

Copyright  
by  
Jing Lin  
2014

The Dissertation Committee for Jing Lin  
certifies that this is the approved version of the following dissertation:

**Robust Transceivers to Combat Impulsive Noise  
in Powerline Communications**

Committee:

---

Brian L. Evans, Supervisor

---

Todd E. Humphreys

---

Alexis Kwasinski

---

Ahmed H. Tewfik

---

Haris Vikalo

**Robust Transceivers to Combat Impulsive Noise  
in Powerline Communications**

by

**Jing Lin, B.E., M.S.E.**

**DISSERTATION**

Presented to the Faculty of the Graduate School of  
The University of Texas at Austin  
in Partial Fulfillment  
of the Requirements  
for the Degree of

**DOCTOR OF PHILOSOPHY**

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2014

Dedicated to the memory of my grandfather.

## Acknowledgments

Foremost, I would like to express my sincere appreciation to my advisor Prof. Brian L. Evans for the continuous support of my PhD study and research. His commitment, enthusiasm and patience have enabled me to pursue my research interests, and to accumulate diverse experience and skills that will remain beneficial for my future career. In addition, I would like to thank the rest of my dissertation committee: Prof. Todd E. Humphreys, Prof. Alexis Kwasinski, Prof. Ahmed H. Tewfik and Prof. Haris Vikalo for their insightful comments and suggestions on this dissertation. My sincere thanks also goes to Prof. Andreas Gerstlauer, who served as a technical co-advisor during my Master's study on multiprocessor mapping for synchronous dataflow models. I have been greatly benefited from the insights I gained during my collaboration and discussion with Prof. Gerstlauer on system-level design and modeling, which also enriches the diversity of my academic experience.

I would like to express my special gratitude to Dr. Anuj Batra, Dr. Anand Dabak, Dr. Il Han Kim, Ms. Xiaolin Lu, Dr. Tarkesh Pande and other members of the Smart Grid R&D Business Unit at Texas Instruments, for offering me the challenging and inspiring summer internship opportunities. These internships as well as our collaborative research projects have motivated my dissertation research on powerline communications. In addition, the

knowledge and professionalism I learnt from them are tremendously helpful in my current study, and will remain so in my future career development.

I have been lucky to be surrounded by brilliant colleagues, who have provided invaluable feedback or suggestions on almost every aspect of my graduate study. I would like to thank Dr. Marcel Nassar, Mr. Karl F. Nieman, Mr. Yousof Mortazavi, Dr. Aditya Chopra and Mr. Akshaya Srivatsa, who have contributed bits and bytes to my research. In addition, I would like to thank my Chinese friends at UT Austin, for sharing the most enjoyable leisure time with me.

Last but not the least, I would like to thank my family: my father Yaohua Lin, my mother Manhua Gao, my husband Yanguang Lin, and my grandparents. Thank you for your most unselfish and generous love and support throughout my life.

# **Robust Transceivers to Combat Impulsive Noise in Powerline Communications**

Publication No. \_\_\_\_\_

Jing Lin, Ph.D.

The University of Texas at Austin, 2014

Supervisor: Brian L. Evans

Future smart grid systems will intelligently monitor and control energy flows in order to improve the efficiency and reliability of power delivery. This monitoring and control requires low-power, low-cost and highly reliable two-way communications between customers and utilities. To enable these two-way communication links, powerline communication (PLC) systems are attractive because they can be deployed over existing outdoor and indoor power lines. Power lines, however, have traditionally been designed for one-directional power delivery and remain hostile environments for communication signal propagation. In particular, non-Gaussian noise that is dominated by asynchronous impulsive noise and periodic impulsive noise, is one of the primary factors that limit the communication performance of PLC systems.

For my PhD dissertation, I propose transmitter and receiver methods to mitigate the impact of asynchronous impulsive noise and periodic impulsive

noise, respectively, on PLC systems. The methods exploit sparsity and/or cyclostationarity of the noise in both time and frequency domains, and require no or minor training overhead prior to data transmission. Compared to conventional PLC systems, the proposed transceivers achieve dramatic improvement (up to 1000x) in coded bit error rates in simulations, while maintaining similar throughput.



# Table of Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>List of Tables</b>	<b>xii</b>
<b>List of Figures</b>	<b>xiii</b>
<b>Chapter 1. Introduction</b>	<b>1</b>
1.1 Smart Grid Communications . . . . .	2
1.1.1 Wireless Communications . . . . .	3
1.1.2 Powerline Communications . . . . .	5
1.2 Interference in Powerline Communications . . . . .	9
1.3 Impact of Impulsive Noise in OFDM Systems . . . . .	12
1.4 Dissertation Summary . . . . .	14
1.4.1 Thesis Statement . . . . .	14
1.4.2 Summary of Contributions . . . . .	15
1.5 Organization . . . . .	17
1.6 List of Acronyms . . . . .	18
<b>Chapter 2. Background</b>	<b>22</b>
2.1 OFDM Systems . . . . .	22
2.2 Narrowband PLC Standards . . . . .	24
2.3 Statistical Modeling of Impulsive Noise . . . . .	30
2.3.1 Asynchronous Impulsive Noise Modeling . . . . .	31
2.3.2 Periodic Impulsive Noise Modeling . . . . .	32
2.4 The Bayesian Framework . . . . .	35
2.4.1 Conjugate Priors . . . . .	36
2.4.2 Expectation Maximization Algorithm . . . . .	38

2.5	Sparse Bayesian Learning . . . . .	39
2.5.1	Single Measurement Vector . . . . .	40
2.5.2	Sequential Sparse Bayesian Learning . . . . .	43
2.5.3	Multiple Measurement Vector . . . . .	44
2.6	Conclusion . . . . .	46
<b>Chapter 3. Nonparametric Mitigation of Impulsive Noise</b>		<b>48</b>
3.1	Introduction . . . . .	49
3.2	Prior Work . . . . .	50
3.2.1	Asynchronous Impulsive Noise Mitigation . . . . .	51
3.2.2	Periodic Impulsive Noise Mitigation . . . . .	53
3.3	System Model . . . . .	55
3.4	Nonparametric Impulsive Noise Estimation . . . . .	59
3.4.1	Estimation Using Null and Pilot Tones . . . . .	60
3.4.2	Estimation Using All Tones . . . . .	61
3.4.3	Decision Feedback Estimation . . . . .	63
3.5	Low-Complexity Implementation . . . . .	66
3.6	Simulation Results . . . . .	67
3.6.1	Performance in Asynchronous Impulsive Noise . . . . .	68
3.6.2	Performance in Periodic Impulsive Noise . . . . .	71
3.7	Conclusion . . . . .	74
<b>Chapter 4. Time-Frequency Modulation Diversity to Combat Periodic Impulsive Noise</b>		<b>80</b>
4.1	Introduction . . . . .	81
4.2	Prior Work . . . . .	82
4.3	Time-Frequency Diversity Modulation . . . . .	85
4.4	Soft-Output Diversity Demodulator . . . . .	89
4.5	Noise Power Estimation . . . . .	94
4.5.1	Problem Formulation . . . . .	97
4.5.2	Noise Power Spectrum Estimation Using T-MSBL . . . . .	99
4.6	Simulation Results . . . . .	105
4.6.1	Performance with Offline Noise Power Estimation . . . . .	107

4.6.2	Performance with Semi-Online Noise Power Estimation	109
4.7	Conclusion . . . . .	110
<b>Chapter 5.</b>	<b>Conclusion</b>	<b>112</b>
5.1	Summary of Contributions . . . . .	112
5.1.1	Nonparametric mitigation of impulsive noise . . . . .	112
5.1.2	Time-frequency modulation diversity to combat periodic impulsive noise. . . . .	113
5.2	Suggestions for Future Work . . . . .	114
<b>Bibliography</b>		<b>120</b>
<b>Vita</b>		<b>134</b>

## List of Tables

1.1	Categories of powerlines communication systems. . . . .	7
1.2	Typical pathloss values (in dB/km) for low-voltage (LV) and medium-voltage (MV) power lines [28]. . . . .	8
2.1	Physical layer parameters of PRIME, G3, IEEE P1901.2 and G.hnem standards for narrowband PLC over the CENELEC A and FCC bands. . . . .	27
2.2	Statistical-physical models of interference in PLC networks categorized by network types. Parameters are given in Figure 2.1 and $M$ is the number of interferers. . . . .	31
3.1	Complexity per iteration of the proposed algorithms. $N_c$ is the FFT size, $N_0$ is the number of null and pilot tones, and $N_B$ is the number of model bases in the current iteration. . . . .	67
3.2	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. . . . .	68
3.3	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. . . . .	73
4.1	Optimal parameters for the Hochwald/Sweldens code in <i>i.i.d.</i> flat Rayleigh fading and AWGN [42]. $N_d$ is the length of a codeword. . . . .	86
4.2	Parameters of the simulated OFDM system. . . . .	105

## List of Figures

1.1	A smart grid for smart metering applications. . . . .	4
1.2	The spectrogram of a noise trace measured in the 45–450 kHz band at an outdoor low-voltage power line [62]. . . . .	10
2.1	Interference scenarios in PLC networks. Each interferer emits a random sequence of emissions onto the power line, which add up at the receiver. An interferer is described statistically by a mean number of emission events $\mu$ , mean duration between emission events $\lambda$ , and the pathloss to the receiver $\gamma$ . . . . .	30
2.2	Asynchronous impulsive noise simulated from a Gaussian mixture distribution with $\boldsymbol{\pi} = [0.9, 0.07, 0.03]$ and $\boldsymbol{\gamma} = [1, 100, 1000]$ . . . . .	33
2.3	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. . . . .	33
2.4	A time-domain trace and spectrogram of the periodic impulsive noise synthesized from a linear periodically time varying system model. . . . .	34
2.5	The density of the prior $p(\mathbf{x}; \mathbf{a}, \mathbf{b})$ , where $\mathbf{x}$ is a length-2 vector, and the parameters $\boldsymbol{\gamma}$ have been integrated out. . . . .	41
3.1	A conventional baseband coded OFDM system, where +CP (-CP) means inserting (removing) the cyclic prefix to (from) each OFDM symbol, and FEQ stands for frequency-domain equalization. . . . .	55
3.2	A time-domain interleaved OFDM system. $\Pi$ denotes the sample-level interleaver, and $\Pi^{-1}$ the corresponding deinterleaver. . . . .	56
3.3	A decision feedback impulsive noise (IN) estimator for an OFDM receiver. . . . .	66
3.4	Uncoded (top) and coded (bottom) BER performance of the proposed algorithms in Gaussian mixture modeled asynchronous impulsive noise, in comparison with the conventional OFDM system without noise mitigation, the compressed sensing based algorithm, and two parametric MMSE detectors with and without noise state information. . . . .	76

3.5	Uncoded (top) and coded (bottom) BER performance of the proposed algorithms in Middleton Class A modeled asynchronous impulsive noise, in comparison with the conventional OFDM system without noise mitigation, the compressed sensing based algorithm, and two parametric MMSE detectors with and without noise state information. . . . .	77
3.6	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% (top) to 30% (bottom) of a period. . . . .	78
3.7	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 the AC cycle. The burst interval is 30% of a period. . . . .	79
4.1	An example of time-frequency modulation diversity for length-two MD codes. Components of an MD codeword (marked in the same color) are allocated to subcarriers $i$ and $((i + \Delta K) \bmod N_c)$ in OFDM symbols $j$ and $(j + \Delta T)$ , respectively, $\forall i, j$ , where $N_c$ is the number of subcarriers in an OFDM symbol. . . . .	88
4.2	The block diagram of a narrowband PLC transmitter using time-frequency modulation diversity. The proposed time-frequency (TF) mapper allocates components of every Hochwald/Sweldens codeword to different subcarriers in various OFDM symbols. $\Pi$ denotes the bit-level block interleaver. . . . .	89
4.3	The block diagram of a narrowband PLC receiver with non-coherent diversity demodulation. The time-frequency (TF) demapper groups subcarriers in OFDM symbols into MD codewords. The diversity demodulator outputs soft information based on the noise power spectrum, which can be estimated before or during transmission (as will be discussed in Section 4.5). $\Pi^{-1}$ denotes the bit-level block de-interleaver. . . . .	94
4.4	Coded BER performance of proposed TFMD transceiver in both non-coherent (left) and coherent (right) systems, using offline noise power estimation. Performance is compared with conventional DBPSK (or BPSK) OFDM systems with bit-level interleaving over 1 or 4 OFDM symbols, and with the frequency modulation diversity (FMD) system. $N_d$ denotes the length of MD codewords. . . . .	108

4.5	Coded BER performance of proposed TFMD transceiver in coherent systems, using semi-online noise power estimation. Both length-two (left) and length-three (right) MD codes are tested. Performance is compared with conventional BPSK OFDM system with bit-level interleaving over 1 OFDM symbol, the TFMD system using offline noise power estimation, and the TFMD system with heuristic semi-online noise power estimation. . . . .	111
-----	--	-----

## List of Algorithms

1	T-MSBL [91] . . . . .	47
2	Impulsive noise estimation using null and pilot tones . . . . .	60
3	Impulsive noise estimation using all tones . . . . .	64
4	Impulsive noise estimation using decision feedback . . . . .	75



# Chapter 1

## Introduction

The growth of energy demand has outpaced the rate at which generation capacity can grow by traditional means. The International Energy Outlook 2013 projects that world energy consumption will grow by 56% from 2010 to 2040 [8]. The rise in energy consumption is not only due to population growth, but also the proliferation of consumer electronics and other plug-in devices such as electric vehicles. Traditional ways to meet increasing demand include constructing large-scale power plants and laying new transmission lines, which are expensive and time-consuming. For example, a 1000 MW nuclear plant costs up to \$2B and takes over 15 years to build; and laying transmission lines costs \$0.6M per kilometer and takes 5–10 years [7].

To solve the energy crisis in a more sustainable way, distributed energy resources (DER) are used to provide an alternative to or an enhancement of the traditional centralized power plants [7]. DERs are small-scale power generation sources (typically in the range of 1 kW to  $10^4$  kW), including renewable energy resources (e.g. wind turbines and solar panels) and decentralized energy storage (e.g. the battery of electric vehicles and photovoltaic batteries). These resources may deliver low-cost power to the grids during peak hours,

or provide standby energy for emergency uses. In addition, they also have a lower impact to the environment compared to traditional power plants.

The integration of DERs have transformed the power grids into much more dynamic and complex systems than they used to be. To manage such systems, a large amount of information needs to be measured, communicated and analyzed in real time. On the utility's side, transducers such as the Phasor Measurement Units are deployed over the grids to precisely measure AC voltages and currents at high speeds (typically 30 observations per second), with time resolutions better than  $1 \mu s$ . These measurements provide grid operators with a picture of real-time grid conditions, and are helpful to speed up their response in the case of unexpected demand disturbance and power outage. On the customer's side, advanced metering infrastructure is deployed to implement time-dependent and load-dependent electricity rates and to make energy usage profiles available to the customer. Such information gives them more incentives to better manage their energy consumption. During peak hours, the advanced metering infrastructure also enables customers to offer power-to-grid from their personal owned DERs at the market price.

## **1.1 Smart Grid Communications**

The desire to use two-way flow of information to create an intelligent energy delivery network has inspired the concept of the "smart grid". A smart grid couples a traditional power grid with a communication network. Within a smart grid, communication technologies for smart metering applications have

recently obtained great interest. Such communication network typically consists of three primary components [75], as illustrated in Figure 1.1:

- Home area networks (HAN) that connect smart appliances and sensors on indoor power lines to smart meters;
- Neighborhood area networks (NAN) that connect smart meters to data concentrators that are deployed by local utilities on medium-voltage (MV) lines (in the US) or low-voltage (LV) lines (in Europe); and
- Communication backhaul that carries traffic between data concentrators and local utilities.

Smart grid communications will likely be supported by a heterogeneous set of network technologies, ranging from wireless to powerline, since no single solution fits all scenarios [28]. In this section, I give a brief overview of the wireless and powerline communication technologies to support the three types of communication links in Figure 1.1 for smart metering applications.

### **1.1.1 Wireless Communications**

Among the wireless solutions, cellular communication was once the primary communication technology in smart meters, thanks to the rapid development of mobile networks [6]. In particular, WiMAX operating in the 2.3, 2.5 and 3.5 GHz bands was used to provide up to 70 Mbps over distances up to 48 km [36, 73]. A major disadvantage of cellular technologies is the high



door communications between smart meters and smart appliances at multiple data rates in the 450 MHz–2.4 GHz band. The IEEE 802.11ah standard for sensor network applications, including smart metering, targets data rates of a few hundred kbps and communication range of 200 meters in the sub-1GHz unlicensed bands.

### 1.1.2 Powerline Communications

In the wireline alternatives, fiber optic networks have been deployed in parts of the US as the communication backhaul connecting data concentrators to local utilities [28]. Fiber optics has the advantages of high transport capacity and the immunity to electromagnetic interference and radio frequency interference. However, due to the high deployment and maintenance cost, it has been limited to backhaul communication links.

Powerline communications (PLC), thanks to the widespread availability of power line infrastructure, is considered as a low-cost alternative for communications in home area networks and neighborhood area networks. Depending on the operating bands, PLC systems can be divided into three categories [28, 67]:

1. Ultra-Narrowband (UNB) PLC. UNB PLC systems operate in the 0.3–3 kHz band to provide about 100 bps over long distances (over 150 km). An example of such systems is the Two-Way Automatic Communications System that has been deployed by hundreds of utilities over the last

two decades for automatic meter reading, outage detection and voltage monitoring [28, 67].

2. Narrowband (NB) PLC. This type of PLC system operates in the 3–500 kHz bands, including the European CENELEC (Committee for Electrotechnical Standardization) bands (3–148.5 kHz), the US FCC (Federal Communications Commission) band (10–490 kHz), the Japanese ARIB (Association of Radio Industries and Businesses) band (10–450 kHz) and the Chinese band (3–500 kHz). Narrowband PLC systems are able to deliver either a few kbps using single-carrier communication, or several hundred kbps using multicarrier communication such as Orthogonal Frequency Division Multiplexing (OFDM). The former, a.k.a. low data rate narrowband PLC, has been used for commercial building automation (e.g. automatic lighting and air conditioning), and has been standardized in BacNet and LonTalk [28, 67]. The latter, a.k.a. high data rate narrowband PLC, has gained tremendous interest as a solution for neighborhood area networks in smart grid communications. In fact, it is currently the most adopted (60% market share) communication technology in smart meters [61]. High data rate narrowband PLC systems are exemplified in the industry developed standards G3 [3] and PRIME (Powerline Intelligent Metering Evolution) [5], and the recent international standards IEEE P1901.2 [2] and ITU-T G.hnem [72].
3. Broadband (BB) PLC. Broadband PLC systems target home area networks, and use OFDM or wavelet OFDM to provide several Mbps to

Categories	Operating Bands	Data Rates	Standards
UNB PLC	0.3–3 kHz	~100 bps	-
NB PLC	3–500 kHz	several kbps	BacNet, LonTalk
		up to 800 kbps	PRIME, G3 IEEE P1901.2 ITU-T G.hnem
BB PLC	1.8–250 MHz	up to 200 Mbps	HomePlug 1.0 IEEE P1901 ITU-T G.hn

Table 1.1: Categories of powerlines communication systems.

several hundred Mbps in the 1.8–250 MHz band. Standards for broad-band PLC include TIA-1113 (a.k.a. HomePlug 1.0), IEEE P1901 and ITU-T G.hn recommendations.

Table 1.1 summarizes the operating bands, supported data rates and standards of the three categories of PLC systems.

Apart from the low deployment costs, one of the advantages that make PLC an appealing candidate for smart grid communications is the predictable propagation channel. While wireless channels are likely to be time varying due to mobility in the propagation environment, transfer function of a powerline channel, as long as it does not cross the MV-LV transformer, can be determined based on cable properties and grid topology [27, 67]. Furthermore, pathloss over LV and MV power lines is much smaller than over typical wireless channels (Table 1.2). In particular, PLC can work where radio propagation is weak or completely blocked. For example, a data concentrator in the basement of a building can barely send or receive wireless signal, while PLC on the other

Power Lines	100 kHz	10 MHz
LV	1.5-3	160-200
MV (Overhead)	0.5-1	30-50
MV (Underground)	1-2	50-80

Table 1.2: Typical pathloss values (in dB/km) for low-voltage (LV) and medium-voltage (MV) power lines [28].

hand can traverse the power lines to reach smart meters in apartment units upstairs.

However, PLC needs to overcome several communication challenges to make itself practical. In the US, since data concentrators are typically deployed on MV lines, communication between data concentrators and smart meters need to traverse across MV-LV transformers. One of the primary channel impairments in this situation is the severe attenuation on signal crossing MV-LV transformers. The attenuation is frequency selective and periodically varying [67], and causes significant performance degradation.

On the other hand, as many wireless communication networks, PLC systems are interference limited. A communication system is said to be interference limited if the interference power from other communication and non-communication devices is increasingly dominating the background noise power. In PLC, such noise and interference 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 [28]. The interference may also lead to severe performance deterioration in PLC systems. This dissertation focuses on addressing the challenge of combating noise and interference in PLC



systems.

## 1.2 Interference in Powerline Communications

Broadly speaking, interference or noise refers to the disturbance energy, resulting from either natural or man-made sources, that adds to the transmitted signal at the receiver and degrades its ability to successfully detect the transmitted information. Throughout this dissertation, I use noise and interference interchangeably. In the 3–500 kHz (narrowband PLC) and 1.8–250 MHz (broadband PLC) bands, powerline noise can generally be decomposed into four classes [21, 67, 70, 96]:

1. *Spectrally-shaped background noise.* The background noise is the summation of numerous low-power noise sources. Its power spectral density slowly varies over time (in minutes or even hours), and exhibits a  $1/f$ -type decay due to the decreasing concentration of noise sources with frequency.
2. *Periodic impulsive noise.* This category of noise includes all noise components whose statistics vary periodically with half the AC cycle. The aggregated noise exhibits cyclostationarity in both time and frequency domain (Figure 1.2). The periodic impulsive noise can be further decomposed into two distinctive sub-components:
  - (a) Periodic impulsive noise synchronous to the main powerline frequency. This type of noise consists of a series of isolated impulses

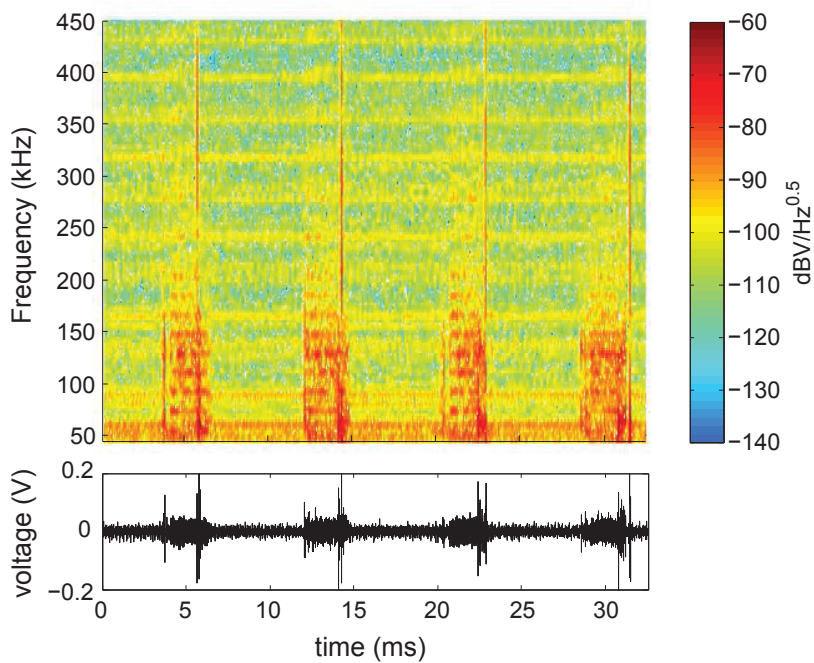


Figure 1.2: The spectrogram of a noise trace measured in the 45–450 kHz band at an outdoor low-voltage power line [62].

of considerable duration and amplitude. The impulses has a repetition rate equal to twice the main powerline frequency and always appear at the same instant of the AC cycle. They are typically caused by nonlinear power electronic devices, such as silicon controlled rectifiers and diodes, that switch on and off with the AC cycle while generating abrupt switching transients.

- (b) Periodic impulsive noise asynchronous to the main powerline frequency. This noise component takes the form of impulse trains with repetition rates unrelated to and much higher than the main

powerline frequency. In addition to the high repetition frequencies, it also exhibits an underlying period equal to half the AC cycle, i.e., it is cyclostationary. The impulses in such noise typically have much lower amplitudes and shorter durations than the ones in the synchronous impulsive noise. In the frequency domain, such noise can be clearly identified as harmonic clusters with significantly higher power spectral density (more than 30 dB) over the background noise. A primary contributor to this type of noise is switching mode power supplies, such as inverters and DC-DC converters, which contain MOSFET switches operating at frequencies above 20 kHz and up to several hundred kHz. These circuits output inband harmonic contents that cannot be perfectly removed by analog filtering.

