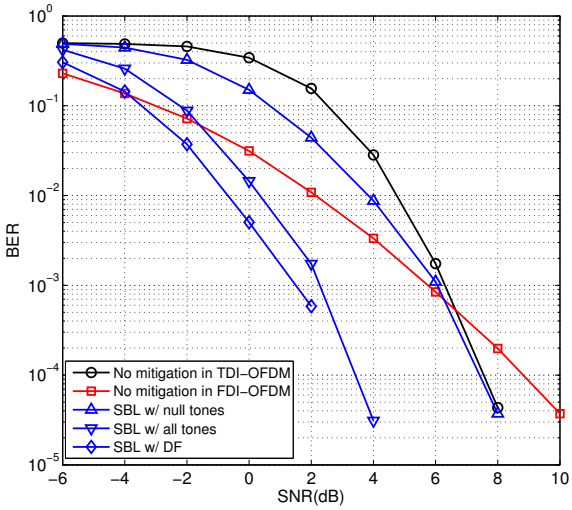

Fig. 10. Coded BER performance of the proposed algorithms in periodic impulsive noise, in comparison with the TDI-OFDM and FDI-OFDM systems without noise mitigation. The interleaving is done over half an AC cycle. The burst interval is fixed at 30% of a period.

performs worse than the conventional FDI-OFDM system until the SNR reaches 9 dB in the 10% burst case. This corresponds well to the results in [16] that the BER improvement of TDI-OFDM over FDI-OFDM can only be achieved above certain SNR threshold. By embedding the three SBL-based denoising algorithms into the TDI-OFDM framework, we are able to lower such SNR threshold to 6 dB, 0 dB and -3 dB, respectively. As the length of noise bursts increases to 30% of a period, the TDI-OFDM system without noise mitigation starts to show BER improvement over the FDI-OFDM system earlier at 7 dB. Embedding our SBL-based estimators into the TDI-OFDM system, especially the ones using all tones and decision feedback, further lowers the SNR threshold to about -1.5 dB and -4 dB, respectively. We notice that the SNR gains obtained by our proposed algorithms over the TDI-OFDM system itself are smaller than in the previous 10% burst case.

System	Noise	SBL w/ null tones	SBL w/ all tones	SBL w/ DF
Uncoded	GM	8 dB	10 dB	-
	MCA	6 dB	7 dB	-
Coded	GM	2 dB	7 dB	9 dB
	MCA	1.5 dB	6.3 dB	9.3 dB
	Periodic	0.8 dB	4.8 dB	6.8 dB

TABLE V
SNR GAINS (MEASURED AT BER=10^{-4}) OF THE PROPOSED IMPULSIVE NOISE MITIGATION ALGORITHMS OVER THE CONVENTIONAL OFDM SYSTEM WITHOUT INTERLEAVING IN GM AND MCA MODELED ASYNCHRONOUS IMPULSIVE NOISE, AND OVER THE FDI-OFDM SYSTEM IN PERIODIC IMPULSIVE NOISE WITH 30% BURST.

The SBL algorithm using null tones even performs slightly worse than the TDI-OFDM system without noise mitigation as the SNR grows above 6.5 dB. The reason is that in the 30% burst case, after deinterleaving, the number of impulses per OFDM symbol increases to a level where the performance of the SBL technique begins to saturate.

To demonstrate the robustness of our proposed algorithms to different interleaver sizes, we simulate the algorithms with a shorter interleaver spanning about half an AC cycle, while fixing the noise burst duration to 30% of a period. Since both interleaver sizes are an integer multiple of the noise period, in theory, after the deinterleaving, the noise within an OFDM symbol should have the same average sparseness. Therefore the same BER performance can be expected from our proposed algorithms. Comparing the BER performance in Fig. 10 to Fig. 9, we observe that decreasing the interleaver size leads to negligible effects on all BER curves, except for the marginal BER loss for the TDI-OFDM system without noise mitigation at SNRs above 6dB. This is because the TDI-OFDM system itself assumes AWGN, and a larger interleaver is useful to make noise samples within an OFDM symbol less correlated, i.e., closer to AWGN in statistics.

In all simulated noise scenarios, including both asynchronous and periodic impulsive noise, our proposed algorithms achieve significant BER improvement over conventional OFDM systems without noise mitigation in various SNR regions. For clarity purposes, we measured the approximate SNR gains of the proposed algorithms over the conventional OFDM system without any interleaving (in asynchronous impulsive noise), and with frequency-domain interleaving (in periodic impulsive noise) at a target BER of 10^{-4}, as listed in TABLE V.