3. *Asynchronous impulsive noise.* This type of noise consists of short duration, high power impulses (up to 50 dB above background noise power [95]) 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 [28] is shown to be asynchronous impulsive noise in nature [65].
4. *Narrowband interference.* Broadcast stations in the long-wave (153–279 kHz), medium-wave (540 kHz–1.61 MHz in the US) and short-wave (2.3

MHz–26.1 MHz) bands introduce narrowband interference to PLC systems. The interference exhibits amplitude or frequency modulated sinusoids in the time domain, which correspond to harmonic clusters in the noise spectrum. The interference level generally varies slowly with daytime, and in some cases varies periodically with half the AC cycle.

Recent field measurements on both MV and LV lines have identified the dominant noise component in the 3–500 kHz band to be periodic impulsive noise [46, 62, 70]. In the 1.8–250 MHz band, asynchronous impulsive noise becomes the dominant noise component.

### **1.3 Impact of Impulsive Noise in OFDM Systems**

Both narrowband and broadband PLC standards adopt OFDM as the modulation technique, since it offers great advantages in combating frequency selective channel. Furthermore, OFDM is inherently more robust against impulsive noise than single-carrier communication systems.

Compared to single-carrier systems, OFDM is more resilient to asynchronous impulsive noise, whose impulses are much shorter than an OFDM symbol. The inverse Fast Fourier Transform (IFFT) at the transmitter can be considered as a precoder that spreads the information carried by each subcarrier over all samples in the time domain. This provides time-domain diversity that allows successful recovery of the information even if a few samples are corrupted by noise impulses. Alternatively, the Fast Fourier Transform (FFT)

operation at the receiver smears the impulsive energy, and hence averages its impact across all subcarriers. As such, OFDM provides significant gain over single-carrier systems at moderate to high SNR values [56]. However, in the low SNR regime, the smeared impulsive energy significantly raises the noise power over all subcarriers, causing dramatic performance degradation in conventional OFDM systems.

OFDM systems also provide inherent immunity to periodic impulsive noise, in which a single noise burst may span multiple consecutive OFDM symbols<sup>1</sup>. As mentioned in Section 1.2, periodic impulsive noise consists of strong harmonic components similarly to narrowband interference (NBI). As such, only the subcarriers close to the harmonic frequencies will be severely corrupted while the rest are affected by low-power background noise. In other words, not all data symbols are contaminated during the noise bursts as would happen in single-carrier systems. In spite of this, the rich harmonic contents in the noise power spectrum still cause significant corruption over a considerable portion of the transmission band, and hence limits the bit error rate (BER) performance of the OFDM system.

---

<sup>1</sup>For example, the OFDM symbol duration in the G3 standard operating in the CENELEC-A band from 3–95 kHz is 695  $\mu$ s (see Table 2.1). Typical noise bursts last from 10% to 30% of a period (i.e., half the AC cycle). A noise burst lasting for 30% of a period will contaminate up to 4 consecutive OFDM symbols.

## 1.4 Dissertation Summary

This dissertation aims to improve communication performance of OFDM-based PLC systems in the presence of asynchronous impulsive noise and periodic impulsive noise, respectively, without significantly reducing throughput. I seek to achieve the goal using two distinctive approaches.

In the first approach, I design nonparametric noise mitigation algorithms at the receiver, which estimate and subtract asynchronous impulsive noise or periodic impulsive noise from received signal. Using recent results in compressed sensing, the algorithms exploit sparsity of impulsive noise in time domain to estimate the noise impulses from their projection on various sub-carriers of received signal. The algorithms do not assume prior knowledge on the statistical noise models, and hence do not impose training overhead prior to data transmission.

The second approach, on the other hand, is a parametric, joint transmitter and receiver design to combat periodic impulsive noise. The proposed transceivers use modulation diversity in both time and frequency domains, which exploits the knowledge on the periodically varying and spectrally shaped sub-channel SNRs to improve communication reliability. The noise power spectral density can be estimated primarily during data transmission, which entails minimal training overhead prior to data transmission.

### 1.4.1 Thesis Statement

In this dissertation, I defend the following thesis statement:

*Reliability of smart grid communications over power lines can be dramatically improved without sacrificing throughput by exploiting sparsity and cyclostationarity of the impulsive noise in both time and frequency domains.*

#### **1.4.2 Summary of Contributions**

The primary contributions of this dissertation can be summarized as follows.

##### **1. Nonparametric mitigation of asynchronous impulsive noise.**

In this contribution, I propose nonparametric algorithms at OFDM PLC receivers to estimate and subtract asynchronous impulsive noise from received signal. Exploiting the sparsity of the noise in time domain, I apply sparse Bayesian learning (SBL) techniques and propose three iterative algorithms, with different complexity vs. performance trade-offs, that

- (a) utilize the noise projection onto null and pilot tones;
- (b) add the information in the data tones to perform joint noise estimation and symbol detection; and
- (c) use decision feedback from the decoder to further enhance the accuracy of noise estimation.

All the methods are nonparametric, i.e., they do not require prior knowledge on the statistical noise model or model parameters. In the simulations, the proposed receivers provide up to 9 dB gain in signal-to-noise power ratio

(SNR), or alternatively over 1000x reduction in coded BER, compared to conventional coherent receivers without impulsive noise mitigation.

## **2. Nonparametric mitigation of periodic impulsive noise.**

In this contribution, I utilize the three algorithms developed in the previous contribution to mitigate periodic impulsive noise. Since periodic impulsive noise consists of long noise bursts that span multiple OFDM symbols, I adopt a time-domain block interleaving OFDM transceiver structure, where the deinterleaver at the receiver spreads noise bursts into short impulses. I then apply the noise mitigation algorithms developed in the previous contribution to estimate and remove the short impulses from received signal. In the simulation results, the proposed receivers provide up to 6.8 dB gain in SNR, or alternatively over 100x reduction in coded BER, compared to conventional coherent receivers with frequency domain interleaving and without impulsive noise mitigation.

## **3. Time-frequency modulation diversity to combat periodic impulsive noise.**

In this contribution, I develop a robust transmission scheme and corresponding receiver methods to combat periodic impulsive noise. Towards that end, I propose (1) a time-frequency modulation diversity scheme at the transmitter and a diversity receiver to improve communication reliability without decreasing throughput; and (2) an online noise power estimator that exploits sparsity of periodic impulsive noise in the frequency domain



to estimate the noise power spectrum for reliable decoding at the diversity receiver. The simulation results show more than 1000x reductions in coded BER compared to conventional coherent and non-coherent OFDM systems at typical SNRs in narrowband PLC.

## 1.5 Organization

The rest of the dissertation is organized as follows.

Chapter 2 gives a brief overview of several digital communication concepts and statistical methodologies used in this dissertation. It starts with the general OFDM system model, and describes various OFDM communication and signal processing techniques that are tailored and adopted in current PLC standards. Then it introduces the statistical models for asynchronous impulsive noise and periodic impulsive noise, respectively. After that, it presents the Bayesian inference framework, followed by a detailed description of the sparse Bayesian learning algorithms.

Chapter 3 presents the first two contributions on nonparametric mitigation of asynchronous impulsive noise and periodic impulsive noise, respectively. It starts with a literature review on existing receiver methods to combat both types of impulsive noise. Then it describes the system model, and proposes three nonparametric impulsive noise mitigation algorithms. The chapter is closed with a discussion of simulation results.

Chapter 4 discusses the third contribution on modulation diversity to

combat periodic impulsive noise. First, it discusses previous transmitter and receiver methods to combat periodic impulsive noise. After a brief review of modulation diversity techniques, it introduces the proposed time-frequency modulation diversity, and the diversity combining demodulator. Then it discusses the offline estimator for the noise power spectrum, and proposes an online alternative that reduces the training overhead prior to data transmission. Simulation results are shown and discussed at the end of the chapter.

Finally, Chapter 5 concludes the dissertation with a summary of proposed contributions and an outline of interesting perspectives for future research.

## 1.6 List of Acronyms

<b>AC</b>	alternating current
<b>AMI</b>	advanced metering infrastructure
<b>AR</b>	autoregressive
<b>ARIB</b>	Association of Radio Industries and Businesses
<b>ASK</b>	Amplitude shift keying
<b>AWGN</b>	additive white Gaussian noise
<b>BB</b>	broadband
<b>BER</b>	bit error rate
<b>BPSK</b>	binary phase shift keying
<b>CEN A</b>	CENELEC-A
<b>CENELEC</b>	Committee for Electrotechnical Standardization

<b>Conv</b>	Convolutional
<b>CP</b>	cyclic prefix
<b>DARMA</b>	deseasonalized autoregressive moving average
<b>DC</b>	direct current
<b>DCM</b>	dual carrier modulation
<b>DER</b>	distributed energy resource
<b>DPSK</b>	differential phase shift keying
<b>EM</b>	expectation maximization
<b>FCC</b>	Federal Communications Commission
<b>FDI-OFDM</b>	frequency-domain block interleaving OFDM
<b>FEC</b>	forward error correction
<b>FEQ</b>	frequency-domain channel equalizer
<b>FFT</b>	Fast Fourier Transform
<b>GMM</b>	Gaussian mixture model
<b>HAN</b>	Home area network
<b>HS</b>	Hochwald/Sweldens
<b>IEEE</b>	Institute of Electrical and Electronics Engineers
<b>IFFT</b>	inverse Fast Fourier Transform
<b>ISI</b>	inter-symbol interference
<b>ITU-T</b>	International Telecommunication Union Telecommunication Standardization Sector
<b>LLR</b>	log-likelihood ratio
<b>LPTV</b>	linear periodically time varying

**LTI** linear time invariant

**LV** low-voltage

**MIMO** multiple-input multiple-output

**MRC** maximum ratio combining

**NB** narrowband

**NBI** narrowband interference

**NMLSE** non-coherent maximum likelihood sequence estimator

**MAP** maximum *a posteriori*

**MCA** Middleton Class A

**MMSE** minimum mean square error

**MMV** multiple measurement vector

**MOSFET** metaloxidesemiconductor field-effect transistor

**MSDD** multiple-symbol differential detector

**MV** medium-voltage

**NAN** Neighborhood area network

**OFDM** orthogonal frequency division multiplexing

**PHY** physical layer

**PLC** powerline communications

**PRIME** PoweRline Intelligent Metering Evolution

**QAM** quadrature amplitude modulation

**PSK** phase shift keying

**RS** Reed-Solomon

**SBL** sparse Bayesian learning

**SC** selection combining  
**SER** symbol error rate  
**SMV** single measurement vector  
**SNR** signal-to-noise power ratio  
**TDI-OFDM** time-domain block interleaving OFDM  
**T-SBL** temporally correlated SMV sparse Bayesian learning  
**T-MSBL** temporally correlated MMV sparse Bayesian learning  
**UNB** ultra-narrowband  
**WiMAX** Worldwide Interoperability for Microwave Access

# Chapter 2

## Background

In this chapter, I review several digital communication concepts and statistical methodologies upon which my contributions are based. I begin in Section 2.1 by introducing the general system model for OFDM. In Section 2.2, I briefly describe how various OFDM communication and signal processing techniques are tailored and applied in current PLC standards. In Section 2.3, I give an overview of the statistical models for asynchronous impulsive noise and periodic impulsive noise, respectively. I then turn to discussing the Bayesian inference framework in Section 2.4. It provides a basis for the sparse Bayesian learning algorithms for the single measurement vector model and the multiple measurement vector model, as discussed in Section 2.5.

### 2.1 OFDM Systems

As mentioned in the previous chapter, OFDM has been adopted in existing narrowband PLC and broadband PLC standards, since it offers great advantages in combating frequency selective channel and impulsive noise. This section briefly describes the system model of OFDM. The purpose is to allow a better understanding of the impact of impulsive noise on OFDM systems (as

described in Section 1.3), and to define a basic system model upon which my proposed algorithms will be built.

OFDM divides the transmission band into a number of equal-width subbands. Each subband is associated with a subcarrier sinusoid whose frequency equals to the center frequency of the subband. There are in general three types of subcarriers (a.k.a. tones): data subcarriers, null subcarriers and pilot subcarriers. Data subcarriers modulate data symbols. Pilot subcarriers carry known reference symbols that may be used by the receiver for channel estimation, synchronization, and other purposes. Null subcarriers are loaded with zero power, and are typically located on both edges of the transmission band to reduce out-of-band emissions, ease design of the analog transmit and receive filter and so on. All subcarriers are orthogonal to each other, the collection of which constitutes an OFDM symbol in the frequency domain.

The modulation of an OFDM symbol is generally implemented by inverse fast Fourier transform (IFFT). Let  $N_c$  denote the total number of subcarriers in an OFDM system. Define  $\mathbf{s}$  as a length- $N_c$  vector consisting of complex signals modulated on all subcarriers; i.e.,  $\mathbf{s}$  is an OFDM symbol in the frequency domain. An OFDM modulator takes the IFFT of  $\mathbf{s}$  to obtain the time-domain OFDM symbol. A cyclic prefix (CP), which is a replica of the last  $N_p$  samples of the time-domain OFDM symbol, is inserted to the beginning of the symbol to reduce inter-symbol interference (ISI). The time-domain OFDM symbol with CP is then transmitted through a multipath fading channel and is further corrupted by additive noise and interference at the receiver. The

receiver removes the CP from each OFDM symbol, which results in

$$\mathbf{r}^{(t)} = \mathbf{H}\mathbf{F}^*\mathbf{s} + \mathbf{e} + \mathbf{n}.$$

Here  $\mathbf{r}^{(t)}$  denotes the received signal in time domain after CP removal,  $\mathbf{e}$  and  $\mathbf{n}$  represent the time-domain impulsive noise and background noise, respectively,  $\mathbf{F}$  is the  $N_c$ -point FFT matrix, and  $\mathbf{H}$  is called the convolutional channel matrix. Assuming that the CP is longer than the channel delay spread (i.e., length of the channel impulse response),  $\mathbf{H}$  is a circulant matrix whose rows are circular shifts of the channel impulse response.

The receiver demodulates an OFDM symbol by taking the FFT of  $\mathbf{r}^{(t)}$ :

$$\begin{aligned} \mathbf{r} &= \mathbf{F}(\mathbf{H}\mathbf{F}^*\mathbf{s} + \mathbf{e} + \mathbf{n}) \\ &= \mathbf{\Lambda}\mathbf{s} + \mathbf{F}\mathbf{e} + \mathbf{F}\mathbf{n}. \end{aligned} \tag{2.1}$$

It can be shown that since  $\mathbf{H}$  is a circulant matrix,  $\mathbf{\Lambda} \triangleq \mathbf{F}\mathbf{H}\mathbf{F}^*$  is a diagonal matrix, with  $\{\Lambda_{ii}\}_{i=1}^{N_c}$  equivalent to the  $N_c$ -point FFT coefficients of the channel impulse response. Effectively, the signals in  $\mathbf{s}$  are transmitted independently over  $N_c$  narrowband flat sub-channels. The OFDM system thereby transforms a multipath channel to  $N_c$  flat sub-channels, each of which can be characterized by a one-tap finite impulse response filter with the coefficient  $\Lambda_{ii}$  ( $\forall i = 1, \dots, N_c$ ) and hence allows simple one-tap equalization.

## 2.2 Narrowband PLC Standards

When applied to PLC, OFDM systems need to be tailored for accommodating specific channel impairments, including fading and impulsive noise. In



this section, I discuss how various digital communications and signal processing techniques can be integrated into OFDM-based PLC systems, using existing narrowband PLC standards as examples. The discussion will shed some light on the communication performance vs. complexity trade-offs when designing PLC systems. More importantly, it also allows a better understanding of various assumptions and simulation settings for the proposed contributions.

PRIME and G3 are the well-known industry-developed OFDM-based narrowband PLC standards. Both operate in the CENELEC-A (CEN A) band from 3 to 95 kHz (ITU-T G.9955/G.9956 and IEEE P1901.2 extend G3 to the FCC band from 159.4 to 478.1 kHz), and adopt non-coherent modulation (i.e., differential phase shift keying, or DPSK) in combination with forward error correction (FEC) coding (e.g. convolutional coding (Conv) and Reed-Solomon coding (RS)) to combat adverse channel impairments. Details about both standards and the comparisons between them are well-described in [40] and [47]. PRIME and G3 PLC modems have been developed by several companies, and have demonstrated success in field trials [10, 79].

To resolve the interoperability issues among the existing technologies, unified international standards have been developed for the next generation narrowband PLC technology. The ITU-T Recommendations G.9955 and G.9956 specify the physical layer (PHY) and the data link layer, respectively, of G.hnem and document PRIME and G3 to facilitate the transition period[72]. Similarly, the very recently published IEEE P1901.2 leverages IEEE 802.15.4g and provides advanced features based on G3. Both these standards enable

scalable data rates of 200–500 kbps over a portion of the CENELEC band (3–95 kHz in CEN A, 95–125 kHz in CEN B and 125–148.5 kHz in CEN CD) and the entire US FCC band (34.375–487.5 kHz). Some PHY parameters of these two international standards are listed in Table 2.1 and compared to those of PRIME and G3. Note that since the PHY specifications of PRIME and G3 are defined in baseband real-valued settings, the FFT length is twice the number of subcarriers, whereas IEEE P1901.2 and G.hnem are specified in complex OFDM settings, where the FFT length is equal to the number of subcarriers.  $M$  denotes the constellation size. The maximum data rates are computed by taking into account of cyclic prefix (CP), FEC coding, and the overhead of frame control headers, preambles, channel estimation symbols and pilots. The two maximum data rates for PRIME in the CEN A band is with the convolutional coding turned on and off, respectively. The two maximum data rates for IEEE P1901.2 in the FCC band is with coherent and non-coherent modulation schemes, respectively.

All the narrowband PLC standards listed in Table 2.1 employ a tone map that specifies the locations of data subcarriers and the number of bits loaded on each data subcarrier. Due to complexity constraints, current standards adopt uniform tone mapping, which assigns the same number of bits to all data subcarriers. In PRIME, each transmission uses a static tone map that is decided prior to the transmission. Other standards allow adaptive tone mapping, which adjusts the number of bits uniformly across all data subcarriers based on the average SNR. Recently, observing the periodically varying and

Table 2.1: Physical layer parameters of PRIME, G3, IEEE P1901.2 and G.hnem standards for narrowband PLC over the CENELEC A and FCC bands.

Parameter		PRIME	G3	IEEE P1901.2	G.hnem
Frequency Range	CEN A	42–89 kHz	35.9–90.6 kHz	35.9–90.6 kHz	35.9–90.6 kHz
	FCC	/	159.4–478.1 kHz	35.9–487.5 kHz	34.4–478.1 kHz
Sampling Frequency	CEN A	250 kHz	400 kHz	400 kHz	200 kHz
	FCC	/	1.2 MHz	1.2 MHz	800 kHz
FFT Length	CEN A	512	256	256	128
	FCC	/	256	256	256
Cyclic Prefix (Duration)	CEN A	48 (192 $\mu$ s)	30 (75 $\mu$ s)	30 (75 $\mu$ s)	20/32 (100 $\mu$ s/160 $\mu$ s)
	FCC	/	30 (25 $\mu$ s)	30 (25 $\mu$ s)	40/64 (50 $\mu$ s/80 $\mu$ s)
Window Size	CEN A	0	8	8	8
	FCC	/	8	8	16
Effective CP (Duration)	CEN A	48 (192 $\mu$ s)	22 (55 $\mu$ s)	22 (55 $\mu$ s)	12/24 (60 $\mu$ s/120 $\mu$ s)
	FCC	/	22 (18.3 $\mu$ s)	22 (18.3 $\mu$ s)	24/48 (30 $\mu$ s/60 $\mu$ s)
Subcarrier Spacing	CEN A	488 Hz	1.5625 kHz	1.5625 kHz	1.5625 kHz
	FCC	/	4.6875 kHz	4.6875 kHz	3.125 kHz
OFDM Duration	CEN A	2240 $\mu$ s	695 $\mu$ s	695 $\mu$ s	700/760 $\mu$ s
	FCC	/	231.7 $\mu$ s	231.7 $\mu$ s	350/380 $\mu$ s
Modulation		DPSK M=2,4,8	DPSK M=2,4,8	DPSK (QAM) M=2,4,8,16	QAM M=2,4,8,16
FEC		Conv (Optional)	Conv+RS	Conv+RS	Conv+RS
Maximum Data Rate	CEN A	61.4/123 kbps	45 kbps	52.3 kbps	101.3 kbps
	FCC	/	207.6 kbps	203.2/207.6 kbps	821.1 kbps

spectrally shaped statistics of periodic impulsive noise, it was suggested to use cyclic non-uniform tone mapping to improve the throughput of narrowband PLC systems [70]. Such tone mapping varies the number of bits assigned to each data subcarrier across time and frequency based on sub-channel SNRs, and requires considerable amount of feedback from receiver to transmitter. For simplicity, I assume static uniform tone mapping as specified in PRIME throughout the dissertation. Nonetheless, the proposed transceiver methods can be generalized to work with adaptive uniform tone mapping, and even cyclic non-uniform tone mapping with minor modifications.

Comparing the modulation schemes in different narrowband PLC standards, one may notice that while DPSK is the only modulation supported

by PRIME and G3, coherent modulation schemes such as PSK and Quadrature Amplitude Modulation (QAM) are included as an optional mode in IEEE P1901.2, and a mandatory mode in G.hnem. In general, the choice between coherent modulation and non-coherent modulation reflects complexity vs. performance trade-offs. In non-coherent modulation, proper insertion of pilots ensures that the channel phase distortion between two pilots is approximately constant and thus is automatically canceled when subtracting the phases of two consecutive symbols between these pilots. As such, non-coherent modulation relieves the receiver from the computationally expensive channel estimation and carrier frequency offset correction. On the other hand, coherent systems generally require higher computational complexity, as well as training overhead, for channel estimation. For example, in the optional coherent modulation mode in IEEE P1901.2, two preamble symbols are added after the frame control header for channel estimation. Nonetheless, the SNR gain of PSK over DPSK is generally 3 dB in additive white Gaussian noise (AWGN) assuming perfect channel estimation. In non-Gaussian noise environments, the SNR gain of PSK is even larger as the noise becomes more impulsive [44]. This is because the occurrence of high amplitude noise impulses can result in noisy phase reference and severe error propagation at non-coherent receivers.

Performance of non-coherent coded modulation can be improved by increasing the receiver complexity. In [25, 78], the authors proposed a non-coherent maximum likelihood sequence estimator (NMLSE) and a multiple-symbol differential detector (MSDD), respectively, that jointly decode a block

of (more than two) differentially encoded symbols. In AWGN, by increasing the block length, performance of the non-coherent coded modulation with NMLSE or MSDD converges to that of the coherent coded modulation with the Viterbi decoder. However, such performance gain is significantly suppressed in the presence of impulsive noise [85]. In [74], it was shown that non-coherent iterative receivers using soft decision feedback from the convolutional encoder can asymptotically achieve the same communication performance as coherent receivers. A drawback of this approach is the extremely high computational complexity. A lower complexity alternative using hard decision feedback was suggested in [80], which however is likely to encounter performance degradation in impulsive noise due to error propagation.

Performance and complexity evaluation of coherent vs. non-coherent modulation schemes in narrowband PLC still remains an active research area. While non-coherent modulation allows low-complexity implementation, coherent modulation becomes increasingly attractive since it provides higher data rates, or alternatively enhanced reliability, to potentially enable more advanced smart metering applications.

In this dissertation, I consider both non-coherent and coherent modulation schemes, with an emphasis on the latter. The first two contributions are developed and evaluated assuming coherent modulation. In the first contribution, except for the decision feedback noise estimator, which requires channel estimation that is unavailable at non-coherent receivers, all the other algorithms can be immediately applied to non-coherent systems to mitigate

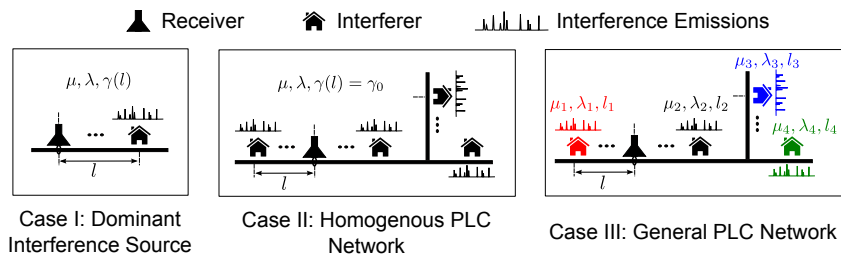


Figure 2.1: Interference scenarios in PLC networks. Each interferer emits a random sequence of emissions onto the power line, which add up at the receiver. An interferer is described statistically by a mean number of emission events  $\mu$ , mean duration between emission events  $\lambda$ , and the pathloss to the receiver  $\gamma$ .

asynchronous impulsive noise. However, in the second contribution, the proposed algorithms for mitigating periodic impulsive noise may not be used in non-coherent systems, since the time-domain block interleaving structure required by these algorithms cannot be easily incorporated into non-coherent OFDM systems. In the third contribution, the proposed transceiver methods are derived and simulated for both non-coherent and coherent systems. While offline training prior to data transmission is required in non-coherent systems, such overhead can be significantly reduced in coherent systems.

## 2.3 Statistical Modeling of Impulsive Noise

In this section, I briefly discuss existing statistical models for asynchronous impulsive noise and periodic impulsive noise in PLC. These statistical models are useful for simulating various impulsive noise environments, under which the robustness of the proposed algorithms can be tested.

Scenario	Example Network	Statistical Model
Dominant Interferer	Rural Area Industrial Area	Middleton Class A: $A = \lambda\mu$ , $\Omega = A\gamma E [h^2 B^2] / 2$
Homogeneous PLC Network	Urban Area Residential Buildings	Middleton Class A: $A = M\lambda\mu$ , $\Omega = A\gamma E [h^2 B^2] / (2M)$
General PLC Network	Dense Urban Area Commercial	Gaussian Mixture: $\pi_k$ and $\gamma_k$ given in [65]

Table 2.2: Statistical-physical models of interference in PLC networks categorized by network types. Parameters are given in Figure 2.1 and  $M$  is the number of interferers.

### 2.3.1 Asynchronous Impulsive Noise Modeling

Time-domain properties of the asynchronous impulsive noise have been empirically modeled in the literature. Many studies targeted indoor broadband PLC, and took noise measurements in higher frequency bands from several hundred kHz to 20 MHz [96], where asynchronous impulsive noise is dominant. To describe the instantaneous amplitude statistics of the noise, various studies empirically fitted the noise data to Nakagami [57], Gaussian mixture and Middleton Class A distributions [17, 24]. By characterizing random emissions of interference events in a PLC network as a temporal Poisson point process, analytical derivation in [65] 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 [65] are given in Figure 2.1 and Table 2.2.

For convenience of the discussion in the rest of the dissertation, I briefly describe the Gaussian mixture and the Middleton Class A models as follows.

- 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.,

$$f_Z(z) = \sum_{k=0}^K \pi_k \cdot \mathcal{N}(z; 0, \gamma_k), \quad (2.2)$$

where  $\mathcal{N}(z; 0, \gamma_k)$  denotes a Gaussian pdf with zero mean and variance  $\gamma_k$ , and  $\pi_k$  is the mixing probability of the  $k$ -th Gaussian component.

- Middleton Class A Model. The Middleton Class A model [59] is parameterized by an overlapping factor  $A$  and background-to-impulsive noise power ratio  $\Omega$  (e.g.  $A \in [10^{-2}, 1]$  and  $\Omega \in [10^{-6}, 1]$  [90]). The Middleton Class A model 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  $K \rightarrow \infty$ . 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 Figure 2.2 and 2.3, respectively.

### 2.3.2 Periodic Impulsive Noise Modeling

Studies in statistical modeling of periodic impulsive noise primarily targeted outdoor narrowband PLC in the 3–500 kHz band [46, 62]. In [46], 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



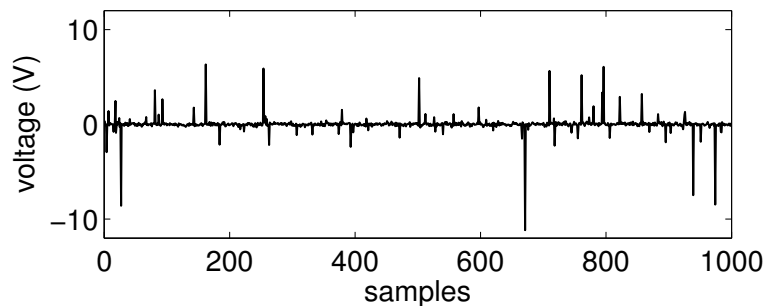


Figure 2.2: Asynchronous impulsive noise simulated from a Gaussian mixture distribution with  $\boldsymbol{\pi} = [0.9, 0.07, 0.03]$  and  $\boldsymbol{\gamma} = [1, 100, 1000]$ .

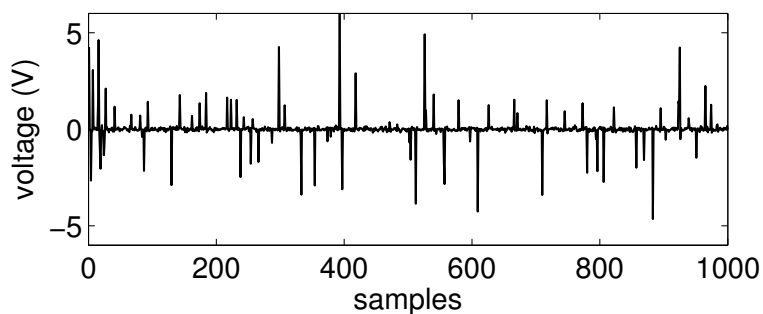


Figure 2.3: Asynchronous impulsive noise simulated from a Middleton Class A distribution with  $A = 0.1$ ,  $\Omega = 0.01$ , and the pdf truncated to the first 10 mixture components.

varying (LPTV) system model was proposed in [62] and was accepted into the IEEE P1901.2 narrowband PLC standard. The model was established on the approximation that each AC cycle can be partitioned into a number of intervals, and 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 partitioned into  $M$  intervals  $\{\mathcal{R}_i\}_{i=1}^M$ , the noise

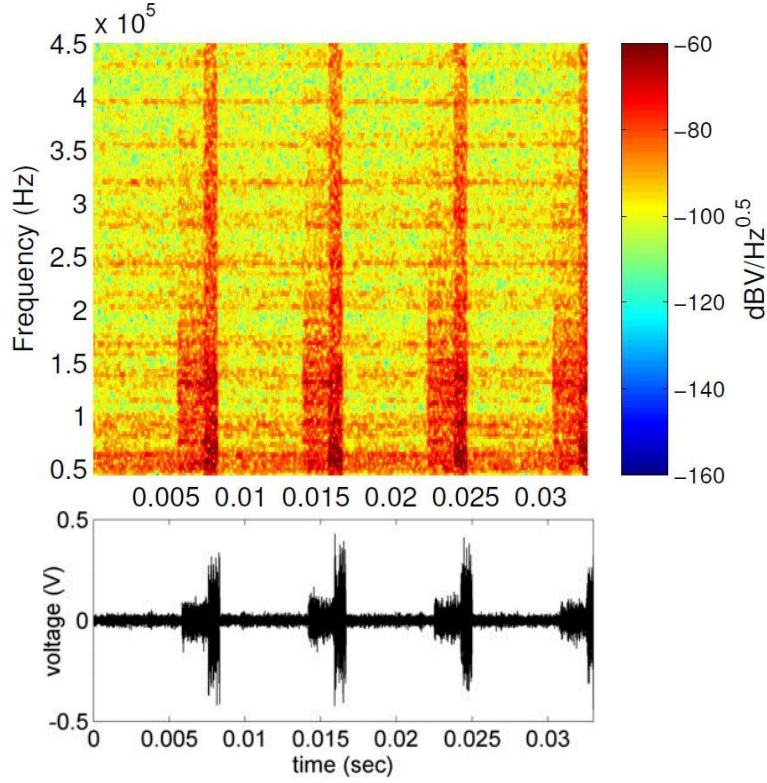


Figure 2.4: A time-domain trace and spectrogram of the periodic impulsive noise synthesized from a linear periodically time varying system model.

samples  $n_k$  can be expressed as

$$n_k = \sum_{i=1}^M \mathbf{1}_{k \in \mathcal{R}_i} \sum_{\tau} h_{\tau}^{(i)} v_{k-\tau}, \quad v_k \sim \mathcal{N}(0, 1). \quad (2.3)$$

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 Figure 2.4. 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 Figure 1.2.

Very recently, cyclostationary noise measurement in indoor broadband PLC channels in the 1–30 MHz band was fitted to a deseasonalized autoregressive moving average (DARMA) model [31]. The DARMA model consists of an LTI ARMA spectral shaping filter followed by a noise power amplifier that scales the noise power periodically. Similarly to [46], this model does not capture the spectral variation of the noise within a period.

## 2.4 The Bayesian Framework

In this section, I provide an overview of the Bayesian inference framework. The purpose is to establish some concepts and methodologies that reappear throughout this dissertation.

Informally speaking, Bayesian inference attempts to extract information about certain unknown parameters  $\boldsymbol{\theta}$  from some known data observations  $\mathbf{y}$ . In particular, one might be interested in the posterior distribution on  $\boldsymbol{\theta}$

$$p(\boldsymbol{\theta}|\mathbf{y}, \boldsymbol{\lambda}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\lambda})}{\int_{\boldsymbol{\theta}} p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\lambda})d\boldsymbol{\theta}}, \quad (2.4)$$

which is a consequence of two probability distributions, namely a prior  $p(\boldsymbol{\theta}|\boldsymbol{\lambda})$ , and a likelihood  $p(\mathbf{y}|\boldsymbol{\theta})$ . In general, the prior is a parameterized distribution with the hyperparameters  $\boldsymbol{\lambda}$ . It is used to incorporate any *a priori* information on the unknown parameters into the inference.

### 2.4.1 Conjugate Priors

There are many perspectives on the choice of prior distribution. An informative prior encodes specific subjective belief about the unknown parameters. On the other hand, a noninformative prior allows the data to speak most loudly. A pragmatic choice, which leads to computationally tractable posterior distributions, is the use of conjugate priors. A prior distribution on  $\boldsymbol{\theta}$  is called a conjugate prior for the likelihood function  $p(\mathbf{y}|\boldsymbol{\theta})$ , if the posterior remains in the same family as the prior. Depending on the choice of hyperparameters, these priors can encode different amount of information (from weakly informative to strongly informative) that influences the analysis of  $\boldsymbol{\theta}$ . In the following, I briefly outline the prior distributions associated with a complex Gaussian distribution, which will be frequently used in the rest of the dissertation.

A single variable complex Gaussian distribution is parameterized by a mean  $\mu$  and a variance  $\sigma^2$

$$p(y; \mu, \gamma) = \frac{1}{\pi\gamma} \exp\left(-\frac{|e - \mu|^2}{\gamma}\right).$$

I denote this Gaussian distribution by  $\mathcal{CN}(y; \mu, \gamma)$ . For fixed mean and unknown variance, the conjugate prior on  $\gamma$  is an inverse Gamma distribution. An inverse Gamma distribution, with shape parameter  $a$  and scale parameter  $b$ , is given by

$$p(\gamma; a, b) = \frac{b^a}{\Gamma(a)} \gamma^{-a-1} \exp\left(-\frac{b}{\gamma}\right), \quad \forall \gamma > 0 \quad (2.5)$$

and denoted by  $\text{IG}(\gamma; a, b)$ . The posterior distribution of  $\gamma$  is an updated inverse Gamma distribution

$$p(\gamma|y, a, b) = \text{IG}(\gamma; \tilde{a}, \tilde{b}), \quad (2.6)$$

where the hyperparameters are updated according to

$$\tilde{a} = a + \frac{1}{2}, \quad \tilde{b} = b + \frac{|y - \mu|^2}{2}. \quad (2.7)$$

More generally, a multivariate complex Gaussian distribution on an  $N \times 1$  vector  $\mathbf{y}$  is parameterized by a mean vector  $\boldsymbol{\mu}$  and a covariance matrix  $\boldsymbol{\Sigma}$

$$p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\pi^N |\boldsymbol{\Sigma}|} \exp[-(\mathbf{y} - \boldsymbol{\mu})^* \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})].$$

Similarly, we denote this distribution by  $\mathcal{CN}(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ . If the covariance is the only uncertain parameter, the conjugate prior on  $\boldsymbol{\Sigma}$  is an inverse Wishart distribution. An inverse Wishart distribution on the positive definite matrix  $\boldsymbol{\Sigma}$  is

$$p(\boldsymbol{\Sigma}; d, \boldsymbol{\Psi}) = \frac{|\boldsymbol{\Psi}|^{d/2}}{2^{\frac{dN}{2}} \Gamma_N(\frac{d}{2})} |\boldsymbol{\Sigma}|^{-\frac{d+N+1}{2}} \exp[-\frac{1}{2} \text{Tr}(\boldsymbol{\Psi} \boldsymbol{\Sigma}^{-1})], \quad (2.8)$$

with the degree of freedom  $d \geq N$  and covariance parameter  $\boldsymbol{\Psi}$ . We denote this distribution by  $\text{IW}(\boldsymbol{\Sigma}; d, \boldsymbol{\Psi})$ . The posterior distribution of  $\boldsymbol{\Sigma}$  is an updated inverse Wishart distribution given by

$$p(\boldsymbol{\Sigma}|\mathbf{y}, \mu, \boldsymbol{\Psi}) = \text{IW}(\boldsymbol{\Sigma}; \tilde{d}, \tilde{\boldsymbol{\Psi}}), \quad (2.9)$$

where the hyperparameters are updated by

$$\tilde{d} = d + 1, \quad \tilde{\boldsymbol{\Psi}} = \boldsymbol{\Psi} + (\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^* \quad (2.10)$$

### 2.4.2 Expectation Maximization Algorithm

One may find the maximum *a posteriori* (MAP) estimate of  $\boldsymbol{\theta}$  by maximizing the posterior distribution (2.4) over  $\boldsymbol{\theta}$ . This can be solved by taking the derivatives of the posterior with respect to  $\boldsymbol{\theta}$ , setting them to zero and solving the resulting equations. In many cases, it is difficult or even impossible to find a closed-form solution to the equations. The expectation maximization (EM) algorithm provides an alternative, simpler way to solve the MAP problem in these cases.

The EM algorithm first introduces a set of latent (hidden) variables  $\mathbf{x}$ , and rewrites the likelihood function as

$$p(\mathbf{y}|\boldsymbol{\theta}) = \sum_{\mathbf{x}} p(\mathbf{y}, \mathbf{x}|\boldsymbol{\theta}).$$

In order to find the MAP solution of the posterior

$$\operatorname{argmax}_{\boldsymbol{\theta}} \sum_{\mathbf{x}} p(\mathbf{y}, \mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\lambda}),$$

the algorithm iteratively apply two steps:

- Expectation step (E-step): compute  $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = \mathbb{E}_{\mathbf{x}|\mathbf{y},\boldsymbol{\theta}^{(k)}} [\log p(\mathbf{y}, \mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\lambda})]$
- Maximization step (M-step): solve  $\boldsymbol{\theta}^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ .

The convergence of the EM algorithm requires the unknown parameters  $\boldsymbol{\theta}$  to be continuous valued. In this case, starting with certain initial values of  $\boldsymbol{\theta}^{(0)}$ , the algorithm monotonically converges to a local maximum of the posterior, which is exactly the MAP solution if the posterior is a concave function.

## 2.5 Sparse Bayesian Learning

Having established the Bayesian inference framework, I will discuss how the Bayesian approach can be used for the recovery of a sparse signal from its linear observations, *a.k.a.* compressed sensing.

The basic mathematical model for compressed sensing is

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{v}, \quad (2.11)$$

where  $\mathbf{y}$  is an  $M \times 1$  observation vector,  $\mathbf{v}$  is an unknown noise vector,  $\Phi \triangleq [\Phi_1 \ \cdots \ \Phi_N]$  is an  $M \times N$  known dictionary matrix ( $M < N$ ), and is assumed to be full rank. The task is to estimate the source vector  $\mathbf{x}$ . Although it is generally an ill-conditioned problem, there exists a unique global solution if the number of nonzero entries in  $\mathbf{x}$  is less than a threshold.

This basic single measurement vector (SMV) model can be extended to the multiple measurement vector (MMV) model

$$\mathbf{Y} = \Phi \mathbf{X} + \mathbf{V}, \quad (2.12)$$

where  $\mathbf{Y} \triangleq [\mathbf{y}_1 \ \cdots \ \mathbf{y}_L]$  is a measurement matrix consisting of  $L$  measurement vectors,  $\mathbf{X} \triangleq [\mathbf{x}_1 \ \cdots \ \mathbf{x}_L]$  is an unknown source matrix with each row representing a possible source, and  $\mathbf{V}$  is an unknown noise matrix. In addition, the MMV model assumes that the support of every column in  $\mathbf{X}$  is identical. Similar to the constraint in the SMV model, the number of nonzero rows in  $\mathbf{X}$  has to be below a threshold to ensure a unique global solution.

Among the compressed sensing algorithms, sparse Bayesian learning (SBL) has received much attention. SBL was initially proposed by Tipping [83]. It was introduced to sparse signal recovery for the SMV model by Wipf and Rao [87], and was later extended to the MMV model in [88] and [91]. The key idea of SBL is to impose a sparsity promoting prior on the source vector  $\mathbf{x}$  or the source matrix  $\mathbf{X}$ , which leads to a posterior density that is concentrated over sparse vectors or row-sparse matrices.

Compared to traditional compressed sensing algorithms such as Basis Pursuit [19] and FOCUSS [33], SBL has several unique advantages [87], including (1) the global optimum is always the sparsest solution; (2) all local optimal solutions are sparse; and (3) the number of local optima is the smallest. All these contribute to the excellent recovery performance of the SBL algorithm.

### 2.5.1 Single Measurement Vector

For the SMV model, SBL imposes a hierarchical parameterized Gaussian prior on  $\mathbf{x}$

$$p(\mathbf{x}|\boldsymbol{\gamma}) = \mathcal{CN}(\mathbf{x}; \mathbf{0}, \boldsymbol{\Gamma}) \quad (2.13)$$

$$p(\boldsymbol{\gamma}|\mathbf{a}, \mathbf{b}) = \prod_{i=1}^N \text{IG}(\gamma_i; a_i, b_i), \quad (2.14)$$

where  $\boldsymbol{\Gamma} \triangleq \text{diag}(\boldsymbol{\gamma})$ , and  $\mathbf{a}, \mathbf{b}$  are given hyperparameters. From the hierarchical formulation of the prior, it appears that a non-sparse Gaussian prior is used. To discover the true identity of this prior, one may integrate out  $\boldsymbol{\gamma}$  to obtain



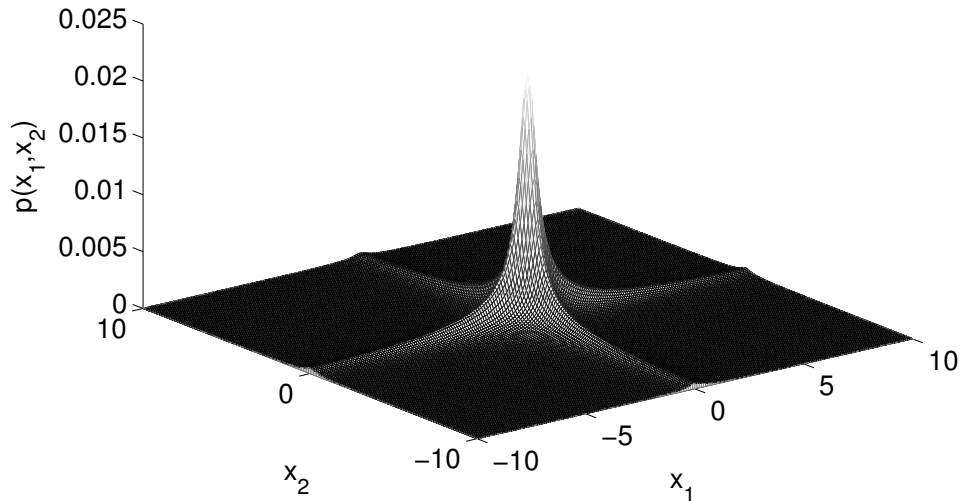


Figure 2.5: The density of the prior  $p(\mathbf{x}; \mathbf{a}, \mathbf{b})$ , where  $\mathbf{x}$  is a length-2 vector, and the parameters  $\boldsymbol{\gamma}$  have been integrated out.

the marginal distribution of  $\mathbf{x}$

$$\begin{aligned}
 p(\mathbf{x}; \mathbf{a}, \mathbf{b}) &= \prod_{i=1}^N \left[ \int p(x_i | \gamma_i) p(\gamma_i | a_i, b_i) d\gamma_i \right] \\
 &= \prod_{i=1}^N \frac{b_i^{a_i} \Gamma(a_i + \frac{1}{2})}{(2\pi)^{\frac{1}{2}} \Gamma(a_i)} \left( b_i + \frac{|x_i|^2}{2} \right)^{-(a_i + \frac{1}{2})} \quad (2.15)
 \end{aligned}$$

Each term in (2.15) corresponds to a Student- $t$  distribution. That is, the prior imposes an independent Student- $t$  distribution on each dimension of  $\mathbf{x}$ . A visualization of the prior (2.15) for  $N = 2$  (Figure 2.5) clearly shows its sparsity encouraging property. We observe that the probability mass is concentrated both at the origin, i.e.,  $\mathbf{x} = 0$ , and along the two axes, where one of the two elements of  $\mathbf{x}$  is zero. When  $a_i = b_i = 0$ ,  $\forall i$ , the inverse Gamma distribution

reduces to a uniform distribution over the positive support. With the uniform hyperprior, the marginalized prior becomes  $p(\mathbf{x}; \mathbf{a}, \mathbf{b}) \propto \prod_{i=1}^N 1/|x_i|$ , which is also a sparse prior since it is sharply peaked at zero.

Assuming that the noise  $\mathbf{v}$  is AWGN with zero mean and unknown variance, i.e.,  $\mathbf{v} \sim \mathcal{CN}(0, \sigma^2)$ . The likelihood of the observations  $\mathbf{y}$  given the unknown parameters  $\boldsymbol{\gamma}$  and  $\sigma$  is Gaussian

$$p(\mathbf{y}|\boldsymbol{\gamma}, \sigma) = \mathcal{CN}(\mathbf{y}; \mathbf{0}, \boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^* + \sigma^2\mathbf{I}_M), \quad (2.16)$$

where  $\mathbf{I}_M$  denotes the  $M \times M$  identity matrix. The posterior distribution of  $\boldsymbol{\gamma}$  and  $\sigma$  is therefore

$$p(\boldsymbol{\gamma}, \sigma | \mathbf{y}, \mathbf{a}, \mathbf{b}) \propto \mathcal{CN}(\mathbf{y}; \mathbf{0}, \boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^* + \sigma^2\mathbf{I}_M) \prod_{i=1}^N \text{IG}(\gamma_i; a_i, b_i). \quad (2.17)$$

To find the MAP estimate of  $\boldsymbol{\gamma}$  and  $\sigma$ , one may employ the EM algorithm, which treats  $\mathbf{x}$  as latent variable. Details about the derivation of the E-step and M-step can be found in [83]. The  $k$ -th iteration of the EM algorithm is summarized as:

$$\begin{aligned} \gamma_i^{(k+1)} &= \underset{\gamma_i \geq 0}{\operatorname{argmax}} \mathbb{E}_{\mathbf{x}|\mathbf{y}; \boldsymbol{\gamma}^{(k)}, \sigma^{(k)}} [\log p(\mathbf{y}, \mathbf{x}|\boldsymbol{\gamma}, \sigma^{(k)})p(\boldsymbol{\gamma}|\mathbf{a}, \mathbf{b})] \\ &= \underset{\gamma_i \geq 0}{\operatorname{argmax}} \mathbb{E}_{\mathbf{x}|\mathbf{y}; \boldsymbol{\gamma}^{(k)}, \sigma^{(k)}} [\log p(\mathbf{y}|\mathbf{x}, \sigma^{(k)})p(\mathbf{x}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}|\mathbf{a}, \mathbf{b})] \\ &= \frac{\boldsymbol{\Sigma}_{\mathbf{x}, ii}^{(k)} + (\boldsymbol{\mu}_{\mathbf{x}, i}^{(k)})^2 + 2b_i}{1 + 2a_i}, \end{aligned} \quad (2.18)$$

$$\begin{aligned} (\sigma^{(k+1)})^2 &= \underset{\sigma \geq 0}{\operatorname{argmax}} \mathbb{E}_{\mathbf{x}|\mathbf{y}; \boldsymbol{\gamma}^{(k)}, \sigma^{(k)}} [\log p(\mathbf{y}|\mathbf{x}, \sigma)] \\ &= \frac{1}{M} \{ \|\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{x}}^{(k)}\|^2 + (\sigma^{(k)})^2 \sum_{i=1}^N [1 - (\gamma_i^{(k)})^{-1} \boldsymbol{\Sigma}_{\mathbf{x}, ii}^{(k)}] \}. \end{aligned} \quad (2.19)$$

Note that the parameters are updated one at a time while keeping the others fixed. Here  $\boldsymbol{\mu}_{\mathbf{x}}$  and  $\boldsymbol{\Sigma}_{\mathbf{x}}$  are the posterior mean and covariance of  $\mathbf{x}$ , given  $\mathbf{y}$  and the current estimates of  $\boldsymbol{\gamma}$  and  $\sigma$ . It can be shown that the posterior density of  $\mathbf{x}$  is also a Gaussian distribution [83]

$$\begin{aligned}
p(\mathbf{x}|\mathbf{y}; \boldsymbol{\gamma}^{(k)}, \sigma^{(k)}) &= \mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x}}^{(k)}, \boldsymbol{\Sigma}_{\mathbf{x}}^{(k)}), \\
\boldsymbol{\mu}_{\mathbf{x}}^{(k)} &= (\sigma^{(k)})^{-2} \boldsymbol{\Sigma}_{\mathbf{x}}^{(k)} \boldsymbol{\Phi}^* \mathbf{y}, \\
\boldsymbol{\Sigma}_{\mathbf{x}}^{(k)} &= \boldsymbol{\Gamma}^{(k)} - \boldsymbol{\Gamma}^{(k)} \boldsymbol{\Phi}^* \left( (\sigma^{(k)})^2 \mathbf{I}_M + \boldsymbol{\Phi} \boldsymbol{\Gamma}^{(k)} \boldsymbol{\Phi}^* \right)^{-1} \boldsymbol{\Phi} \boldsymbol{\Gamma}^{(k)}.
\end{aligned}
\tag{2.20}$$

Upon convergence, most components of  $\boldsymbol{\gamma}$  are driven to zero. A point estimate of  $\mathbf{x}$  can be given by the posterior mean  $\boldsymbol{\mu}_{\mathbf{x}}$ . One can see from (2.20) and (2.21) that as most components of  $\boldsymbol{\gamma}$  go to zero,  $\boldsymbol{\mu}_{\mathbf{x}}$  becomes a sparse vector as well.