VIII. CONCLUSION

This paper proposes three methods for improving communication performance of OFDM PLC systems in the presence of asynchronous impulsive noise and periodic impulsive noise. To mitigate asynchronous impulsive noise, we apply sparse Bayesian learning (SBL) techniques to estimate the impulsive noise from the received signal by observing information either on the null and pilot subcarriers or on all subcarriers. Under periodic impulsive noise, we adopt a time-domain interleaving OFDM transceiver structure to break long noise bursts that span multiple OFDM symbols into short bursts, and then apply the SBL techniques. All the 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 asynchronous impulsive noise and periodic impulsive noise simulated from various statistical models.

REFERENCES

[1] J. Lin, M. Nassar, and B. Evans, "Non-parametric impulsive noise mitigation in OFDM systems using sparse Bayesian learning," *Proc. IEEE Global Comm. Conf.*, 2011.

[2] G. Bumiller, L. Lampe, and H. Hrasnica, "Power line communication networks for large-scale control and automation systems," *IEEE Comm. Magazine*, vol. 48, no. 4, pp. 106–113, 2010.

[3] S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: The role of power line communications in the smart grid," *Proc. of the IEEE*, vol. 99, no. 6, pp. 998–1027, 2011.

[4] A. Moscatelli, "Technologies evolution for smart grid applications," in *IEEE Int. Symp. Power Line Comm. and Appl.*, 2011. [Online]. Available: http://www.ieee-isplc.org/2011/programme.html#keynotes

[5] M. Zimmermann and K. Dostert, "An analysis of the broadband noise scenario in powerline networks," *Proc. Int. Symp. Power Line Comm. and Appl.*, pp. 5–7, 2000.

[6] M. Nassar, K. Gulati, Y. Mortazavi, and B. Evans, "Statistical modeling of asynchronous impulsive noise in powerline communication networks," *Proc. IEEE Global Comm. Conf.*, pp. 1–6, 2011.

[7] M. Nassar, A. Dabak, I. Kim, T. Pande, and B. Evans, "Cyclostationary noise modeling in narrowband powerline communication for smart grid applications," *Proc. IEEE Int. Conf. on Acoustics, Speech and Sig. Proc.*, pp. 3089–3092, 2012.

[8] K. F. Nieman, J. Lin, M. Nassar, K. Waheed, and B. L. Evans, "Cyclic spectral analysis of power line noise in the 3-200 khz band," in *Proc. IEEE Int. Symp. Power Line Comm. and Appl.*, 2013.

[9] M. Nassar, J. Lin, Y. Mortazavi, A. Dabak, I. H. Kim, and B. Evans, "Local utility power line communications in the 3-500 khz band: Channel impairments, noise, and standards," *IEEE Sig. Proc. Mag.*, vol. 29, no. 5, pp. 116–127, 2012.

[10] Texas Instruments, "Data sheet for powerline communications analog front-end AFE031," 2012. [Online]. Available: http://www.ti.com/lit/ds/symlink/afe031.pdf

[11] J. Lin and B. Evans, "Cyclostationary noise mitigation in narrowband powerline communications," *Proc. APSIPA Annual Summit Conf.*, 2012.

[12] Y. Yoo and J. Cho, "Asymptotic analysis of CP-SC-FDE and UW-SC-FDE in additive cyclostationary noise," *Proc. IEEE Int. Conf. Comm.*, pp. 1410–1414, 2008.

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

[14] 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.

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

[16] A. Al-Dweik, A. Hazmi, B. Sharif, and C. Tsimenidis, "Efficient interleaving technique for OFDM system over impulsive noise channels," in *Proc. IEEE Int. Symp. Pers. Indoor and Mobile Radio Comm.*, 2010.

[17] 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.

[18] H. Meng, Y. L. Guan, and S. Chen, "Modeling and analysis of noise effects on broadband power-line communications," *IEEE Trans. Power Delivery*, vol. 20, no. 2, pp. 630–637, 2005.

[19] M. Chan and R. Donaldson, "Amplitude, width, and interarrival distributions for noise impulses on intrabuilding power line communication networks," *IEEE Trans. Electromag. Compat.*, vol. 31, no. 3, 1989.

[20] D. B. L., P. Caldera, D. Schwingshackl, and A. M. Tonello, "On noise modeling for power line communications," in *Proc. IEEE Int. Symp. Power Line Comm. and Appl.* IEEE, 2011, pp. 283–288.

[21] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromag. Compat.*, no. 3, pp. 106–127, 2007.

[22] S. M. Zabin and H. V. Poor, "Efficient estimation of class a noise parameters via the em algorithm," *IEEE Trans. Info. Theory*, vol. 37, no. 1, pp. 60–72, 1991.