### 2.5.2 Sequential Sparse Bayesian Learning

The core SBL algorithm in (2.21) involves an  $M \times M$  matrix inversion, which has a computational complexity of  $\mathcal{O}(M^3)$  and hence might be practically infeasible on a hardware platform if  $M$  is large. An accelerated version of SBL that utilizes the properties of the marginal likelihood has been proposed in [84]. Using these properties, the accelerated SBL algorithm performs a sequential addition and deletion of candidate basis functions given by the columns of  $\boldsymbol{\Phi}$  in (2.11), while keeping the same reconstruction performance.

A non-zero entry in vector  $\mathbf{x}$  at index  $i$  contributes  $x_i \boldsymbol{\phi}_i$  to the measurement vector  $\mathbf{y}$ , where  $\boldsymbol{\phi}_i$  denotes the  $i$ -th column in  $\boldsymbol{\Phi}$ . The accelerated

algorithm will sequentially add, remove, or update basis  $\phi_i$  until convergence. At convergence, the bases that remains in the model will indicate the support of vector  $\mathbf{x}$ . Given that  $\mathbf{x}$  is sparse, the number of bases that remains in each iteration is generally smaller than  $M$ . As such, the sequential SBL algorithm reduces the computational complexity of the matrix inversion operation in each iteration from  $\mathcal{O}(M^3)$  to  $\mathcal{O}(N_B^3)$ , where  $N_B$  is the number of remaining bases and varies among different iterations. Furthermore, significant computational savings can be obtained if the background noise power  $\sigma^2$  is known. This will allow for efficient calculations without any matrix inversions as given in [84]. Please refer to [84] for more mathematical details.

### 2.5.3 Multiple Measurement Vector

To solve the MMV model, the SBL algorithm was extended to the temporally correlated MMV sparse Bayesian learning (T-MSBL) algorithm in [91].

Let  $\mathbf{X}_i$  denote the  $i$ -th row of  $\mathbf{X}$ . The algorithm imposes a parameterized Gaussian prior on  $\mathbf{X}_i$ .

$$p(\mathbf{X}_i; \gamma_i, \mathbf{B}) = \mathcal{CN}(\mathbf{X}_i; 0, \gamma_i \mathbf{B}), \quad \forall i = 1, \dots, N \quad (2.22)$$

where  $\boldsymbol{\gamma} \triangleq [\gamma_1 \cdots \gamma_N]$  are nonnegative parameters controlling the row sparsity of  $\mathbf{X}$ , and  $\mathbf{B}$  is a positive definite matrix that captures the covariance of  $\mathbf{X}_i$ . In addition, hyperpriors on  $\boldsymbol{\gamma}$  can be imposed to form a hierarchical prior as (2.13) in the SMV model. However, for simplicity, one could simply assume a uniform hyperprior, i.e.,  $\gamma_i \sim \text{IG}(\gamma_i; 0, 0)$ , and adopt (2.22) as the prior.

The MMV model (2.12) can be transformed to a block sparsity SMV model by vectorizing the matrices  $\mathbf{Y}$ ,  $\mathbf{X}$  and  $\mathbf{V}$ . Defining  $\bar{\mathbf{y}} \triangleq \text{vec}(\mathbf{Y}^T)$ ,  $\bar{\mathbf{x}} \triangleq \text{vec}(\mathbf{X}^T)$ ,  $\bar{\mathbf{v}} \triangleq \text{vec}(\mathbf{V}^T)$ , and  $\mathbf{D} \triangleq \Phi \otimes \mathbf{I}_L$ , where  $L$  is the number of columns in  $\mathbf{X}$ , we have

$$\bar{\mathbf{y}} = \mathbf{D}\bar{\mathbf{x}} + \bar{\mathbf{v}}. \quad (2.23)$$

Since  $\bar{\mathbf{x}}$  is constructed by concatenating the rows of  $\mathbf{X}$  into a vector, row sparsity in  $\mathbf{X}$  indicates block sparsity in  $\bar{\mathbf{x}}$  with length- $L$  blocks. The prior (2.22) can now be rewritten succinctly as

$$p(\bar{\mathbf{x}}; \boldsymbol{\gamma}, \mathbf{B}) = \mathcal{CN}(\bar{\mathbf{x}}; \mathbf{0}, \boldsymbol{\Sigma}_0), \quad (2.24)$$

with  $\boldsymbol{\Sigma}_0 \triangleq \text{diag}\{\boldsymbol{\gamma}\} \otimes \mathbf{B}$ .

Suppose that the unknown noise  $\mathbf{v}$  can be modeled by AWGN with zero mean and variance  $\lambda$ . The posterior distribution of the unknown parameters  $\boldsymbol{\theta} \triangleq \{\boldsymbol{\gamma}, \mathbf{B}, \lambda\}$  is proportional to the likelihood

$$p(\boldsymbol{\theta}; \bar{\mathbf{y}}) \propto p(\bar{\mathbf{y}}; \boldsymbol{\theta}) = \mathcal{CN}(\bar{\mathbf{y}}; \mathbf{0}, \mathbf{D}\boldsymbol{\Sigma}_0\mathbf{D}^* + \lambda\mathbf{I}). \quad (2.25)$$

The MAP solution of  $\boldsymbol{\theta}$  can then be computed using the EM algorithm, by treating  $\bar{\mathbf{x}}$  as latent variables. The update rules in each iteration of the EM algorithm were derived in [91] and are summarized as follows.

$$\begin{aligned} \gamma_i^{(k+1)} &= \frac{1}{L} \text{Tr}[(\mathbf{B}^{(k)})^{-1}(\boldsymbol{\Sigma}_{\bar{\mathbf{x}},i}^{(k)} + \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^{(k)} \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^{(k)*})], \forall i = 1, \dots, N \\ \mathbf{B}^{(k+1)} &= \frac{1}{N} \sum_{i=1}^N \frac{\boldsymbol{\Sigma}_{\bar{\mathbf{x}},i}^{(k)} + \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^{(k)} \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^{(k)*}}{\gamma_i}, \\ (\sigma^{(k+1)})^2 &= \frac{\|\bar{\mathbf{y}} - \mathbf{D}\boldsymbol{\mu}_{\bar{\mathbf{x}}}^{(k)}\|_2^2 + (\sigma^{(k)})^2 [NL - \text{Tr}(\boldsymbol{\Sigma}_{\bar{\mathbf{x}}}^{(k)}(\boldsymbol{\Sigma}_0^{(k)})^{-1})]}{ML}. \end{aligned} \quad (2.26)$$

Here  $\Sigma_{\bar{\mathbf{x}},i}$  denotes the  $i$ -th  $L \times L$  diagonal block of  $\Sigma_{\bar{\mathbf{x}}}$ , and similarly  $\boldsymbol{\mu}_{\bar{\mathbf{x}},i}$  is the  $i$ -th length- $L$  block of  $\boldsymbol{\mu}_{\bar{\mathbf{x}}}$ . The posterior mean and covariance of  $\bar{\mathbf{x}}$  given the current estimate of the parameters can be expressed as

$$\begin{aligned}\boldsymbol{\mu}_{\bar{\mathbf{x}}}^{(k)} &= \Sigma_0^{(k)} \mathbf{D}^* \left( (\sigma^{(k)})^2 \mathbf{I} + \mathbf{D} \Sigma_0^{(k)} \mathbf{D}^* \right)^{-1} \bar{\mathbf{y}}, \\ \Sigma_{\bar{\mathbf{x}}}^{(k)} &= \Sigma_0^{(k)} - \Sigma_0^{(k)} \mathbf{D}^* \left( (\sigma^{(k)})^2 \mathbf{I} + \mathbf{D} \Sigma_0^{(k)} \mathbf{D}^* \right)^{-1} \mathbf{D} \Sigma_0^{(k)}.\end{aligned}\quad (2.27)$$

The T-SBL algorithm [91] comprised of (2.26) and (2.27) transforms the original MMV model into a higher dimensional SMV model, and learns the parameters in the higher dimensional space. This results in high computational complexity due to the high dimensional matrix/vector operations. In light of this, a low-complexity near-optimal algorithm, namely T-MSBL, was proposed to back-map the T-SBL algorithm to the original lower dimensional space [91]. The EM update rules in the  $k$ -th iteration of T-MSBL are summarized in Algorithm 1, where  $\boldsymbol{\Gamma} \triangleq \text{diag}\{\boldsymbol{\gamma}\}$ , and  $\|\cdot\|_{\mathcal{F}}$  denotes the Frobenius norm of a matrix. Upon convergence, most components of  $\boldsymbol{\gamma}$  are driven to zero, thereby rendering a block sparse estimate of  $\mathbf{x}$ , or equivalently a row sparse estimate of  $\mathbf{X}$ .

## 2.6 Conclusion

In this chapter, I give an extensive review of the basic OFDM system model, and several digital communications and signal processing techniques that are adopted in OFDM-based PLC systems. After a brief description of the statistical models for impulsive noise, I proceed to discuss the general Bayesian

---

**Algorithm 1** T-MSBL [91]

---

- 1: Initialize parameters  $\boldsymbol{\gamma}^{(0)} = \mathbf{1}$ ,  $\mathbf{B}^{(0)} = \mathbf{I}_L$ ,  $\lambda^{(0)} = 0$ .
- 2: **for**  $k = 1, \dots, K$  **do**
- 3:     E-Step:

$$\begin{aligned}\boldsymbol{\Xi}^{(k)} &= \left( (\boldsymbol{\Gamma}^{(k)})^{-1} + \frac{1}{\lambda^{(k)}} \boldsymbol{\Phi}^* \boldsymbol{\Phi} \right)^{-1}, \\ \hat{\mathbf{X}}^{(k)} &= \boldsymbol{\Gamma}^{(k)} \boldsymbol{\Phi}^* (\lambda^{(k)} \mathbf{I}_L + \boldsymbol{\Phi} \boldsymbol{\Gamma}^{(k)} \boldsymbol{\Phi}^*)^{-1} \mathbf{Y}.\end{aligned}\quad (2.28)$$

- 4:     M-Step:

$$\begin{aligned}\gamma_i^{(k+1)} &= \frac{1}{L} \hat{\mathbf{X}}_{i \cdot}^{(k)} (\mathbf{B}^{(k)})^{-1} \hat{\mathbf{X}}_{i \cdot}^{(k)*} + \boldsymbol{\Xi}_{ii}^{(k)}, \forall i = 1, \dots, N \\ \tilde{\mathbf{B}}^{(k+1)} &= \sum_{i=1}^N \frac{\hat{\mathbf{X}}_{i \cdot}^{(k)*} \hat{\mathbf{X}}_{i \cdot}^{(k)}}{\gamma_i^{(k)}}, \\ \mathbf{B}^{(k+1)} &= \tilde{\mathbf{B}}^{(k)} / \|\tilde{\mathbf{B}}^{(k)}\|_{\mathcal{F}} \\ \lambda^{(k+1)} &= \frac{1}{ML} \|\mathbf{Y} - \boldsymbol{\Phi} \hat{\mathbf{X}}^{(k)}\|_{\mathcal{F}}^2 + \\ &\quad \frac{\lambda^{(k)}}{M} \text{Tr}[\boldsymbol{\Phi} \boldsymbol{\Gamma}^{(k)} \boldsymbol{\Phi}^* (\lambda^{(k)} \mathbf{I}_M + \boldsymbol{\Phi} \boldsymbol{\Gamma}^{(k)} \boldsymbol{\Phi}^*)^{-1}].\end{aligned}\quad (2.29)$$

- 5: **end for**
  - 6: Return  $\boldsymbol{\gamma}^{(K)}$ ,  $\hat{\mathbf{X}}^{(K)}$ .
- 

inference framework, and in particular the sparse Bayesian learning algorithms. All of these establish the fundamental system model and methodologies upon which the contributions of the dissertation will be built.

## Chapter 3

### Nonparametric Mitigation of Impulsive Noise

Asynchronous impulsive noise and periodic impulsive noises limit communication performance in OFDM-based PLC 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, I exploit its sparsity in the time domain, and apply sparse Bayesian learning methods introduced in Section 2.5.1 to estimate and subtract out the noise impulses. I propose three iterative algorithms, with different complexity vs. performance trade-offs, which (1) utilize the noise projection onto null and pilot subcarriers; (2) add the information to (1) in the data subcarriers to perform joint noise estimation and symbol detection; and (3) 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 6.8 dB in coded systems in periodic impulsive noise.

### 3.1 Introduction

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 [52], minimum mean square error (MMSE) equalizer [89] or decoder [38, 39] to compensate the performance loss due to non-Gaussianity of the noise. Such approaches, however, entail significant training overhead prior to data transmission for model parameter estimation, and may be vulnerable to parameter estimation errors.

In this part of the dissertation, I aim to mitigate asynchronous impulsive noise and periodic impulsive noise, respectively, at OFDM-based PLC receivers. My work distinguishes from the above approaches in two perspectives: (1) I propose “nonparametric” algorithms that do not make any assumptions on statistical noise models and hence do not require extra training; and (2) my 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, I develop three denoising algorithms based on SBL techniques that were proposed in [87] and reviewed in Section 2.5.1. I 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, I 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, I rely on a time-domain block interleaving OFDM transceiver structure (as previously proposed in [9]), 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 this chapter is organized as follows. In Section 3.2, I review existing receiver algorithms for impulsive noise mitigation. Having established the system model in Section 3.3, I propose three nonparametric impulsive noise mitigation algorithms in Section 3.4. Then I perform complexity analysis and present a low-complexity implementation of the first proposed algorithm in Section 3.5. To demonstrate the performance of the proposed algorithms, simulation results are presented and discussed in Section 3.6.

## **3.2 Prior Work**

Prior work on combating asynchronous impulsive noise and periodic impulsive noise involves efforts from both transmitter and receiver's perspectives. In Section 2.3, I have described various statistical models to capture

temporal and spectral properties of asynchronous impulsive noise and periodic impulsive noise, respectively. These models have been exploited to derive parametric receiver methods based on the estimated model parameters. Alternatively, nonparametric receiver methods that do not rely on any assumptions on the noise models have also been investigated in the literature.

### 3.2.1 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 [35]. 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 [64, 66], nulling and clipping methods [93], MMSE symbol-by-symbol detectors [38], and iterative decoders [39, 63]. 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/or can suffer from performance degradation when the noise model or parameters mismatch the possibly time-varying noise statistics. In [68], low-complexity message passing OFDM receivers were proposed that leverage recent results in approximate message passing [77] and sparse Bayesian learning [84] for joint channel / interference estimation and data decoding. While achieving huge improvements in communication performance without training overhead, the approach in [68] assumes Gaussian

mixture modeled or hidden Markov modeled impulsive noise, and might expect performance degradation if actual noise statistics does not follow either of these models.

Recently, there has been growing interest in developing nonparametric denoising methods that exploit the sparse structure of the asynchronous impulsive noise in the time domain. In particular, [16] applied Basis Pursuit [19] based compressed sensing 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 FFT 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 approach was extended in [50] 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.

My work seeks to develop nonparametric mitigation algorithms that are applicable to all asynchronous impulsive noise scenarios. Towards this end, I extend the Basis Pursuit based algorithm in [16] to a sparse Bayesian learning (SBL) approach [87] for improved performance and robustness.

### 3.2.2 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 [14, 30], 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 [20, 89], 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 [29, 51] or whiten [52] the periodic impulsive noise at the NB PLC receivers. In particular, in [52], the noise was fitted to a periodically switching autoregressive (AR) process by nonparametric Bayesian learning. Based on the estimated AR model, a peri-

odically switching moving average filter was adopted at the receiver for noise whitening. A common drawback of these filter-based methods [29, 51, 52] 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 [9, 55, 60]. Since periodic impulsive noise occurs in bursts that typically span more than one OFDM symbol, 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 [55] and time-domain interleaving [9, 60]. In particular, a time-domain block interleaving OFDM (TDI-OFDM) transceiver structure was proposed in [9] to cope with bursty impulsive noise. Unlike conventional frequency-domain interleaving OFDM (FDI-OFDM) systems where the interleaver is placed before the IFFT at the transmitter, the TDI-OFDM scheme interleaves and deinterleaves the signal in the time domain, i.e., post-IFFT at the transmitter and pre-FFT 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 [9] 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 [67], in which TDI-OFDM generally has diminishing gains over FDI-OFDM and

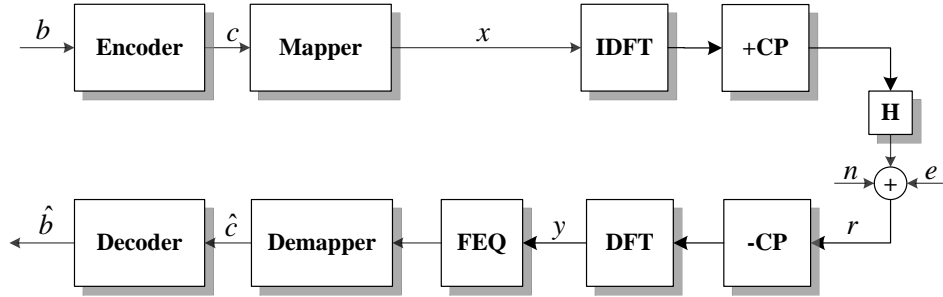


Figure 3.1: A conventional baseband coded OFDM system, where +CP (-CP) means inserting (removing) the cyclic prefix to (from) each OFDM symbol, and FEQ stands for frequency-domain equalization.

even performs worse towards lower SNRs.

Although the TDI-OFDM scheme itself does not provide much benefit in NB PLC, I will show that by embedding the proposed SBL-based denoising algorithms into the TDI-OFDM framework, significant BER improvement over conventional FDI-OFDM systems can be achieved even at low 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 I have developed for asynchronous impulsive noise mitigation.

### 3.3 System Model

For asynchronous impulsive noise mitigation in BB PLC, I consider a conventional OFDM system whose complex baseband equivalent representation is shown in Fig. 3.1. At the transmitter, a convolutionally coded data stream is mapped to a set of PSK or QAM symbols. Suppose that each

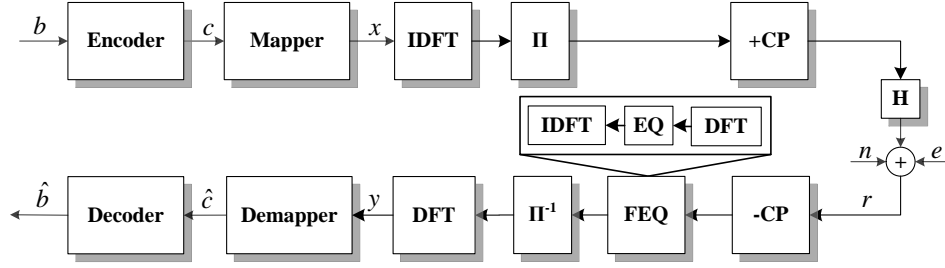


Figure 3.2: A time-domain interleaved OFDM system.  $\Pi$  denotes the sample-level interleaver, and  $\Pi^{-1}$  the corresponding deinterleaver.

OFDM symbol has  $N_c$  subcarriers,  $N_0$  of which are non-data subcarriers (i.e., null tones and pilot tones). The PSK/QAM symbols are then modulated onto the  $N_c - N_0$  data subcarriers of a number of OFDM symbols. As interpreted in Section 2.1, if the cyclic prefix (CP) is longer than the channel delay spread, a received and demodulated OFDM symbol can be expressed as

$$\mathbf{r} = \mathbf{\Lambda} \mathbf{s} + \mathbf{F} \mathbf{e} + \mathbf{g}. \quad (3.1)$$

Here I adopt the same notations as defined in Section 2.1, where  $\mathbf{s}$  denotes the transmitted OFDM symbol in the frequency domain,  $\mathbf{e}$  and  $\mathbf{g}$  represent the time-domain impulsive noise and the AWGN modeled background noise in the frequency domain, respectively,  $\mathbf{F}$  is the  $N_c$ -point FFT matrix, and  $\mathbf{\Lambda}$  is a diagonal matrix, with  $\{\Lambda_{ii}\}_{i=1}^{N_c}$  equal to the  $N_c$ -point FFT coefficients of the channel impulse response.

Let  $\mathcal{J}$  denote the index set of the non-data subcarriers, and define  $(\cdot)_{\mathcal{J}}$  as the sub-matrix (or sub-vector) corresponding to the rows (or elements) indexed by the set  $\mathcal{J}$ . The non-data subcarriers of the transmitted OFDM symbol,



written as  $\mathbf{s}_j$ , are known in advance at the receiver since they are either zeros or predefined pilots. Assuming perfect channel estimation, subtracting  $(\mathbf{\Lambda}\mathbf{s})_j$  from  $\mathbf{r}_j$  results in linear observations of the impulsive noise

$$\mathbf{y} \triangleq \mathbf{r}_j - (\mathbf{\Lambda}\mathbf{s})_j = \mathbf{F}_j\mathbf{e} + \mathbf{g}_j. \quad (3.2)$$

Recognizing the sparse nature of  $\mathbf{e}$  (since impulsive noise has very few non-zero samples in the time domain), the recovery of  $\mathbf{e}$  from  $\mathbf{y}$  based on the under-determined  $N_0 \times N_c$  linear system in (3.2) is the standard SMV compressed sensing problem as in (2.11).

I would like to use the estimated impulsive noise to improve the detection of  $\mathbf{s}$ . 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

$$\begin{aligned} \hat{\mathbf{r}}_{\bar{j}} &= \mathbf{r}_{\bar{j}} - \mathbf{F}_{\bar{j}}\hat{\mathbf{e}} \\ &= (\mathbf{\Lambda}\mathbf{s})_{\bar{j}} + \mathbf{g}_{\bar{j}} + \mathbf{F}_{\bar{j}}(\mathbf{e} - \hat{\mathbf{e}}). \end{aligned} \quad (3.3)$$

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

For periodic impulsive noise mitigation in NB PLC, I consider a TDI-OFDM system [9] as shown in Fig. 3.2. At the transmitter, multiple OFDM symbols are interleaved using a sample-level block interleaver after the IFFT 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. Similar 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 FFT. Assuming perfect channel estimation, the demodulated OFDM signal  $\mathbf{y}$  can be expressed as

$$\mathbf{r} = \mathbf{s} + \mathbf{F}\mathbf{e}_\pi + \mathbf{F}\mathbf{n}_\pi = \mathbf{s} + \mathbf{F}\mathbf{e}_\pi + \mathbf{g}_\pi. \quad (3.4)$$

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

One may notice that (3.4) can be considered as a special case of (2.1) with  $\mathbf{\Lambda} = \mathbf{I}$ . Furthermore, with appropriate selection of the interleaver size, the deinterleaved noise  $\mathbf{e}_\pi$  assumes similar sparsity structure as the asynchronous impulsive noise. Therefore, all the algorithms derived from (3.1) can also be directly applied to the periodic impulsive noise case, simply by replacing  $\mathbf{\Lambda}$  by  $\mathbf{I}$ , and  $\mathbf{e}, \mathbf{g}$  by  $\mathbf{e}_\pi, \mathbf{g}_\pi$ , respectively.

Given a particular noise scenario, the size of the interleaver is an important design factor that determines the sparsity level of  $\mathbf{e}_\pi$ , which affects the recovery performance of the compressed sensing algorithms. The sparsity level of a vector is defined as the number of non-zero elements in the vector. The key is to maintain the sparsity level of  $\mathbf{e}_\pi$  below a certain threshold to

ensure a unique global solution. 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  $38.92 \text{ ms}$ , 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, I 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. These interleavers will result in  $\mathbf{e}_\pi$  with a sparsity level ranging from 10% to 30% of its total length. While being challenging for traditional compressed sensing algorithms such as Basis Pursuit, such sparsity level can be robustly handled by the SBL-based algorithms, as will be demonstrated by the simulation results.

### 3.4 Nonparametric Impulsive Noise Estimation

The estimation of impulsive noise converts to solving an underdetermined linear regression problem in (3.2) under sparsity constraints. Among various compressed sensing algorithms, SBL has become increasingly attractive due to its improved robustness over deterministic approaches such as Basis Pursuit. 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 I apply the SBL techniques and propose three nonparametric algorithms for impulsive

noise estimation, with different complexity vs. performance trade-offs. For conciseness purpose, I only present the algorithms for estimating asynchronous impulsive noise. The corresponding estimators for periodic impulsive noise can be immediately deduced by replacing  $\mathbf{\Lambda}$  by  $\mathbf{I}$ , and  $\mathbf{e}, \mathbf{g}$  by  $\mathbf{e}_\pi, \mathbf{g}_\pi$ , respectively.

### 3.4.1 Estimation Using Null and Pilot Tones

Since (3.2) is an example of the SMV model, an estimate of  $\mathbf{e}$  can be obtained by applying the SBL algorithm. More specifically, without further prior information, I set  $\mathbf{a} = \mathbf{b} = \mathbf{0}$ , and substitute  $\mathbf{\Phi} = \mathbf{F}_j$ ,  $\mathbf{x} = \mathbf{e}$  and  $\mathbf{v} = \mathbf{g}_j$  into (2.18), (2.19), (2.20) and (2.21), resulting in the update rules for the iterative noise estimator using null and pilot tones, as summarized in Algorithm 2.

---

**Algorithm 2** Impulsive noise estimation using null and pilot tones

---

- 1: Initialize parameters  $\gamma^{(0)} = \mathbf{1}, \sigma^{(0)} = 0$ .
- 2: **for**  $k = 1, \dots, K$  **do**
- 3:   Compute the posterior mean and variance of  $\mathbf{e}$  (E-Step):

$$\boldsymbol{\mu}_e^{(k)} = (\sigma^{(k)})^{-2} \boldsymbol{\Sigma}_e^{(k)} \mathbf{F}_j^* \mathbf{y}, \quad (3.5)$$

$$\boldsymbol{\Sigma}_e^{(k)} = \boldsymbol{\Gamma}^{(k)} - \boldsymbol{\Gamma}^{(k)} \mathbf{F}_j^* \left( (\sigma^{(k)})^2 \mathbf{I}_M + \mathbf{F}_j \boldsymbol{\Gamma}^{(k)} \mathbf{F}_j^* \right)^{-1} \mathbf{F}_j \boldsymbol{\Gamma}^{(k)}. \quad (3.6)$$

- 4:   Update the parameters (M-Step):

$$\gamma_i^{(k+1)} = \boldsymbol{\Sigma}_{e,ii}^{(k)} + (\boldsymbol{\mu}_{e,i}^{(k)})^2, \quad (3.7)$$

$$(\sigma^{(k+1)})^2 = \frac{1}{N_0} \{ \|\mathbf{y} - \mathbf{F}_j \boldsymbol{\mu}_e^{(k)}\|^2 + (\sigma^{(k)})^2 \sum_{i=1}^{N_c} [1 - (\gamma_i^{(k)})^{-1} \boldsymbol{\Sigma}_{e,ii}^{(k)}] \} \quad (3.8)$$

- 5: **end for**
  - 6: Return the MAP estimate  $\hat{\mathbf{e}} = \boldsymbol{\mu}_e^{(K)}$ .
-

Upon termination of the EM algorithm, I obtain the MAP estimate of the time-domain impulsive noise  $\hat{\mathbf{e}} = \boldsymbol{\mu}_e$ . I then transform  $\hat{\mathbf{e}}$  to the frequency domain and subtract it from the received signal in the data tones according to (3.3).

### 3.4.2 Estimation Using All Tones

Similar to other compressed sensing techniques, assuming full-rank dictionaries, the recovery performance of the SBL algorithm will be improved as the number of observations increases. In this case, the sub-FFT matrix  $\mathbf{F}_j$  (i.e., the dictionary) is always full-rank, and therefore the more non-data tones used in the estimation, the more accurate are the impulsive noise estimates. However in a given OFDM system, having more non-data tones indicates reduced throughput. Furthermore, in realistic OFDM systems, part of the null tones are designated for adjacent band interference suppression. These null tones are likely to be suppressed by analog filtering at the receiver front end, and hence cannot be used to estimate impulsive noise accurately.

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, define a length- $N_c$  vector  $\tilde{\mathbf{g}}$  such that  $\tilde{\mathbf{g}}_j = \mathbf{g}_j$  and  $\tilde{\mathbf{g}}_{\bar{j}} = (\boldsymbol{\Lambda}\mathbf{s})_{\bar{j}} + g_{\bar{j}}$ . The SMV model in (3.1) can be complemented by an additional set of linear observations of the impulsive noise

$$\tilde{\mathbf{y}} \triangleq \begin{bmatrix} \mathbf{y} \\ \mathbf{r}_{\bar{j}} \end{bmatrix} = \mathbf{F}\mathbf{e} + \tilde{\mathbf{g}}. \quad (3.9)$$

The extra observations come from the demodulated data tones,  $\mathbf{r}_{\bar{j}}$ . Due to the presence of unknown data signal, the associated noise  $\tilde{\mathbf{g}}_{\bar{j}}$  can be modeled as Gaussian with unknown mean and variance, i.e.,  $p(\tilde{\mathbf{g}}_{\bar{j}}) = \mathcal{CN}((\mathbf{\Lambda}\mathbf{s})_{\bar{j}}, \sigma^2\mathbf{I}_{N_c-N_0})$ .

The SBL algorithm needs to be modified to incorporate the additional unknown parameters  $\mathbf{s}_{\bar{j}}$ . It is straightforward to show that the posterior density of  $\mathbf{e}$  given  $\tilde{\mathbf{y}}$  and the current estimates of  $\gamma, \sigma$  and  $\mathbf{s}_{\bar{j}}$  is Gaussian whose mean and covariance take the same forms as (3.5) and (3.6), while replacing  $\mathbf{F}_{\bar{j}}$  by  $\mathbf{F}$ , and  $\mathbf{y}$  by  $\tilde{\mathbf{y}} - \mathbf{\Lambda}\mathbf{s}^{(k)}$ . Here  $\mathbf{s}^{(k)}$  combines the known null and pilot tones with the estimated data tones in the current iteration. Now let us take a look at the update rules for the three sets of parameters,  $\gamma, \sigma$  and  $\mathbf{s}_{\bar{j}}$ .

- The update rule for  $\gamma$  remains in the same form as (3.7), since according to the EM algorithm,

$$\begin{aligned}
\gamma_i^{(k+1)} &= \operatorname{argmax}_{\gamma_i \geq 0} \mathbb{E}_{\mathbf{e}|\tilde{\mathbf{y}};\gamma^{(k)},\sigma^{(k)},\mathbf{s}_{\bar{j}}^{(k)}} [\log p(\tilde{\mathbf{y}}|\mathbf{e}, \sigma^{(k)}, \mathbf{s}_{\bar{j}}^{(k)})p(\mathbf{x}|\gamma)] \\
&= \operatorname{argmax}_{\gamma_i \geq 0} \mathbb{E}_{\mathbf{e}|\tilde{\mathbf{y}};\gamma^{(k)},\sigma^{(k)},\mathbf{s}_{\bar{j}}^{(k)}} [\log p(\mathbf{x}|\gamma)] \\
&= \Sigma_{\mathbf{e},ii}^{(k)} + (\boldsymbol{\mu}_{\mathbf{e},i}^{(k)})^2.
\end{aligned} \tag{3.10}$$

In other words, the estimate of  $\mathbf{s}_{\bar{j}}^{(k)}$  affects the update rule for  $\gamma$  only through the posterior mean and covariance of  $\mathbf{e}$ .

- The update rule for  $\sigma$  now becomes

$$\begin{aligned}
(\sigma^{(k+1)})^2 &= \operatorname{argmax}_{\sigma \geq 0} \mathbb{E}_{\mathbf{e}|\tilde{\mathbf{y}};\gamma^{(k)},\sigma^{(k)},\mathbf{s}_{\bar{j}}^{(k)}} [\log p(\tilde{\mathbf{y}}|\mathbf{e}, \sigma, \mathbf{s}_{\bar{j}}^{(k)})] \\
&= \frac{1}{N_c} \{ \|\tilde{\mathbf{y}} - \mathbf{\Lambda}\mathbf{s}^{(k)} - \mathbf{F}\boldsymbol{\mu}_{\mathbf{e}}^{(k)}\|^2 \\
&\quad + (\sigma^{(k)})^2 \sum_{i=1}^{N_c} [1 - (\gamma_i^{(k)})^{-1} \Sigma_{\mathbf{e},ii}^{(k)}] \}. \tag{3.11}
\end{aligned}$$

- Although  $\mathbf{s}_{\bar{j}}$  consists of discrete valued constellation points, I temporarily relax it to be continuous to ensure the convergence of the EM algorithm.

The update rule for  $\mathbf{s}_{\bar{j}}$  can be easily derived as

$$\begin{aligned}
\mathbf{s}_{\bar{j}}^{(k+1)} &= \operatorname{argmax}_{\mathbf{s}_{\bar{j}}} \mathbb{E}_{\mathbf{e}|\tilde{\mathbf{y}};\gamma^{(k)},\sigma^{(k)},\mathbf{s}_{\bar{j}}^{(k)}} [\log p(\tilde{\mathbf{y}}|\mathbf{e}, \sigma^{(k)}, \mathbf{s}_{\bar{j}})] \\
&= \operatorname{argmin}_{\mathbf{s}_{\bar{j}}} |\tilde{\mathbf{y}}_{\bar{j}} - (\mathbf{\Lambda}\mathbf{s})_{\bar{j}} - \mathbf{F}_{\bar{j}}\boldsymbol{\mu}_{\mathbf{e}}^{(k)}|^2 \\
&= \mathbf{\Lambda}_{\bar{j}}^{-1}(\tilde{\mathbf{y}}_{\bar{j}} - \mathbf{F}_{\bar{j}}\boldsymbol{\mu}_{\mathbf{e}}^{(k)}). \tag{3.12}
\end{aligned}$$

The entire EM algorithm is summarized in Algorithm 3. Intuitively, the algorithm estimates the impulsive noise and the signal transmitted on the data tones iteratively, assuming the knowledge of one when estimating the other.

### 3.4.3 Decision Feedback Estimation

The estimators described above do not impose any prior information on the impulsive noise other than the fact that it is a sparse vector. As mentioned in Section 2.5.1, additional knowledge about the impulsive noise can be incorporated into the SBL inference using the hierarchical prior in (2.13) and (2.14) with non-zero values of  $\mathbf{a}$  and  $\mathbf{b}$ . Suppose that in addition

---

**Algorithm 3** Impulsive noise estimation using all tones
 

---

- 1: Initialize parameters  $\boldsymbol{\gamma}^{(0)} = \mathbf{1}$ ,  $\sigma^{(0)} = 0$ ,  $\mathbf{s}_{\bar{j}}^{(0)} = \mathbf{0}$ .
- 2: **for**  $k = 1, \dots, K$  **do**
- 3:   Compute the posterior mean and variance of  $\mathbf{e}$  (E-Step):

$$\boldsymbol{\mu}_{\mathbf{e}}^{(k)} = (\sigma^{(k)})^{-2} \boldsymbol{\Sigma}_{\mathbf{e}}^{(k)} \mathbf{F}^* (\tilde{\mathbf{y}} - \boldsymbol{\Lambda} \mathbf{s}^{(k)}), \quad (3.13)$$

$$\boldsymbol{\Sigma}_{\mathbf{e}}^{(k)} = \boldsymbol{\Gamma}^{(k)} - \boldsymbol{\Gamma}^{(k)} \mathbf{F}^* \left( (\sigma^{(k)})^2 \mathbf{I}_M + \mathbf{F} \boldsymbol{\Gamma}^{(k)} \mathbf{F}^* \right)^{-1} \mathbf{F} \boldsymbol{\Gamma}^{(k)}. \quad (3.14)$$

- 4:   Update the parameters (M-Step):

$$\gamma_i^{(k+1)} = \boldsymbol{\Sigma}_{\mathbf{e}, ii}^{(k)} + (\boldsymbol{\mu}_{\mathbf{e}, i}^{(k)})^2, \quad (3.15)$$

$$\begin{aligned} (\sigma^{(k+1)})^2 &= \frac{1}{N_c} \{ \|\tilde{\mathbf{y}} - \boldsymbol{\Lambda} \mathbf{s}^{(k)} - \mathbf{F} \boldsymbol{\mu}_{\mathbf{e}}^{(k)}\|^2 \\ &\quad + (\sigma^{(k)})^2 \sum_{i=1}^{N_c} [1 - (\gamma_i^{(k)})^{-1} \boldsymbol{\Sigma}_{\mathbf{e}, ii}^{(k)}] \}, \end{aligned} \quad (3.16)$$

$$\mathbf{s}_{\bar{j}}^{(k+1)} = \boldsymbol{\Lambda}_{\bar{j}}^{-1} (\tilde{\mathbf{y}}_{\bar{j}} - \mathbf{F}_{\bar{j}} \boldsymbol{\mu}_{\mathbf{e}}^{(k)}). \quad (3.17)$$

- 5: **end for**

- 6: Return the MAP estimate  $\hat{\mathbf{e}} = \boldsymbol{\mu}_{\mathbf{e}}^{(K)}$ .
- 

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  $\boldsymbol{\gamma}$  given  $\hat{\mathbf{e}}'$ . Since the hyper-prior in (2.14) is conjugate to (2.13), the posterior density of  $\boldsymbol{\gamma}$  given  $\hat{\mathbf{e}}'$ ,  $\mathbf{a}$  and  $\mathbf{b}$  is also an inverse Gamma distribution, i.e.,

$$P(\boldsymbol{\gamma} | \hat{\mathbf{e}}'; \mathbf{a}, \mathbf{b}) = \prod_{i=1}^{N_c} \text{IG}(\gamma_i; \tilde{a}_i, \tilde{b}_i) \quad (3.23)$$



with the updated parameters

$$\begin{aligned}\tilde{a}_i &= a_i + \frac{1}{2}, \\ \tilde{b}_i &= b_i + \frac{|e_i|^2}{2}, \quad \forall i = 1, \dots, N_c.\end{aligned}\tag{3.24}$$

One can then proceed with the SBL inference using (3.23) as the hyper-prior.

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  $\hat{\mathbf{c}}$  one can recover the data tones of the OFDM symbols by appropriate constellation mapping. This gives an estimate of  $\hat{\mathbf{s}}_{\bar{j}}$ , which is multiplied by the channel frequency response  $\mathbf{\Lambda}_{\bar{j}}$ , transformed to the time domain and subtracted from the received signal  $\mathbf{r}$  to generate the estimate  $\hat{\mathbf{e}}'$ . Then I 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, I 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. The algorithm is summarized in Algorithm 4 and the receiver structure is depicted in Fig. 3.3. Compared to the estimator using all tones in Section 3.4.2, the decision feedback estimator is expected to have better performance by exploiting the redundant information (i.e., coding structure) on the data tones.

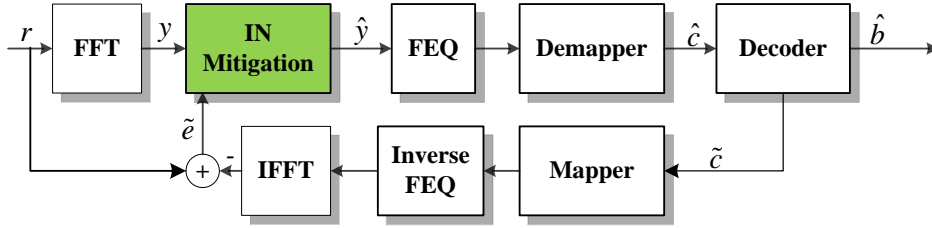


Figure 3.3: A decision feedback impulsive noise (IN) estimator for an OFDM receiver.

### 3.5 Low-Complexity Implementation

The proposed noise estimators require a matrix inversion operation per iteration. For example, the estimator using null tones requires an  $N_0 \times N_0$  matrix inversion in (3.6), where  $N_0$  is the number of null and pilot tones. For typical values of  $N_0$  in PLC, such matrix inversions might be practically infeasible on a hardware platform. To allow low-complexity implementation of the proposed algorithms while maintaining the same performance, one could apply the sequential SBL algorithm introduced in Section 2.5.2.

Table 3.1 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 for example in (3.6). Compared to the estimator using null and pilot tones, the estimator using all tones increases the complexity from  $\mathcal{O}(N_c^2 N_0)$  per iteration to  $\mathcal{O}(N_c^3)$  per iteration, where  $N_c$  is the FFT size. On the other hand, each iteration of the sequential SBL involves matrix multiplications and inversions that have complexities of  $\mathcal{O}(N_c^2 N_B)$  and  $\mathcal{O}(N_B^3)$ ,

<b>Estimator</b>	<b>Operation</b>	<b>Complexity</b>
Using null and pilot tones	Matrix multiply	$\mathcal{O}(N_c^2 N_0)$
	Matrix inversion	$\mathcal{O}(N_0^3)$
Using all tones	Matrix multiply	$\mathcal{O}(N_c^3)$
	Matrix inversion	$\mathcal{O}(N_c^3)$
Sequential SBL w/ unknown background noise power	Matrix multiply	$\mathcal{O}(N_c^2 N_B)$
	Matrix inversion	$\mathcal{O}(N_B^3)$
Sequential SBL w/ known background noise power	Matrix multiply	$\mathcal{O}(N_c^2 N_B)$

Table 3.1: Complexity per iteration of the proposed algorithms.  $N_c$  is the FFT size,  $N_0$  is the number of null and pilot tones, and  $N_B$  is the number of model bases in the current iteration.

respectively, where  $N_B$  is the number of bases that remains in that particular iteration. As mentioned in Section 2.5.2, further complexity reduction is possible by making use of the knowledge of background noise power, which eliminates any matrix inversion operations.

### 3.6 Simulation Results

To evaluate the performance of the proposed algorithms, I simulate a complex baseband OFDM system over a flat channel. The system parameters are listed in Table 3.2 and compared with those in the G3 standard operating in the CENELEC-A band. In all simulations, I use the SBL algorithm in its full-complexity version, since the sequential SBL algorithm has the same reconstruction performance.

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 3.2: 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.

### 3.6.1 Performance in Asynchronous Impulsive Noise

I 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 [95]. The noise samples are assumed to be independent and identically distributed (i.i.d.). Traces of the simulated asynchronous impulsive noise have been shown earlier in Figures 2.2 and 2.3.

In asynchronous impulsive noise, I compare the BER performance of the proposed algorithms with the Basis Pursuit based algorithm [16]. To com-

pare the performance of these nonparametric methods to the parametric ones, I also implement the two MMSE detectors in [38], since they are optimal in the MMSE sense among other parametric methods such as nulling and clipping [93]. Both MMSE detectors assume perfect knowledge of the GM model parameters (with the truncated MCA as a special case), and one even assumes perfect noise state information (NSI), i.e., noise variance at each time instance. In reality, the NSI can be estimated using decision feedback. More specifically, based on the decision feedback, the receiver can estimate the impulsive noise samples in time domain. The NSI can be obtained by categorizing each noise sample into one of the multiple Gaussian components in the GM distribution. Performance of such MMSE detector with decision feedback is upper-bounded by the MMSE detector with perfect NSI. 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. 3.5. For conciseness purposes, I 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 (decision feedback)”, the Basis Pursuit based algorithm as “CS”, and the two MMSE detectors as “MMSE w/o NSI” and “MMSE w/ NSI”, respectively.

In the uncoded system, the proposed estimator using null tones achieves 6–8 dB SNR gain over conventional OFDM receivers. One can obtain addi-

tional 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 3.4.2). However, such error quickly becomes negligible as the SNR increases. All the 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, the 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 receivers. The estimator using all tones provides an additional 2–5 dB gains. Furthermore, using decision feedback from the convolutional decoder, I 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 the proposed estimators. As mentioned previously in Section 3.2, 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 the system settings.

### 3.6.2 Performance in Periodic Impulsive Noise

I generate periodic impulsive noise using the LPTV system model in [62]. I divide one period of the noise into three intervals, each assuming an individual spectral shape (Figure 2.4). The spectral shapes are fitted to noise measurement collected at an outdoor low-voltage site as shown in [62]. I 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, I simulate the 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 [38] 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, I 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 3.3 spans 2.3 AC cycles.

With the interleaver size fixed at approximately an AC cycle, I increase the noise burst duration from 10% to 30% of a period. The BER performance of all algorithms are plotted in Figure 3.7. Without any noise mitigation, the TDI-OFDM system 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 [9] 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, I am 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 the 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. I notice that the SNR gains obtained by the proposed algorithms over the TDI-OFDM system itself are smaller than in the previous 10% burst case. 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 the proposed algorithms to different interleaver sizes, I 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 the proposed algorithms. In Figure 3.7, comparing the BER performance in the bottom right plot to that in the top right plot, I observe



<b>System</b>	<b>Noise</b>	<b>SBL w/ null tones</b>	<b>SBL w/ all tones</b>	<b>SBL w/ DF</b>
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 3.3: 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.

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, the proposed algorithms achieve significant BER improvement over conventional OFDM systems without noise mitigation in various SNR regions. For clarity purposes, I 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 3.3.

### 3.7 Conclusion

In this part of the dissertation, I propose 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, I 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. In periodic impulsive noise, I 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 nonparametric; i.e. they do not require prior knowledge on the statistical noise model or model parameters. I validate the proposed algorithms based on asynchronous impulsive noise and periodic impulsive noise simulated from various statistical models.

---

**Algorithm 4** Impulsive noise estimation using decision feedback
 

---

1: Initialize parameters  $\boldsymbol{\gamma}^{(0)} = \mathbf{1}$ ,  $\sigma^{(0)} = 0$ ,  $\mathbf{a}^{(0)} = 0$ ,  $\mathbf{b}^{(0)} = 0$ .

2: **for**  $k = 1, \dots, K$  **do**

3:   Compute the posterior mean and variance of  $\mathbf{e}$  (E-Step):

$$\boldsymbol{\mu}_{\mathbf{e}}^{(k)} = (\sigma^{(k)})^{-2} \boldsymbol{\Sigma}_{\mathbf{e}}^{(k)} \mathbf{F}_{\mathbf{j}}^* \mathbf{y}, \quad (3.18)$$

$$\boldsymbol{\Sigma}_{\mathbf{e}}^{(k)} = \boldsymbol{\Gamma}^{(k)} - \boldsymbol{\Gamma}^{(k)} \mathbf{F}_{\mathbf{j}}^* \left( (\sigma^{(k)})^2 \mathbf{I}_M + \mathbf{F}_{\mathbf{j}} \boldsymbol{\Gamma}^{(k)} \mathbf{F}_{\mathbf{j}}^* \right)^{-1} \mathbf{F}_{\mathbf{j}} \boldsymbol{\Gamma}^{(k)}. \quad (3.19)$$

4:   Update the parameters (M-Step):

$$\gamma_i^{(k+1)} = \frac{\boldsymbol{\Sigma}_{\mathbf{e},ii}^{(k)} + (\boldsymbol{\mu}_{\mathbf{e},i}^{(k)})^2 + 2b_i^{(k)}}{1 + 2a_i^{(k)}}, \quad \forall i = 1, \dots, N_c, \quad (3.20)$$

$$(\sigma^{(k+1)})^2 = \frac{1}{N_0} \left\{ \|\mathbf{y} - \mathbf{F}_{\mathbf{j}} \boldsymbol{\mu}_{\mathbf{e}}^{(k)}\|^2 + (\sigma^{(k)})^2 \sum_{i=1}^{N_c} [1 - (\gamma_i^{(k)})^{-1} \boldsymbol{\Sigma}_{\mathbf{e},ii}^{(k)}] \right\} \quad (3.21)$$

5:   Denoise the received signal:  $\hat{\mathbf{y}}^{(k)} = \mathbf{y} - \mathbf{F} \boldsymbol{\mu}_{\mathbf{e}}^{(k)}$ , equalize and decode  $\hat{\mathbf{y}}^{(k)}$ .

6:   Estimate slicing error  $\hat{\boldsymbol{\epsilon}}'$ , and update parameters:

$$\begin{aligned} a_i^{(k+1)} &= a_i^{(k)} + \frac{1}{2}, \\ b_i^{(k+1)} &= b_i^{(k)} + \frac{|\hat{\epsilon}'_i|^2}{2}, \quad \forall i = 1, \dots, N_c. \end{aligned} \quad (3.22)$$

7: **end for**

8: Return the decoder output  $\hat{\mathbf{b}}$ .

---

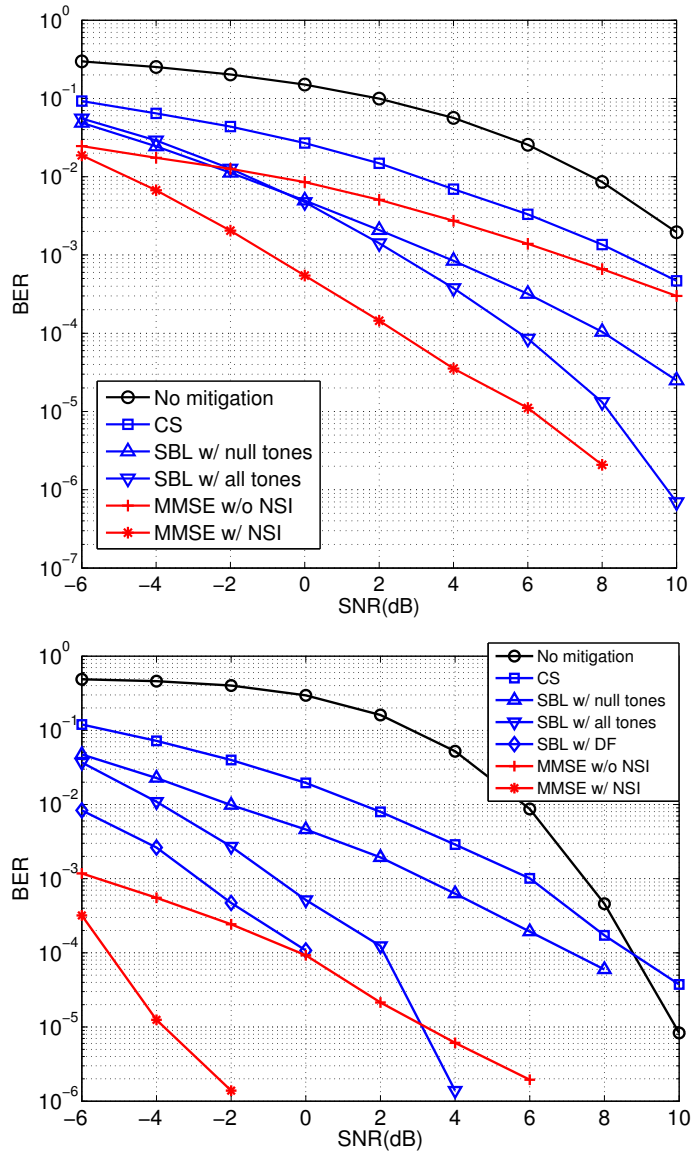


Figure 3.4: Uncoded (top) and coded (bottom) BER performance of the proposed algorithms in Gaussian mixture modeled asynchronous impulsive noise, in comparison with the conventional OFDM system without noise mitigation, the compressed sensing based algorithm, and two parametric MMSE detectors with and without noise state information.