[23] M. Katayama, T. Yamazato, and H. Okada, "A mathematical model of noise in narrowband power line communication systems," *IEEE J. Sel. Areas Comm.*, vol. 24, no. 7, pp. 1267–1276, 2006.

[24] 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. Sig. Proc.*, vol. 58, no. 12, pp. 6207–6222, 2010.

[25] M. Nassar, K. Gulati, A. Sujeeth, N. Aghasadeghi, B. Evans, and K. Tinsley, "Mitigating near-field interference in laptop embedded wireless transceivers," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Sig. Proc.*, 2008, pp. 1405–1408.

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

[27] S. V. Zhidkov, "Analysis and comparison of several simple impulsive noise mitigation schemes for ofdm receivers," *IEEE Trans. Comm.*, vol. 56, no. 1, pp. 5–9, 2008.

[28] M. Nassar and B. Evans, "Low complexity EM-based decoding for OFDM systems with impulsive noise," *Proc. IEEE Asilomar Conf. on Sig., Sys. and Computers*, pp. 1943–1947, 2011.

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

[30] L. Lampe, "Bursty impulse noise detection by compressed sensing," *Proc. IEEE Int. Symp. Power Line Comm. and Appl.*, pp. 29–34, 2011.

[31] W. Gardner and C. Spooner, "The cumulant theory of cyclostationary time-series. I. foundation," *IEEE Trans. Sig. Proc.*, vol. 42, no. 12, pp. 3387–3408, 1994.

[32] F. Bonnardot, R. Randall, and F. Guillet, "Extraction of second-order cyclostationary sourcesapplication to vibration analysis," *Mech. Sys. and Sig. Proc.*, vol. 19, no. 6, pp. 1230–1244, 2005.

[33] J. Cho, "Joint transmitter and receiver optimization in additive cyclostationary noise," *IEEE Trans. Info. Theory*, vol. 50, no. 12, 2004.

[34] R. García, L. Díez, J. Cortes, and F. Canete, "Mitigation of cyclic short-time noise in indoor power-line channels," *Proc. IEEE Int. Symp. Power Line Comm. and Appl.*, pp. 396–400, 2007.

[35] A. Liano, A. Sendin, A. Arzuaga, and S. Santos, "Quasi-synchronous noise interference cancellation techniques applied in low voltage PLC," *Proc. IEEE Int. Symp. Power Line Comm. and Appl.*, 2011.

[36] M. Luby, M. Watson, T. Gasiba, and T. Stockhammer, "High-quality video distribution using power line communication and application layer forward error correction," in *IEEE Int. Symp. Power Line Comm. and Appl.* IEEE, 2007, pp. 431–436.

[37] J. Mitra and L. Lampe, "Convolutionally coded transmission over markov-gaussian channels: Analysis and decoding metrics," *IEEE Trans. Comm.,*, vol. 58, no. 7, pp. 1939–1949, 2010.

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

[39] M. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

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

[41] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse bayesian models," *Proc. Int. Workshop on AI and Stats.*, 2003.

Jing Lin received a B.S. degree in electrical engineering in 2008 from Tsinghua University, China, and an M.S. degree in electrical and computer engineering in 2010 from the University of Texas at Austin, where she is currently pursuing a PhD degree. Her research has been focused on noise and interference mitigation in powerline communications for smart grid applications.

Marcel Nassar received a B.E. degree in computer and communications engineering, with a minor degree in mathematics, from the American University of Beirut in 2006, and an M.S. degree in electrical and computer engineering from The University of Texas at Austin in 2008, where he is currently pursuing a PhD degree. His interests lie at the intersection of signal processing and machine learning with application to statistical modeling and mitigation of interference in communication systems.

Brian L. Evans holds the Engineering Foundation Professorship at UT Austin. He earned his BSEECS (1987) degree from the Rose-Hulman Institute of Technology, and his MSEE (1988) and PhDEE (1993) degrees from the Georgia Institute of Technology. From 1993 to 1996, he was a post-doctoral researcher at the University of California, Berkeley. In 1996, he joined the faculty at UT Austin.

Prof. Evans research bridges the gap between digital signal processing theory and embedded real-time implementation. His research interests include wireless interference mitigation, smart grid communications, smart phone video acquisition, multicore computing. and cloud radio access networks.

Prof. Evans has published more than 200 refereed conference and journal papers, and graduated 20 PhD and 9 MS students. He received a Top 10% Paper Award at the 2012 IEEE Multimedia Signal Processing Workshop. He has received three teaching awards at UT Austin: Gordon T. Lepley Memorial ECE Teaching Award (2008), Texas Exes Teaching Award (2011) and HKN Outstanding ECE Professor (2012). He received a 1997 US National Science Foundation CAREER Award.