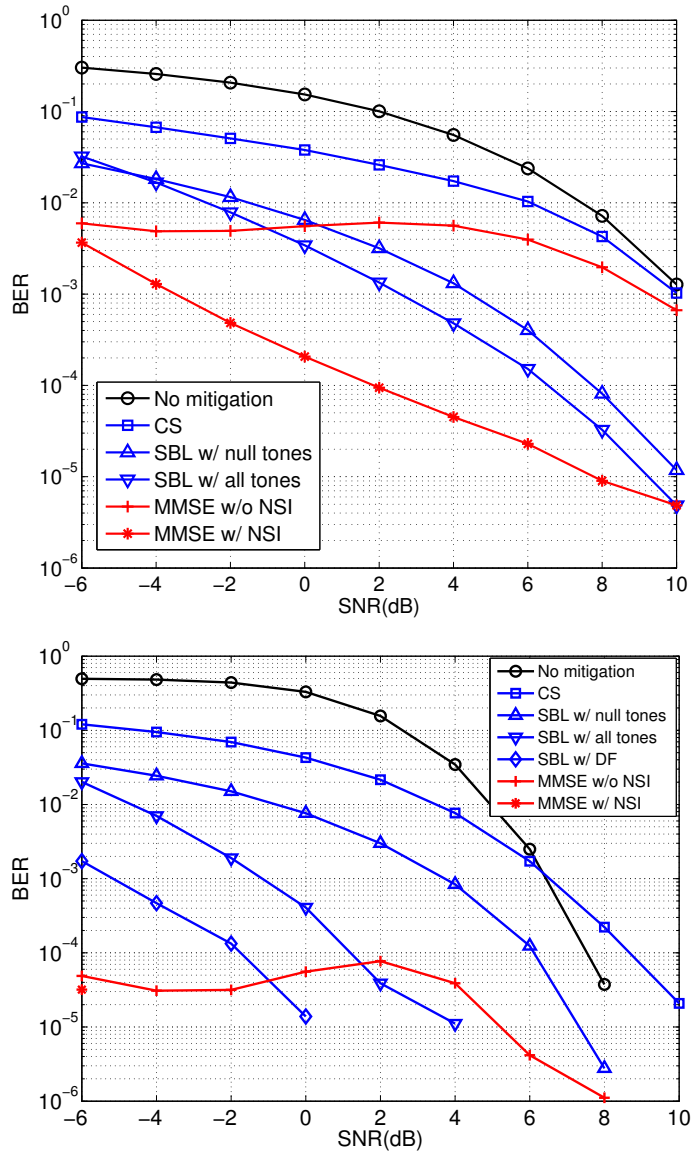


Figure 3.5: Uncoded (top) and coded (bottom) BER performance of the proposed algorithms in Middleton Class A modeled asynchronous impulsive noise, in comparison with the conventional OFDM system without noise mitigation, the compressed sensing based algorithm, and two parametric MMSE detectors with and without noise state information.

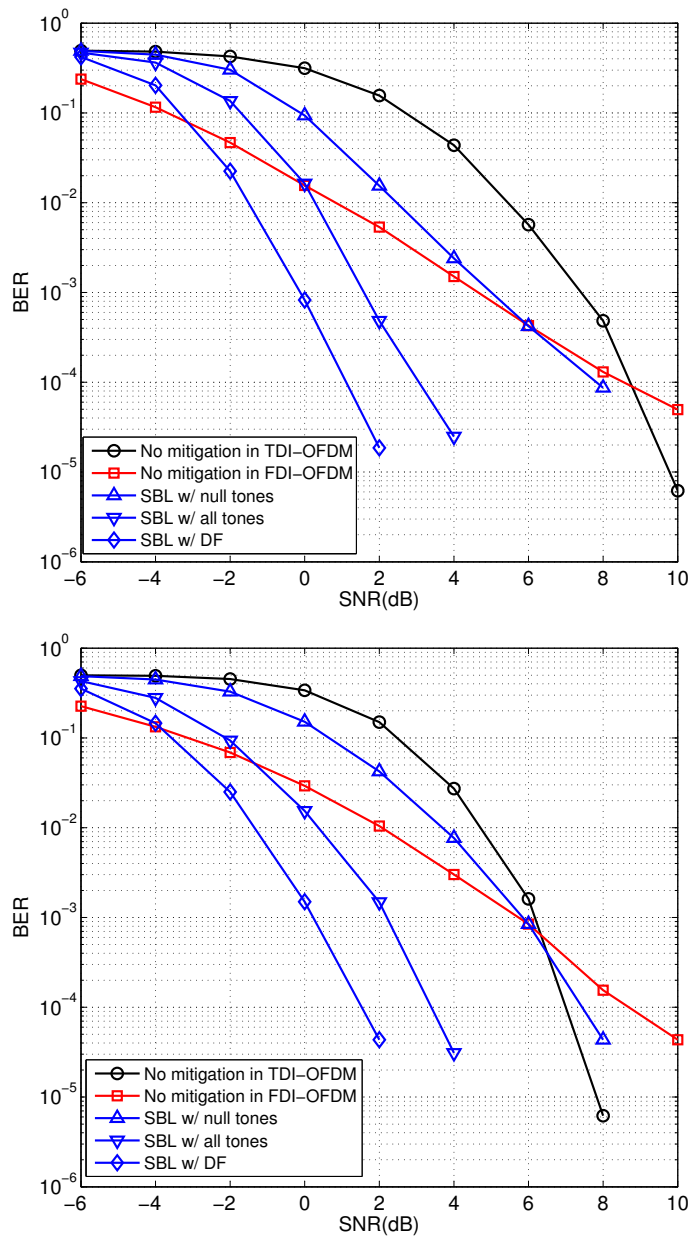


Figure 3.6: Coded BER performance of the proposed algorithms in periodic impulsive noise, in comparison with the TDI-OFDM and FDI-OFDM systems without noise mitigation. The interleaving is done over an entire AC cycle. The burst interval varies from 10% (top) to 30% (bottom) of a period.

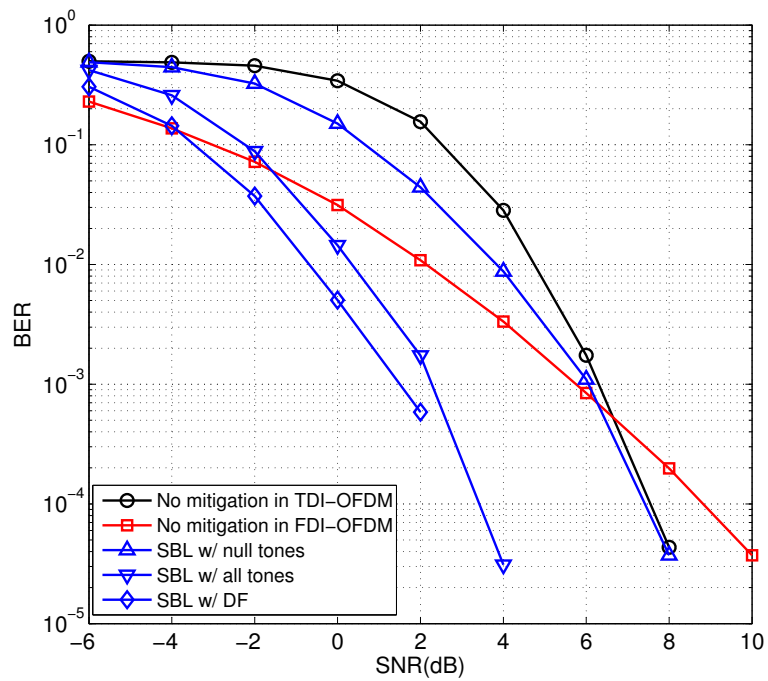


Figure 3.7: Coded BER performance of the proposed algorithms in periodic impulsive noise, in comparison with the TDI-OFDM and FDI-OFDM systems without noise mitigation. The interleaving is done over half the AC cycle. The burst interval is 30% of a period.

## Chapter 4

# Time-Frequency Modulation Diversity to Combat Periodic Impulsive Noise

Periodic impulsive noise synchronous to the main powerline frequency severely limits the communication performance of narrowband PLC systems. The periodic impulsive noise statistics deviate significantly from that of additive white Gaussian noise (AWGN), thereby causing dramatic performance degradation in conventional narrowband PLC systems. In this part of the dissertation, I propose a robust transmission scheme and corresponding receiver methods to combat periodic impulsive noise in OFDM-based narrowband PLC. Towards that end, I propose (1) a time-frequency modulation diversity scheme at the transmitter and a diversity demodulator at the receiver to improve communication reliability without decreasing data rates; and (2) a semi-online algorithm that exploits the sparsity of the noise in the frequency domain to estimate the noise power spectrum for reliable decoding at the diversity demodulator. In the simulations, our proposed transceiver methods reduce the bit error rates in conventional OFDM systems by more than 1000x at typical SNRs in narrowband PLC.



## 4.1 Introduction

In this part of the dissertation, I aim to improve the reliability of OFDM-based narrowband PLC systems in periodic impulsive noise by joint transmitter and receiver design. At the transmitter, I propose a time-frequency modulation diversity technique that can be embedded into existing narrowband PLC standards to improve the robustness of transmission in periodically varying and spectrally shaped noise. On the receiver side, I present a diversity combining demodulator that reliably decodes the signal based on the periodically varying noise power spectrum. To estimate the noise power spectrum, I discuss an offline algorithm based on noise measurements during no-transmission intervals. Furthermore, I apply sparse Bayesian learning (SBL) techniques and develop a semi-online algorithm that estimates the noise power spectrum primarily during data transmission, by exploiting its sparsity in frequency domain.

The rest of this chapter is organized as follows. Section 4.2 gives an overview of existing transmitter and receiver methods to combat periodic impulsive noise. In Section 4.3, I briefly review the principle of modulation diversity, and introduce the time-frequency modulation diversity technique. I present in Section 4.4 the diversity combining demodulator for non-coherent (differential) and coherent systems, respectively. I then discuss the offline training based estimator for the noise power spectrum, and derive an online alternative based on SBL in Section 4.5. Finally, the performance of the proposed transceiver methods is evaluated and discussed in Section 4.6.

## 4.2 Prior Work

Prior work on combating periodic impulsive noise in OFDM systems involves efforts from both the transmitter and receiver's perspectives. In [70], it was demonstrated that non-uniform bit loading based on the periodically-varying and frequency-selective sub-channel SNRs provides significant improvement in reliability for a given throughput. On the other hand, assuming no knowledge of channel/noise information at transmitter, transmitters specified in existing narrowband PLC standards rely on forward error correction coding and frequency-domain block interleaving to cope with impulsive noise. In particular, it was suggested to use concatenated forward error correction codes (i.e., convolutional, Reed-Solomon and repetition codes) to enhance the error correction capability in harsh channel and noise environments. Heavy coding, however, sacrifices throughput (or equivalently, requires bandwidth expansion) for improved reliability. Most narrowband PLC standards adopt frequency-domain interleaving, i.e., bit-level interleaving prior to the IFFT. Alternatively, sample-level time-domain and time-frequency-domain interleaving were proposed in [9] and [69] due to their superior performance over frequency-domain interleaving at high SNRs. However, in order to store the continuous-valued time-domain signal, these interleavers require considerably larger memory than the bit-level frequency-domain interleavers, while the performance improvement is not as significant at low to moderate SNRs.

On the receiver side, pre-processing methods have been developed to mitigate the impact of periodic impulsive noise. These methods exploited var-

ious statistical properties of the noise, according to which they can be divided into three categories. The first type of approach is based on the cyclostationary noise model [62]. These methods parameterized the second-order noise statistics in each stationary interval by either a correlation matrix [89] or filter coefficients [29, 51, 52], and estimated the parameters by training. Based on the estimated parameters, equalizers [52, 89] or prediction filters [29, 51] were designed to essentially transform the spectrally shaped noise into AWGN. These parametric approaches are sensitive to parameter estimation errors. Accurate parameter estimation generally requires large training overhead and a huge memory that might not exist at narrowband PLC receivers.

A second class of receiver methods utilize the impulsive nature of the noise to estimate and subtract the noise impulses from the received signal. A common approach among these methods is to observe the impulsive noise on the null and pilot subcarriers of the received OFDM symbols, and apply compressed sensing techniques to recover the noise, assuming its sparse structure in time domain. For example, to estimate noise bursts shorter than an OFDM symbol, a block compressed sensing algorithm was applied in [50] to estimate the locations and amplitudes of the bursts. To estimate longer noise bursts that span more than one OFDM symbol, the previous chapter of this dissertation adopts time-domain block interleaving to spread the bursts into short impulses over multiple OFDM symbols, and estimated the deinterleaved noise by sparse Bayesian learning.

Recognizing the harmonic structure of the noise, another type of re-

ceiver methods reuses existing techniques for narrowband interference mitigation. Simple techniques include time-domain Nyquist windowing applied to the received signal to suppress the spectral leakage from the narrowband interference [26]. Given a model of the power spectral density of the narrowband interference, linear minimum mean square error estimators were used to estimate the spectral leakage of the narrowband interference by observing a number of subcarriers in an OFDM symbol [71]. Assuming that the narrowband interference consists of a single frequency sinusoid, subspace methods were used to extract the sinusoid from the received OFDM symbol [23]. More recently, by exploiting sparsity of narrowband interference in the frequency domain, compressed sensing algorithms were proposed to estimate the narrowband interference from the time-domain guard intervals in a zero-padded OFDM system [32]. However, most existing algorithms were designed to combat narrowband interference with limited number (e.g. one or two) of peaks, and may not be able to handle the rich harmonic components in the periodic impulsive noise (as shown in Figure 2.4).

My work distinguishes from previous studies in two aspects. First of all, unlike concatenated coding, the proposed diversity modulation scheme improves communication reliability without decreasing data rates. On the other hand, while most existing receiver techniques attempt to whiten or remove the noise in the pre-processing stage, the proposed receiver design alters the demodulator as well to incorporate knowledge on the periodic impulsive noise statistics. Furthermore, I will demonstrate that the proposed receiver methods

(1) do not require the memory-consuming time-domain interleaving as in the second contribution; and (2) are able to mitigate noise that has a low sparsity level in time or frequency domain.

### 4.3 Time-Frequency Diversity Modulation

In general, modulation diversity (MD) refers to modulation schemes that jointly map a number of bits to a multi-dimensional constellation point (a.k.a. an MD codeword), each dimension of which is a real or complex symbol. The goal is to improve communication reliability without throughput reduction or bandwidth expansion. This is achieved by spreading information over multiple symbols that are transmitted over independent channels, thereby exploiting channel diversity.

Several categories of MD codes have been investigated in the literature [15, 81], among which the Hochwald/Sweldens codes have attracted a lot of attention [41, 42]. A Hochwald/Sweldens codebook defines a one-to-one mapping from any group of  $N_d R$  bits to a length- $N_d$  vector of phase shift keying (PSK) symbols:

$$f : \mathbf{c} \in \{0, 1\}^{N_d R} \rightarrow \mathbf{s} \in \mathcal{C}^{N_d}. \quad (4.1)$$

Here  $R$  denotes the data rate in bits per symbol, and  $\mathcal{C}$  denotes a  $2^{N_d R}$ -PSK constellation. The  $n$ -th component of a particular codeword  $\mathbf{s}^{(m)}$  has the form

$$s_n^{(m)} = \exp(j2\pi u_n m / 2^{N_d R}), \quad (4.2)$$

$\forall m = 1, \dots, 2^{N_d R}, \forall n = 1, \dots, N_d$ . When  $u_n = 1, \forall n$ , the Hochwald/Sweldens

code reduces to a PSK repetition code [81]. In general,  $\mathbf{u} \triangleq [u_1 \cdots u_{N_d}]$  need to be optimized to minimize the symbol error rate in specific channel conditions. Assuming that all components of a codeword are transmitted over static flat channels or *i.i.d.* flat Rayleigh fading channels and corrupted by AWGN, the optimal values of  $\mathbf{u}$  for  $R = 1$  and  $N_d = 2$  to 4 have been found by exhaustive search [42] and are summarized in Table 4.1.

A second type of MD codes defines a multi-dimensional rotated cubic lattice constellation [15, 81]. The constellation can be viewed as carved from a translated and scaled version of the cubic lattice  $\mathbb{Z}^{N_d}$  ( $\mathbb{Z}$  is the integer set), and rotated by a certain linear transform. The rotation increases the minimum number of distinct components between any two constellation points, i.e., the diversity order. Suppose that some dimensions of the received signal are attenuated by the channel, the compressed constellation in the subspace spanned by other dimensions offers more protection, since no or very few points collapse together as would happen in unrotated constellations.

MD has demonstrated significant communication performance improvement in multi-antenna communications [41] and in OFDM systems [12]. In [41], a unitary space-time modulation scheme was derived based on the MD codes. In [12], dual carrier modulation was proposed to exploit channel di-

Table 4.1: Optimal parameters for the Hochwald/Sweldens code in *i.i.d.* flat Rayleigh fading and AWGN [42].  $N_d$  is the length of a codeword.

$N_d$	2	3	4
$[u_1 \cdots u_{N_d}]$	[1 1]	[1 1 3]	[1 3 5 7]

iversity in frequency domain in ultra-wideband wireless communications. Dual carrier modulation adopts a special MD codebook that maps every four bits to two different 16-QAM symbols. The symbols are transmitted on two subcarriers that are separated apart in the spectrum, and are supposed to experience independent sub-channels.

In OFDM-based narrowband PLC systems, recognizing the periodically varying spectrally shaped statistics of periodic impulsive noise, I propose to apply modulation diversity across time and frequency, which leads to the time-frequency modulation diversity (TFMD). In TFMD, components of an MD codeword are allocated to various subcarriers in multiple OFDM symbols. Locations of the subcarriers are chosen so that narrowband interference or a deep fade in frequency domain impacts as few symbols of the same codeword as possible. Likewise, the designated OFDM symbols are separated apart in time to prevent being simultaneously contaminated by noise bursts or impulses. An example of the time-frequency mapping is depicted in Figure 4.1, assuming  $N_d = 2$ .

I adopt a Hochwald/Sweldens codebook since it produces PSK symbols (PSK is a mandatory modulation scheme in PRIME, G3 and IEEE P1901.2 narrowband PLC standards). Ideally, the parameters  $\mathbf{u}$  need to be optimized taking into account the PLC channel models. In particular, amplitude of PLC channels is typically characterized by log-normal distributions rather than Rayleigh fading [27, 67]. Furthermore, for a given PLC link, the channel frequency response is fairly static and can be determined from the topology and

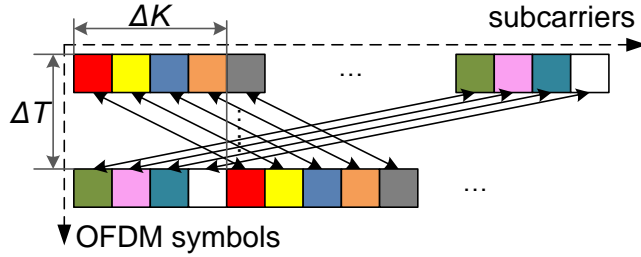


Figure 4.1: An example of time-frequency modulation diversity for length-two MD codes. Components of an MD codeword (marked in the same color) are allocated to subcarriers  $i$  and  $((i + \Delta K) \bmod N_c)$  in OFDM symbols  $j$  and  $(j + \Delta T)$ , respectively,  $\forall i, j$ , where  $N_c$  is the number of subcarriers in an OFDM symbol.

impedance properties of the cable [22, 67]. Since this chapter is focused on combating periodic impulsive noise, I leave the design of optimal codebooks based on PLC channel models for future work. The receiver methods presented in the following can be applied to arbitrary values of  $\mathbf{u}$ .

The TFMD scheme can be integrated into existing narrowband PLC standards as a robust transmission mode in harsh channel and noise environments. The block diagram for an OFDM-based narrowband PLC transmitter using TFMD is shown in Figure 4.2. The binary data is protected by a convolutional code followed by a bit-level block interleaver. The diversity modulator maps every  $N_d$  bits to  $N_d$  PSK symbols, which are then allocated to designated time-frequency slots by the time-frequency (TF) mapper. An optional differential encoder follows to differentially encode the OFDM symbols. Without loss of generality, I assume that the differential encoding is performed in the frequency domain (i.e., each subcarrier is differentially encoded upon the



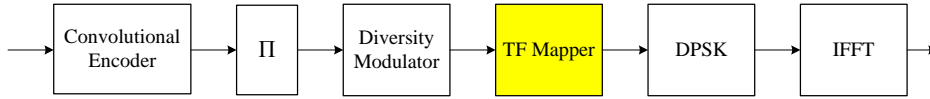


Figure 4.2: The block diagram of a narrowband PLC transmitter using time-frequency modulation diversity. The proposed time-frequency (TF) mapper allocates components of every Hochwald/Sweldens codeword to different subcarriers in various OFDM symbols.  $\Pi$  denotes the bit-level block interleaver.

previous subcarrier in the same OFDM symbol). The OFDM symbols are converted to the time domain via IFFT. Finally a cyclic prefix (CP) is inserted to the beginning of each OFDM symbol to prevent inter-symbol interference.

#### 4.4 Soft-Output Diversity Demodulator

In this section, I present a diversity demodulator that generates soft decisions on the modulated bits based on the periodically varying spectrally shaped statistics of periodic impulsive noise. I consider both non-coherent and coherent systems. As will be shown by the end of the section, the non-coherent and coherent demodulators essentially take the same parametric form.

Let us focus on a particular group of bits,  $\mathbf{c} \in \{0, 1\}^{N_d}$ , that is jointly mapped to an MD codeword. Suppose that the  $n$ -th component of the codeword is mapped to subcarrier  $i_n$  in the  $j_n$ -th OFDM symbol,  $\forall n = 1, \dots, N_d$ , which is denoted by  $v_{i_n j_n}$ . To simplify notations, define index sets  $I \triangleq \{i_n\}_{n=1}^{N_d}$  and  $J \triangleq \{j_n\}_{n=1}^{N_d}$ , and a column vector  $\mathbf{v}_{IJ} \triangleq \{v_{i_n j_n}\}_{n=1}^{N_d}$ . In non-coherent systems, a differential encoder follows the time-frequency mapping, which results

in

$$\mathbf{s}_{IJ} = \mathbf{v}_{IJ} \odot \mathbf{s}_{I-1,J}, \quad (4.3)$$

where  $\mathbf{s}_{I-1,J} \triangleq \{s_{i_n-1,j_n}\}_{n=1}^{N_d}$ , and  $\odot$  denotes the pointwise product between two vectors. After that, the OFDM symbols are converted to the time domain, with the CP inserted.

Upon receiving an OFDM symbol, the receiver discards the CP and transforms the rest of the symbol to the frequency domain by taking an FFT. One of the most important properties of OFDM is that it divides a frequency-selective channel to multiple narrowband flat sub-channels, each of which is centered at a subcarrier frequency. More specifically, assuming that the CP is longer than the channel delay spread, the output of FFT can be expressed as

$$\mathbf{r}_{IJ} = \mathbf{h}_{IJ} \odot \mathbf{s}_{IJ} + \mathbf{e}_{IJ}. \quad (4.4)$$

Effectively, the DPSK (or PSK) symbol on the  $i_n$ -th subcarrier of the  $j_n$ -th OFDM symbol is transmitted over a flat sub-channel, whose complex amplitude is denoted by  $h_{i_n j_n}$ . The received signal on the subcarrier is corrupted by an additive noise  $e_{i_n j_n}$  in frequency domain.

In non-coherent systems, a soft-output diversity demodulator computes a log-likelihood ratio (LLR) for each bit in  $\mathbf{c}$ , given  $\mathbf{r}_{IJ}$  and the reference signal

$\mathbf{r}_{I-1,J}$ . The conditional LLR for  $c_n$  ( $\forall n \in \{1, \dots, N_d\}$ ) can be written as

$$\begin{aligned} L(c_n) &= \ln \frac{p(c_n = 0 | \mathbf{r}_{IJ}, \mathbf{r}_{I-1,J})}{p(c_n = 1 | \mathbf{r}_{IJ}, \mathbf{r}_{I-1,J})} \\ &= \ln \frac{\sum_{\mathbf{v}_{IJ}:c_n=0} p(\mathbf{r}_{IJ} | \mathbf{v}_{IJ}, \mathbf{r}_{I-1,J})}{\sum_{\mathbf{v}_{IJ}:c_n=1} p(\mathbf{r}_{IJ} | \mathbf{v}_{IJ}, \mathbf{r}_{I-1,J})} \end{aligned} \quad (4.5)$$

$$\approx \ln \frac{\max_{\mathbf{v}_{IJ}:c_n=0} p(\mathbf{r}_{IJ} | \mathbf{v}_{IJ}, \mathbf{r}_{I-1,J})}{\max_{\mathbf{v}_{IJ}:c_n=1} p(\mathbf{r}_{IJ} | \mathbf{v}_{IJ}, \mathbf{r}_{I-1,J})}. \quad (4.6)$$

In (4.5), it is assumed that after the bit-level interleaver,  $c_n = 0$  or  $1$  with equal probability and hence  $\mathbf{v}_{IJ}$  is uniformly distributed over the codebook. In (4.6), the max-sum approximation of the exact LLR is applied to reduce the computational complexity as in [74]. To evaluate the likelihood expressions in (4.6), notice from (4.3) and (4.4) that

$$\begin{aligned} \mathbf{r}_{IJ} &= \mathbf{h}_{IJ} \odot \mathbf{s}_{I-1,J} \odot \mathbf{v}_{IJ} + \mathbf{e}_{IJ} \\ &\approx (\mathbf{r}_{I-1,J} - \mathbf{e}_{I-1,J}) \odot \mathbf{v}_{IJ} + \mathbf{e}_{IJ} \end{aligned} \quad (4.7)$$

$$= \mathbf{r}_{I-1,J} \odot \mathbf{v}_{IJ} + \tilde{\mathbf{e}}_{IJ}, \quad (4.8)$$

where  $\tilde{\mathbf{e}}_{IJ} \triangleq -\mathbf{e}_{I-1,J} \odot \mathbf{v}_{IJ} + \mathbf{e}_{IJ}$ . In (4.7), I make the approximation that the channel frequency response does not vary significantly between adjacent sub-channels, and hence  $\mathbf{h}_{IJ} \approx \mathbf{h}_{I-1,J}$ . In addition, I assume that the additive noise samples on different subcarriers are mutually uncorrelated, and that components of an MD codeword are mapped to different subcarriers. Given these assumptions, and the fact that  $|v_{i_n j_n}| = 1$ , it can be verified that

$$\mathbb{E}[\tilde{\mathbf{e}}_{IJ} \tilde{\mathbf{e}}_{IJ}^*] = \text{diag}\{\tilde{\boldsymbol{\sigma}}_{IJ}\}, \quad (4.9)$$

$$\tilde{\boldsymbol{\sigma}}_{IJ} = \boldsymbol{\sigma}_{I-1,J} + \boldsymbol{\sigma}_{IJ}, \quad (4.10)$$

where  $\sigma_{i_n j_n} \triangleq \mathbb{E}[|e_{i_n j_n}|^2], \forall n$ , are the amplitudes of the periodically varying noise power spectrum. Furthermore, according to the LPTV model for periodic impulsive noise [62], the noise samples are spectrally shaped Gaussian random variables, and therefore  $\tilde{\mathbf{e}}_{IJ} \sim \mathcal{CN}(\mathbf{0}, \tilde{\boldsymbol{\sigma}}_{IJ})$ . From (4.8) and (4.9), the likelihoods in (4.6) can be written as

$$p(\mathbf{r}_{IJ} | \mathbf{v}_{IJ}, \mathbf{r}_{I-1, J}) = \mathcal{CN}(\mathbf{r}_{I-1, J} \odot \mathbf{v}_{IJ}, \tilde{\boldsymbol{\sigma}}_{IJ}). \quad (4.11)$$

Substituting (4.11) into (4.6) results in the LLR expression

$$L_{MRC}(c_n) = \max_{\mathbf{v}_{IJ}: c_n=0} \sum_{k=1}^{N_d} \frac{\text{Re}\{r_{i_k j_k} v_{i_k j_k}^* r_{i_k-1, j_k}^*\}}{\tilde{\sigma}_{i_k j_k}} - \max_{\mathbf{v}_{IJ}: c_n=1} \sum_{k=1}^{N_d} \frac{\text{Re}\{r_{i_k j_k} v_{i_k j_k}^* r_{i_k-1, j_k}^*\}}{\tilde{\sigma}_{i_k j_k}}. \quad (4.12)$$

Compared to conventional DPSK detection [25], the diversity demodulator (4.12) combines the decision metrics from  $N_d$  time-frequency slots, with the weights inversely proportional to the sub-channel noise variances. Due to its similarity to maximum ratio combining (MRC) reception in multi-antenna communications [92], I will refer to (4.12) as the MRC diversity demodulator. For low complexity implementation, one could also consider the sub-optimal selection combining (SC) diversity demodulator

$$L_{SC}(c_n) = \max_{\mathbf{v}_{IJ}: c_n=0} \max_k \frac{\text{Re}\{r_{i_k j_k} v_{i_k j_k}^* r_{i_k-1, j_k}^*\}}{\tilde{\sigma}_{i_k j_k}} - \max_{\mathbf{v}_{IJ}: c_n=1} \max_k \frac{\text{Re}\{r_{i_k j_k} v_{i_k j_k}^* r_{i_k-1, j_k}^*\}}{\tilde{\sigma}_{i_k j_k}}. \quad (4.13)$$

The diversity demodulators (4.12) and (4.13) use two-symbol observation windows for differential detection. They can also be generalized to

use multiple-symbol differential detection, given that increasing the window size generally leads to better performance [25, 74, 80]. The generalized diversity demodulator computes LLRs conditioned on  $\{\mathbf{r}_{IJ}, \mathbf{r}_{I-1,J}, \dots, \mathbf{r}_{I-N_w+1,J}\}$ , where  $N_w$  is the window size. However, multiple-symbol differential detection requires much higher implementation complexity, and hence might be infeasible for low-power low-cost PLC systems. As such, throughout the rest of the chapter, I will only use two-symbol observation windows for differential detection.

In coherent systems, the received signal after the FFT can be written as

$$\mathbf{r}_{IJ} \approx \hat{\mathbf{h}}_{IJ} \odot \mathbf{v}_{IJ} + \mathbf{e}_{IJ}, \quad (4.14)$$

where  $\hat{\mathbf{h}}_{IJ}$  denotes the estimated channel frequency response. Note the similar mathematical forms in (4.14) and (4.8). The diversity demodulator for coherent systems can be derived immediately from (4.12) or (4.13), by replacing  $r_{i_k-1,j_k}$  with  $\hat{h}_{i_k j_k}$ , and  $\tilde{\sigma}_{i_k j_k}$  with  $\sigma_{i_k j_k}$ .

The non-coherent OFDM receiver structure with the diversity demodulator is depicted in Figure 4.3. In coherent systems, the differential detection block is replaced by a frequency-domain channel equalizer. The LLRs generated by the diversity demodulator are first de-interleaved, and then decoded by the convolutional decoder.

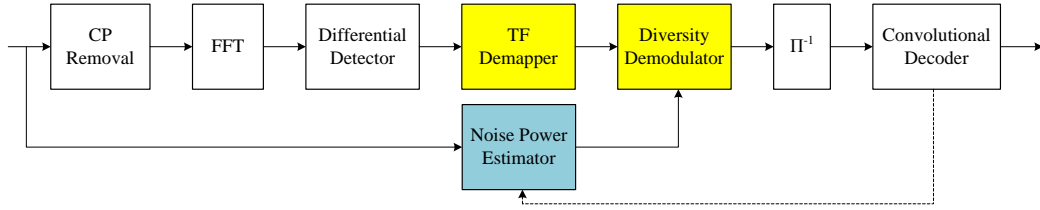


Figure 4.3: The block diagram of a narrowband PLC receiver with non-coherent diversity demodulation. The time-frequency (TF) demapper groups subcarriers in OFDM symbols into MD codewords. The diversity demodulator outputs soft information based on the noise power spectrum, which can be estimated before or during transmission (as will be discussed in Section 4.5).  $\Pi^{-1}$  denotes the bit-level block de-interleaver.

## 4.5 Noise Power Estimation

The soft-output diversity demodulators (4.12) and (4.13) assume knowledge of the noise power spectrum. In this section, I will discuss two distinct approaches to estimate the noise power spectrum, both of which exploit the cyclostationarity of periodic impulsive noise.

PLC systems typically emit periodic bursty transmissions. Considering a single PLC link, the silent (i.e., no-transmission) interval between consecutive bursts typically takes several minutes. For example, a smart meter reports customer load profile to a data concentrator every 15 minutes [13]. During a silent interval, the receiver can collect a noise trace that extends multiple AC cycles, and use that to estimate the periodically varying noise power spectrum. In fact, since a period of noise can be partitioned into  $T$  stationary intervals [62], noise in the  $t$ -th interval ( $\forall t = 1, \dots, T$ ) has a time-invariant power

spectrum

$$\boldsymbol{\sigma}^{(t)} \triangleq \{\sigma_i^{(t)}\}_{i=1}^{N_c}, \quad (4.15)$$

$$\sigma_i^{(t)} \triangleq \sigma_{ij}, \forall j \in \mathbb{S}^{(t)} \quad (4.16)$$

where  $\mathbb{S}^{(t)}$  is the index set for all OFDM symbols received in the  $t$ -th stationary interval. The receiver can estimate the  $T$  noise power spectrums individually, by taking sliding FFTs over corresponding noise samples and averaging the power of the FFT outputs. I will refer to such approach as offline noise power estimation.

In order to obtain accurate estimates, a large number of noise samples need to be recorded, which entails a huge memory that might not exist at typical narrowband PLC receivers. Furthermore, within a PLC network, power lines are a shared medium, and hence PLC devices need to be scheduled according to resource sharing protocols (e.g. carrier sense multiple access or time division multiple access) to limit uncoordinated interference [28]. As the number of PLC devices increases, due to the limited silent intervals between transmissions from different PLC devices, a receiver might not be able to collect enough noise samples to accurately estimate the noise power spectrum.

In light of the drawbacks of offline estimation, it is desirable to develop algorithms to estimate the noise power spectrum primarily during data transmission, with minor effort during silent intervals. Such semi-online noise power estimators are also more adaptive to any fluctuation of noise statistics in addition to the periodic variation.

At coherent receivers, a heuristic semi-online estimator can be developed using decision feedback from the convolutional decoder. During the silent intervals, the receiver only needs to obtain a rough estimate of the start/end instances of all stationary intervals within a period. In general, the average noise power varies significantly among different stationary intervals. Therefore, these intervals can be easily identified by amplitude thresholding in time domain. As the transmission starts, the receiver computes the slicing errors by subtracting, from the received signal, the signal that is reconstructed based on the decoder output. The receiver can then update the estimated noise power spectrum in a stationary interval with the average slicing errors over all OFDM symbols received during that stationary interval. An iterative receiver can thus be formulated, which starts with the AWGN assumption and keeps updating the noise power spectrum through iterations. Note that such semi-online heuristics cannot be used at non-coherent receivers, since it requires channel estimation to reconstruct the received signal from the decoder output.

As will be demonstrated by simulation results, the heuristic semi-online estimator is prone to error propagation, since it completely relies on decision feedback. In the following, I will focus on coherent systems, and develop a semi-online estimator that exploits sparsity of periodic impulsive noise in frequency domain, and uses the CP and decision feedback to obtain more accurate estimate of the noise power spectrum.



### 4.5.1 Problem Formulation

Let  $N_c$  and  $N_p$  denote the FFT size and CP length in an OFDM system, and  $L_h$  the channel delay spread in samples. Most narrowband PLC standards adopt a CP that is much longer than typical channel delay spreads (i.e.,  $N_p \gg L_h$ ). For example, the root-mean-square delay spreads encountered in the field measurements are  $2 - 6\mu s$ , while the effective CP duration for narrowband PLC in the CELENEC-A band is  $55\mu s$  in IEEE P1901.2 and up to  $120\mu s$  in G.hnem [67]. The first  $L_h$  samples inside the CP are affected by inter-symbol-interference. Removing these samples from the  $j$ -th received OFDM symbol in time domain results in

$$\dot{\mathbf{r}}_j = \mathbf{H}_j \mathbf{S} \mathbf{F}_{N_c}^* \mathbf{s}_j + \dot{\mathbf{e}}_j. \quad (4.17)$$

Here  $\mathbf{F}_{N_c}$  is the  $N_c$ -point FFT matrix,  $\mathbf{S} \triangleq \begin{bmatrix} \mathbf{0}_{N_c - N_p + L_h} & \mathbf{I}_{N_p - L_h} \\ & \mathbf{I}_{N_c} \end{bmatrix}$ ,  $\mathbf{H}_j$  is a Toeplitz matrix consisting of a time shifted channel impulse response in each row,  $\dot{\mathbf{r}}_j$  and  $\dot{\mathbf{e}}_j$  denote the time-domain received signal and additive noise, respectively.

Define a matrix  $\mathbf{W}$  as

$$\mathbf{W} \triangleq \begin{bmatrix} \mathbf{A} & \mathbf{0}_{(N_p - L_h) \times (N_c - N_p + L_h)} & -\mathbf{A} \end{bmatrix}, \quad (4.18)$$

where  $\mathbf{A}$  is an arbitrary  $(N_p - L_h) \times (N_p - L_h)$  unitary matrix. It can be easily proved that  $\mathbf{W} \mathbf{H}_j \mathbf{S} = \mathbf{0}, \forall \mathbf{A}$ , as long as  $\mathbf{H}_j$  is Toeplitz. Therefore pre-multiplying  $\dot{\mathbf{r}}_j$  by  $\mathbf{W}$  removes the information bearing portion of the received signal

$$\mathbf{y}_j \triangleq \mathbf{W} \dot{\mathbf{r}}_j = \mathbf{W} \dot{\mathbf{e}}_j. \quad (4.19)$$

As previously mentioned, the power of periodic impulsive noise is concentrated around a few frequencies due to narrowband interference. The noise can therefore be decomposed in frequency domain as

$$\dot{\mathbf{e}}_j = \mathbf{F}_N^*(\mathbf{x}_j + \mathbf{g}_j), \quad (4.20)$$

where  $N = N_c + N_p - L_h$ ,  $\mathbf{x}_j$  and  $\mathbf{g}_j$  represent the narrowband interference and background noise in frequency domain, respectively. Although in reality the background noise  $\mathbf{g}_j$  is spectrally shaped, I model it as AWGN for simplicity. Such approximation will not cause significant estimation error for the overall noise power spectrum, since the background noise has much lower power compared to the narrowband interference. Defining  $\mathbf{\Phi} \triangleq \mathbf{W}\mathbf{F}_N^*$  and  $\mathbf{v}_j \triangleq \mathbf{W}\mathbf{F}_N^*\mathbf{g}_j$ , (4.19) can be succinctly rewritten as

$$\mathbf{y}_j = \mathbf{\Phi}\mathbf{x}_j + \mathbf{v}_j. \quad (4.21)$$

Note that  $\mathbf{v}_j$  is still AWGN since  $\mathbf{A}$  and  $\mathbf{F}_N$  are both unitary.

Suppose that during the silent intervals, the receiver has estimated the start/end instances of all stationary intervals within a period of the noise. We can collect the measurements  $\mathbf{y}_j, \forall j \in \mathbb{S}^{(t)}$ , and expand (4.21) into

$$\mathbf{Y}^{(t)} = \mathbf{\Phi}\mathbf{X}^{(t)} + \mathbf{V}^{(t)}, \quad (4.22)$$

where  $\mathbf{Y}^{(t)}$  is a matrix formed by column vectors  $\{\mathbf{y}_j, \forall j \in \mathbb{S}^{(t)}\}$ , and similarly for  $\mathbf{X}^{(t)}$  and  $\mathbf{V}^{(t)}$ . Recall that  $\{\mathbf{x}_j, \forall j \in \mathbb{S}^{(t)}\}$  are instantaneous spectrums of the narrowband interference in the  $t$ -th stationary interval. Mathematically, all

columns of  $\mathbf{X}^{(t)}$  are sparse vectors that share an identical support. Considering  $\mathbf{Y}^{(t)}$  as a measurement matrix,  $\Phi$  as an  $(N_p - L_h) \times N$  known dictionary matrix,  $\mathbf{X}^{(t)}$  as an unknown source matrix with each row representing a possible source, and  $\mathbf{V}^{(t)}$  as an unknown noise matrix, (4.22) boils down to the standard MMV problem in compressed sensing [91].

#### 4.5.2 Noise Power Spectrum Estimation Using T-MSBL

While promoting sparsity, the prior (2.22) adopted by the T-MSBL algorithm is also a close match to the statistical model for periodic impulsive noise. The prior essentially assumes that the narrowband interference  $\mathbf{x}_j$  follows a Gaussian distribution with zero mean and covariance  $\mathbf{\Gamma}$ . This is in accordance with the LPTV model for periodic impulsive noise [62], except for the additional approximation that the elements of  $\mathbf{x}_j$  are uncorrelated. The uncorrelated assumption can be justified by the fact that the correlation between different frequency components of the narrowband interference is much lower in amplitude compared to the interference peaks; i.e.,  $\mathbb{E}[\mathbf{x}_j \mathbf{x}_j^*]$  can be approximated by a diagonal matrix.

However, applying T-MSBL directly to our problem may lead to inaccurate estimation of  $\gamma$ . This is because to ensure a unique global solution, the number of non-zero rows in  $\mathbf{X}$  has to be below a certain threshold. Unfortunately, given typical system settings of narrowband PLC, such threshold is in general too restrictive compared to the actual number of peaks in the interference spectrum. To overcome this challenge, I propose a T-MSBL based

estimator that incorporates decision feedback from the convolutional decoder into the prior to improve the recovery performance.

The key idea is to use a more informative prior that effectively guides the Bayesian inference to converge to the actual values of  $\boldsymbol{\gamma}$ , despite of the low sparsity level. Such prior can be constructed by imposing a hierarchical prior on  $\boldsymbol{\gamma}$  and  $\mathbf{B}$ , respectively. I adopt conjugate priors since they generally lead to computationally tractable solutions:

$$p(\boldsymbol{\gamma}; \mathbf{a}, \mathbf{b}) = \prod_{i=1}^N \text{IG}(\gamma_i; a_i, b_i), \quad (4.23)$$

$$p(\mathbf{B}; \mu, \boldsymbol{\Psi}) = \text{IW}(\mathbf{B}; \mu, \boldsymbol{\Psi}). \quad (4.24)$$

Here  $\text{IG}(\gamma_i; a_i, b_i)$  is the inverse Gamma distribution with the shape parameter  $a_i$  and scale parameter  $b_i$ , both of which assume non-negative values;  $\text{IW}(\mathbf{B}; \mu, \boldsymbol{\Psi})$  denotes the inverse Wishart distribution, where  $\mu > L$  is the degree of freedom and the scale parameter  $\boldsymbol{\Psi}$  is a positive definite matrix. Particular values of the hyperparameters  $\mathbf{a}, \mathbf{b}, \mu$  and  $\boldsymbol{\Psi}$  reflect certain prior knowledge, or side information, on  $\mathbf{X}$ .

Consider the vectorized MMV model (2.23). To find the hyperparameters  $\boldsymbol{\theta} = \{\boldsymbol{\gamma}, \mathbf{B}, \lambda\}$  that maximize  $p(\boldsymbol{\theta} | \bar{\mathbf{y}}, \mathbf{a}, \mathbf{b}, \mu, \boldsymbol{\Psi})$ , we can employ the EM algorithm. The EM algorithm treats  $\bar{\mathbf{x}}$  as latent variables, and seeks to maxi-

mize

$$\begin{aligned}
Q(\boldsymbol{\theta}) &= \mathbb{E}_{\bar{\mathbf{x}}|\bar{\mathbf{y}};\boldsymbol{\theta}^{(\text{old})}} \left\{ \log [p(\bar{\mathbf{y}}, \bar{\mathbf{x}}|\boldsymbol{\theta})p(\boldsymbol{\theta})] \right\} \\
&= \mathbb{E}_{\bar{\mathbf{x}}|\bar{\mathbf{y}};\boldsymbol{\theta}^{(\text{old})}} [\log p(\bar{\mathbf{y}}|\bar{\mathbf{x}}, \boldsymbol{\theta})] + \\
&\quad \mathbb{E}_{\bar{\mathbf{x}}|\bar{\mathbf{y}};\boldsymbol{\theta}^{(\text{old})}} [\log p(\bar{\mathbf{x}}|\boldsymbol{\gamma}, \mathbf{B})] + \\
&\quad \log p(\boldsymbol{\gamma}) + \log p(\mathbf{B}), \tag{4.25}
\end{aligned}$$

where  $\boldsymbol{\theta}^{(\text{old})}$  denotes the estimate of unknown parameters in the previous iteration.

Since the only term in (4.25) that depends on  $\lambda$  is the first term, the update rule for  $\lambda$  remains the same as in the original T-SBL algorithm [91]. Note that this term is independent of  $\boldsymbol{\gamma}$  and  $\mathbf{B}$ . Therefore to estimate  $\boldsymbol{\gamma}$  and  $\mathbf{B}$ , the cost function (4.25) can be simplified to

$$\begin{aligned}
Q(\boldsymbol{\gamma}, \mathbf{B}) &= \mathbb{E}_{\bar{\mathbf{x}}|\bar{\mathbf{y}};\boldsymbol{\theta}^{(\text{old})}} [\log p(\bar{\mathbf{x}}|\boldsymbol{\gamma}, \mathbf{B})] + \\
&\quad \log p(\boldsymbol{\gamma}) + \log p(\mathbf{B}). \tag{4.26}
\end{aligned}$$

It can be shown that [91]

$$\begin{aligned}
\mathbb{E}_{\bar{\mathbf{x}}|\bar{\mathbf{y}};\boldsymbol{\theta}^{(\text{old})}} [\log p(\bar{\mathbf{x}}|\boldsymbol{\gamma}, \mathbf{B})] &\propto -\frac{L}{2} \log(|\boldsymbol{\Gamma}|) - \frac{N}{2} \log(|\mathbf{B}|) \\
&\quad -\frac{1}{2} \text{Tr}[(\boldsymbol{\Gamma}^{-1} \otimes \mathbf{B}^{-1}) \\
&\quad (\boldsymbol{\Sigma}_{\bar{\mathbf{x}}} + \boldsymbol{\mu}_{\bar{\mathbf{x}}}\boldsymbol{\mu}_{\bar{\mathbf{x}}}^*)], \tag{4.27}
\end{aligned}$$

where  $\boldsymbol{\Sigma}_{\bar{\mathbf{x}}}$  and  $\boldsymbol{\mu}_{\bar{\mathbf{x}}}$  are evaluated as in the original T-SBL algorithm [91]. Replacing  $p(\boldsymbol{\gamma})$  and  $p(\mathbf{B})$  by the probability density function of the inverse

Gamma distribution and inverse Wishart distribution, respectively, we have

$$\log p(\boldsymbol{\gamma}) \propto \sum_{i=1}^N [-a_i \log \gamma_i - b_i/\gamma_i] \quad (4.28)$$

$$\log p(\mathbf{B}) \propto -\frac{\mu + L}{2} \log |\mathbf{B}| - \frac{1}{2} \text{Tr}(\boldsymbol{\Psi} \mathbf{B}^{-1}) \quad (4.29)$$

The derivative of (4.26) with respect to  $\gamma_i (i = 1, \dots, N)$  is

$$\begin{aligned} \frac{\partial Q}{\partial \gamma_i} \propto & -\frac{L}{2\gamma_i} + \frac{1}{2\gamma_i^2} \text{Tr}[\mathbf{B}^{-1}(\boldsymbol{\Sigma}_{\bar{\mathbf{x}},i} + \boldsymbol{\mu}_{\bar{\mathbf{x}},i} \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^*)] \\ & -\frac{a_i + 1}{\gamma_i} + \frac{b_i}{\gamma_i^2}. \end{aligned} \quad (4.30)$$

Setting the derivative to zero, we obtain the update rule for  $\gamma_i$  in an EM iteration:

$$\gamma_i \leftarrow \frac{\text{Tr}[\mathbf{B}^{-1}(\boldsymbol{\Sigma}_{\bar{\mathbf{x}},i} + \boldsymbol{\mu}_{\bar{\mathbf{x}},i} \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^*)] + 2b_i}{L + 2a_i}. \quad (4.31)$$

Similarly, we take the derivative of (4.26) over  $\mathbf{B}$

$$\begin{aligned} \frac{\partial Q}{\partial \mathbf{B}} = & -\frac{N + \mu + L}{2} \mathbf{B}^{-1} + \frac{1}{2} \mathbf{B}^{-1} \\ & \left[ \sum_{i=1}^N \frac{1}{\gamma_i} (\boldsymbol{\Sigma}_{\bar{\mathbf{x}},i} + \boldsymbol{\mu}_{\bar{\mathbf{x}},i} \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^*) + \boldsymbol{\Psi}^* \right] \mathbf{B}^{-1}. \end{aligned} \quad (4.32)$$

Setting the derivative to zero gives the update rule for  $\mathbf{B}$ :

$$\mathbf{B} \leftarrow \frac{1}{N + \mu + L} \left[ \sum_{i=1}^N \frac{1}{\gamma_i} (\boldsymbol{\Sigma}_{\bar{\mathbf{x}},i} + \boldsymbol{\mu}_{\bar{\mathbf{x}},i} \boldsymbol{\mu}_{\bar{\mathbf{x}},i}^*) + \boldsymbol{\Psi}^* \right]. \quad (4.33)$$

Following the same approximation as in [91], we can back-map the update rules in (4.31) and (4.33) to the original space, resulting in the update

rules for  $\boldsymbol{\gamma}$  and  $\mathbf{B}$

$$\begin{aligned} \gamma_i^{(k+1)} &= \frac{\hat{\mathbf{X}}_{i \cdot}^{(k)} (\mathbf{B}^{(k)})^{-1} \hat{\mathbf{X}}_{i \cdot}^{(k)*} + L \boldsymbol{\Xi}_{ii}^{(k)} + 2b_i^{(k)}}{L + 2a_i^{(k)}}, \\ &\forall i = 1, \dots, N, \end{aligned} \quad (4.34)$$

$$\begin{aligned} \tilde{\mathbf{B}}^{(k+1)} &= \sum_{i=1}^N \frac{\hat{\mathbf{X}}_{i \cdot}^{(k)*} \hat{\mathbf{X}}_{i \cdot}^{(k)}}{\gamma_i^{(k)}} + \boldsymbol{\Psi}^{(t)*}, \\ \mathbf{B}^{(k+1)} &= \tilde{\mathbf{B}}^{(k+1)} / \|\tilde{\mathbf{B}}^{(k+1)}\|_{\mathcal{F}}, \end{aligned} \quad (4.35)$$

where  $\hat{\mathbf{X}}^{(k)}$  and  $\boldsymbol{\Xi}^{(k)}$  are evaluated as in (2.28). Compared to (2.29), the prior information contained in  $\mathbf{a}, \mathbf{b}, \boldsymbol{\mu}$  and  $\boldsymbol{\Psi}$  does affect the estimation of  $\boldsymbol{\gamma}$  and  $\mathbf{B}$ .

In coherent systems, we can extract side information on  $\mathbf{X}$  from the output of the convolutional decoder. Using the hard decision output from the decoder, the receiver can reconstruct an estimate of the transmit signal by repeating the entire transmitter chain. The reconstructed transmit signal is then filtered by the estimated channel, and subtracted from the real received signal. The residual in the frequency domain provides a second estimate of  $\mathbf{X}$ , denoted by  $\hat{\mathbf{X}}'$ .

The side information in  $\hat{\mathbf{X}}'$  can be fused into the hierarchical prior formed by (2.22), (4.23) and (4.24) via the hyperparameters. More specifically, the posterior of  $\boldsymbol{\gamma}$  (or  $\mathbf{B}$ ) given  $\hat{\mathbf{X}}'$  is an inverse Gamma (or inverse Wishart)

distribution:

$$\begin{aligned}
p(\mathbf{B}|\hat{\mathbf{X}}'; \mu, \boldsymbol{\Psi}, \boldsymbol{\gamma}) &= \text{IW}(\mathbf{B}; \tilde{\mu}, \tilde{\boldsymbol{\Psi}}), \\
\tilde{\mu} &= \mu + N, \\
\tilde{\boldsymbol{\Psi}} &= \boldsymbol{\Psi} + \sum_{i=1}^N \frac{\hat{\mathbf{X}}_{i \cdot}^* \hat{\mathbf{X}}_{i \cdot}'}{\gamma_i}
\end{aligned} \tag{4.36}$$

$$\begin{aligned}
p(\boldsymbol{\gamma}|\hat{\mathbf{X}}'; \mathbf{a}, \mathbf{b}, \mathbf{B}) &= \prod_{i=1}^N \text{IG}(\gamma_i; \tilde{a}_i, \tilde{b}_i), \\
\tilde{a}_i &= a_i + \frac{L}{2}, \\
\tilde{b}_i &= b_i + \frac{1}{2} \|\hat{\mathbf{X}}_{i \cdot}^* \mathbf{Q} \sqrt{\boldsymbol{\Lambda}}^{-1}\|_2^2,
\end{aligned} \tag{4.37}$$

where the unitary matrix  $\mathbf{Q}$  and the diagonal matrix  $\boldsymbol{\Lambda}$  satisfy that  $\mathbf{B} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^*$ . The hierarchical prior with the updated hyperparameters  $\tilde{\mathbf{a}}, \tilde{\mathbf{b}}, \tilde{\mu}$ , and  $\tilde{\boldsymbol{\Psi}}$  can then be used in the Bayesian inference (4.34) and (4.35) to incorporate the side information provided by the decision feedback.

Upon convergence of the algorithm, we can compute the desired  $N_c$ -point noise power spectrum  $\boldsymbol{\sigma}^{(t)}$  from  $\boldsymbol{\Gamma}^{(t)}$ , since

$$\boldsymbol{\sigma}^{(t)} = \text{diag}\left\{ \mathbf{P}\boldsymbol{\Gamma}^{(t)}\mathbf{P}^* \right\}, \tag{4.38}$$

where  $\mathbf{P} = \mathbf{F}_{N_c} \begin{bmatrix} \mathbf{0}_{N_p-L_h} & \mathbf{I}_{N_c} \end{bmatrix} \mathbf{F}_{N_c}^*$ . The estimated  $\boldsymbol{\sigma}^{(t)}$  is sent to the diversity demodulator to improve its reliability. The soft output from the demodulator is then decoded by the convolutional decoder, the decision from which is used to further update the hyperparameters. As such, I form a decision feedback estimator that transfers information back-and-forth between the noise power spectrum estimator and the convolutional decoder.



## 4.6 Simulation Results

To evaluate the communication performance of the proposed transceiver methods in periodic impulsive noise, I simulate an OFDM system with parameters in Table 4.2. Values assumed by these parameters are typical in narrow-band PLC standards (c.f. [67]). I inject to the received signal a representative noise trace collected at an outdoor low-voltage power line (Figure 2.4) [62]. From the spectrogram of this particular noise trace, it can be visually identified that each period can be approximately partitioned into three stationary intervals: a low-power interval, a moderate-power bursty interval, and a high-power impulsive interval. For simplicity, I assume static flat channel in all the experiments.

At the transmitter, I adopt the length-two and length-three Hochwald /Sweldens codes with parameters specified in Table 4.1. The time-frequency mapping is implemented as follows. Let us start by dividing the transmission band into  $N_d$  subbands ( $N_d = 2$  or  $3$ ), each of which contains the same num-

<b>Parameters</b>	<b>Values</b>
Sampling Frequency	400 kHz
FFT Size $N_c$	256
CP Length $N_p$	30
Number of Subcarriers	128
Data Subcarriers	33 : 104
Convolutional code	rate 1/2, length 7
Packet Size	256 Bytes
Interleaver Size	72 bits (1 OFDM symbol)

Table 4.2: Parameters of the simulated OFDM system.

ber of data subcarriers. Components of an MD codeword are assigned to  $N_d$  subcarriers that are located in different subbands and are equally-spaced in frequency. Furthermore, the  $N_d$  subcarriers are allocated to multiple nonconsecutive OFDM symbols. The separation between these OFDM symbols in time domain is determined based on the duration of noise bursts, which spans approximately three consecutive OFDM symbols in the selected noise trace (Figure 2.4). As such, the two subcarriers that are used to carry a length-two codeword are located in two OFDM symbols separated by two other symbols. This ensures that at most one component of the codeword is contaminated by a noise burst. Similarly, for  $N_d = 3$ , the first two components of a codeword are allocated to the  $j$ -th and the  $(j + 3)$ -th OFDM symbols. The third component, however, is assigned to the  $j$ -th OFDM symbol as well. In fact, the noise power spectrum has a low-pass feature; i.e., the third subband located at the high-frequency end of the spectrum experiences low-power noise throughout the period. Therefore, location of the third component in time might not pose significant impact on communication performance. Constraining an MD codeword within a small number of OFDM symbols also helps to reduce buffer size and processing latency at the receiver.

At the receiver, I implement the MRC diversity demodulator (4.12), due to its optimality and manageable complexity considering the small sized codebooks. The three noise power spectrums (one for each stationary interval) are estimated offline in both coherent and non-coherent systems, or using semi-online methods in coherent systems. The bit error rate (BER) performance

is measured over  $10^4$  packets, within the SNR range of -8 dB to 2 dB, which covers low to moderate SNR values in typical narrowband PLC systems [67].

#### 4.6.1 Performance with Offline Noise Power Estimation

I first evaluate communication performance of the TFMD transceiver with offline noise power estimation. The noise power spectrum for each of the three stationary intervals is estimated by taking  $N_c$ -point sliding FFTs over  $10^4$  noise samples within the same stationary interval and averaging the instantaneous power spectrums.

I compare the proposed transceiver methods with several reference designs in terms of BER performance. A baseline reference is a conventional OFDM system using BPSK (or DBPSK) modulation and bit-level interleaving over one OFDM symbol. Since the proposed time-frequency mapping permutes signal in both time and frequency domains across four OFDM symbols, for fair comparison I also simulate the conventional BPSK OFDM system with interleaving over four OFDM symbols (i.e., 288 bits). In addition, a degenerate case of the TFMD scheme, which maps components of a codeword to different subcarriers within the same OFDM symbol, is implemented, and will be referred to as frequency modulation diversity (FMD) in the following.

The BER performance of various transceiver designs in non-coherent and coherent systems, respectively, is plotted in Figure 4.4. In non-coherent systems, the SNR required for the proposed transceiver to achieve a BER of  $10^{-4}$  is 0.4 dB for the length-two code and -3.5 dB for the length-three code.

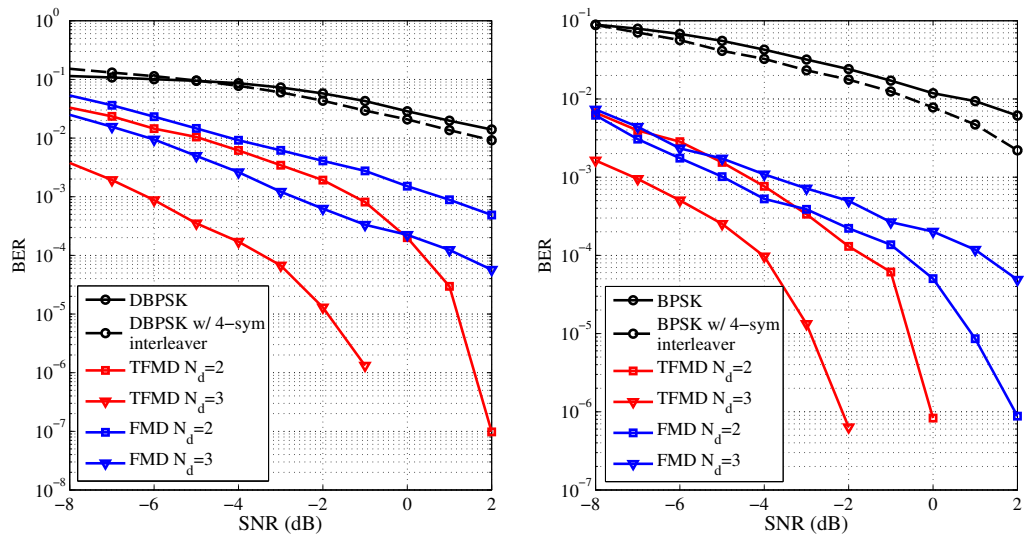


Figure 4.4: Coded BER performance of proposed TFMD transceiver in both non-coherent (left) and coherent (right) systems, using offline noise power estimation. Performance is compared with conventional DBPSK (or BPSK) OFDM systems with bit-level interleaving over 1 or 4 OFDM symbols, and with the frequency modulation diversity (FMD) system.  $N_d$  denotes the length of MD codewords.

Compared to conventional OFDM systems operating at these SNRs, the proposed transceiver reduces the BER by more than 100x and 1000x, respectively. In coherent systems, the SNR required for the proposed transceiver to achieve a BER of  $10^{-4}$  is about -2 dB for the length-two code and -4 dB for the length-three code. In both cases, the reduction in BER compared with the conventional OFDM systems at corresponding SNRs is more than 100x. Notice that increasing the interleaver size in conventional OFDM systems only achieves marginal performance gains in the tested SNR range. Furthermore, comparing TFMD with FMD, we can see that spreading components of a codeword

in different OFDM symbols leads to additional 10x reduction in BER.

#### 4.6.2 Performance with Semi-Online Noise Power Estimation

I now evaluate communication performance of the coherent TFMD transceiver using the proposed semi-online noise power estimator. The iterative receiver is set to run 10 iterations for each packet. For additional comparison, I also simulate the heuristic method for semi-online noise power estimation, as discussed at the beginning of Section 4.5.

The BER performance of the conventional BPSK OFDM system and the TFMD system with various noise power estimation methods is plotted in Figure 4.5. For both the length-two and length-three codes, the heuristic semi-online estimator encounters about 1-2dB loss in SNR compared to the offline estimator, due to error propagation. The performance of the proposed semi-online estimator, however, converges to that of the offline estimator for the length-three code, and is even superior for the length-two code. The latter is primarily because the semi-online estimator, compared to the offline estimator, takes into account any non-periodical fluctuation of noise statistics. Such performance gain is diminishing for longer MD codes, since the diversity demodulator is more robust against inaccurate noise power estimation as the number of diversity branches increases.

## 4.7 Conclusion

In this part of the dissertation, I propose a time-frequency diversity modulation scheme to enhance the communication reliability of narrowband PLC systems in periodic impulsive noise. The time-frequency modulation diversity transmitter jointly encodes multiple bits to multiple PSK symbols, and allocates them to different subcarriers in various OFDM symbols. The receiver linearly combines signals received from corresponding sub-channels / OFDM symbols with weights inversely proportional to the sub-channel SNRs. The periodically varying noise power spectrum can be estimated before or during data transmission using sparse Bayesian learning techniques. I validate the proposed transceiver methods based on noise data collected from narrowband PLC field measurements.

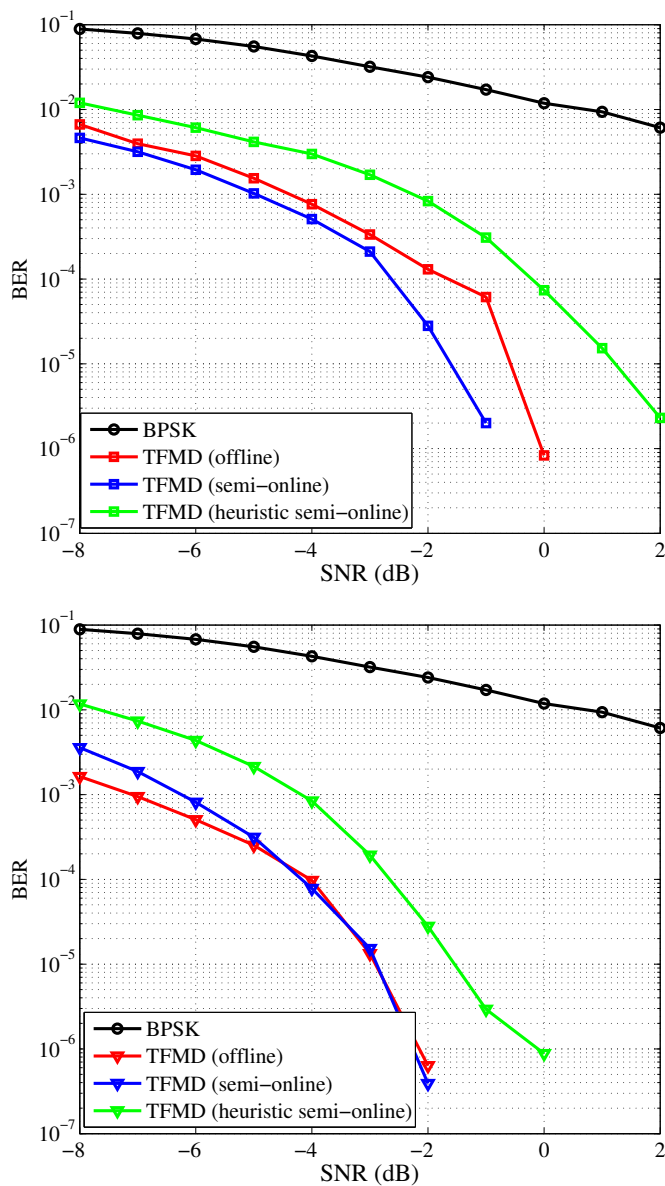


Figure 4.5: Coded BER performance of proposed TFMD transceiver in coherent systems, using semi-online noise power estimation. Both length-two (left) and length-three (right) MD codes are tested. Performance is compared with conventional BPSK OFDM system with bit-level interleaving over 1 OFDM symbol, the TFMD system using offline noise power estimation, and the TFMD system with heuristic semi-online noise power estimation.

# Chapter 5

## Conclusion

In this dissertation, I focus on designing OFDM transmitters and receivers to combat impulsive noise and interference in powerline communications. I propose the following thesis statement:

*Reliability of smart grid communications over power lines can be dramatically improved without sacrificing throughput by exploiting sparsity and cyclostationarity of the impulsive noise in both time and frequency domains.*

### 5.1 Summary of Contributions

I defend this thesis statement via two distinctive perspectives.

#### 5.1.1 Nonparametric mitigation of impulsive noise

In Chapter 3, I develop various receiver algorithms that estimate and subtract impulsive noise from received signal, without prior knowledge on the statistical noise model or model parameters. In coded systems, compared to conventional OFDM receivers without noise mitigation, the proposed receivers provide up to 9 dB and 6.8 dB SNR gains (or alternatively more than 1000x and 100x BER reductions) in asynchronous impulsive noise and periodic im-



pulsive noise, respectively. As such, the proposed receivers enables significant power savings at PLC modems, or alternatively dramatic enhancement of communication reliability. In addition, the algorithms do not impose any training overhead and therefore do not decrease throughput. To achieve these goals, I propose two closely related contributions that tackle asynchronous impulsive noise and periodic impulsive noise, respectively.

1. To mitigate asynchronous impulsive noise, I exploit the sparsity structure of the noise in time domain, and apply sparse Bayesian learning techniques to estimate the noise impulses from (1) null and pilot subcarriers of received OFDM symbols; (2) all subcarriers including data tones; and (3) all subcarriers and decision feedback from the convolutional decoder.
2. For periodic impulsive noise, I observe the cyclostationary property and bursty structure of the noise in time domain, and use a time-domain interleaver that spans about one or two periods of the noise to break long noise bursts into short impulses. I then apply the SBL based methods developed in the previous contribution to estimate the noise impulses after interleaving.

### **5.1.2 Time-frequency modulation diversity to combat periodic impulsive noise.**

In Chapter 4, I propose to use time-frequency diversity modulation to enhance the communication reliability of narrowband PLC systems in periodic impulsive noise. In coded systems, the proposed transceivers achieve more

than 1000x BER reductions compared to conventional OFDM systems, while maintaining the same data rate. Towards that end, my third contribution in this dissertation is comprised of (1) a time-frequency diversity modulation transmitter that jointly encodes multiple bits to a multi-dimensional constellation point, whose components are allocated to time-frequency slots that are separated apart; (2) a linear diversity combining receiver that combines signals received from corresponding time-frequency units with weights inversely proportional to the sub-channel SNRs; and (3) an online algorithm that exploits the frequency-domain sparsity and cyclostationary property of the noise to estimate the noise power spectrum during data transmission.

## 5.2 Suggestions for Future Work

I conclude the dissertation by discussing a variety of open research directions in improving reliability, throughput and energy efficiency of smart grid communications. I will first point out a few potential improvements or extensions based on my current approaches to combat impulsive noise in OFDM-based PLC systems. Finally, I will discuss a novel research direction on PLC-wireless diversity in heterogeneous smart grid communication networks.

- **Nonparametric mitigation of asynchronous bursty impulsive noise.**

The nonparametric methods proposed in Chapter 3 assume *i.i.d.* impulsive noise samples. In bursty impulsive noise, BER performance of these methods will likely be degraded. This is because in some OFDM symbols, the num-

ber of noise impulses might exceed the threshold that guarantees accurate estimation by the SBL algorithms, while in others there is no impulsive noise and hence the SBL algorithms do not provide any performance gain over conventional receivers. To improve the performance of SBL algorithms to recover temporally correlated sparse vectors, one could use impose a Markov random field prior, in addition to the parameterized Gaussian prior, to the sparse vectors. The Markov random field prior is used to capture the temporal correlation between adjacent noise samples.

- **Nonparametric mitigation of asynchronous impulsive noise in non-coherent systems.**

Although derived and evaluated in coherent systems, some of the SBL-based asynchronous impulsive noise mitigation algorithms in Chapter 3 can be extended to work in non-coherent systems as well. In fact, all the methods except for the one using decision feedback can be directly integrated as a pre-processing block into non-coherent systems. The decision feedback estimator does not work in this case since it requires channel estimation, which is unavailable at the non-coherent receivers. It will be interesting to simulate the BER performance of the SBL algorithms using null tones and using all tones in DPSK OFDM receivers, and compare it against that of conventional DPSK OFDM receiver without noise mitigation. One would expect that the SNR gains achieved by the noise mitigation methods be larger than those in coherent systems. As mentioned in Section 2.2, the SNR gap between DPSK and PSK systems is about 3 dB in AWGN, and is

larger than that in impulsive noise [44]. As such, in asynchronous impulsive noise, the SNR gap between conventional OFDM systems with DPSK and PSK modulation schemes is larger than 3 dB. On the other hand, if the noise impulses have been successfully removed, the gap between the SBL receivers with DPSK and PSK modulation schemes should be around 3 dB since the residual background noise is approximately AWGN.

- **Nonparametric mitigation of periodic impulsive noise in non-coherent systems.**

In periodic impulsive noise, the proposed algorithms in Chapter 3 may not be directly applied to non-coherent systems. This is because in non-coherent systems, the use of time-domain block interleaving makes non-coherent demodulation extremely difficult. Instead, one could use a compressed sensing approach similarly to the online noise power spectrum estimator proposed in Chapter 4, which exploits the frequency-domain sparsity of periodic impulsive noise to estimate the harmonic peaks from the CP of received OFDM symbols. Although such approach is able to accurately estimate the noise power spectrum, the estimation of the exact noise samples in frequency domain is difficult, primarily because (1) the number of ISI-free samples in CP is limited; (2) the dictionary matrix involved is highly coherent; and (3) the noise in the frequency domain consists of many harmonic peaks; i.e., it is not very sparse. A possible way to resolve these difficulties is to impose more informative priors on the frequency-domain noise samples. For example, one might use a Markov random field prior to capture the correlation

between the noise amplitudes on adjacent subcarriers.

- **Properties of periodic impulsive noise in other transform domains.**

As mentioned in the previous point, periodic impulsive noise exhibits certain sparsity structure in frequency domain, but is not sparse enough to allow successful recovery by compressed sensing. It will be interesting to transform such noise into other domains, e.g. wavelet domains, and explore the statistical distribution and/or sparsity of the noise in those domains. This could potentially lead to novel statistical models of periodic impulsive noise and corresponding receiver methods that exploit the statistical models or sparsity of the noise in the transform domain.

- **Modulation diversity code design in periodic impulsive noise.**

In Chapter 4, for simplicity, I adopt an MD code that is optimized for Rayleigh fading channel and AWGN. Assuming perfect channel estimation and prior knowledge of the cyclostationary noise statistics, it is possible to design an MD code with minimum symbol error rate (in the average sense) under the specific channel and noise condition. It is very likely that the optimal solution cannot be obtained exactly, and hence sub-optimal methods that allow practical implementation will be interesting.

- **PLC-Wireless Diversity**

As mentioned in Section 1.1, smart grid communications will likely be supported by a variety of network technologies, including wireless and powerline

communications. Wireless and PLC systems were primarily considered as competing candidates for enabling neighborhood area networks and home area networks. Nonetheless, recent study has suggested combining PLC and wireless links to form a multiple-input multiple-output (MIMO) system that exploit the diversity of PLC and wireless channels [37, 48, 49]. In particular, narrowband PLC in the 3–500 kHz band and wireless communications in the unlicensed 920–928 MHz band can be used simultaneously for enabling neighborhood area networks. In [37], the authors assumed flat-fading PLC and wireless channels and made the questionable assumption of a single SNR metric for both the PLC and wireless channels although the noise distributions on those channels are different. This assumption was not made in [49]; however, their investigation assumed in-home PLC transmissions in the 2-30 MHz and wireless transmissions in the 2.4 GHz band which have different channel and interference characteristics from the frequency bands used in neighborhood area networks. Moreover, in [49] the PLC and wireless channel responses were assumed to be deterministic and perfectly known at the receiver and no performance analysis results were presented.

A key observation on PLC-wireless diversity is the independent and non-identical nature of the frequency-selective fading and interference characteristics experienced on both links. One may exploit this observation to enhance transmission reliability and/or increase the coverage range for the smart meter infrastructure using receiver diversity combining schemes. The novelty here, compared to conventional receiver diversity combining

schemes widely used in wireless systems, is due to : (1) different types of non-Gaussian noise dominant in PLC and wireless links (i.e., periodic impulsive noise in PLC vs. asynchronous impulsive noise in wireless); (2) the possibly different modulation schemes used by PLC (e.g. DPSK, PSK, QAM) and wireless (e.g. PSK, QAM or FSK) systems; and (3) the distinct statistical characteristics (e.g. root mean square delay spread, power delay profile, etc.) and fading distributions (e.g. log-normal vs. Rayleigh) of PLC and wireless channels.

One could derive the optimal PLC-wireless receiver combining scheme in the presence of periodic impulsive noise. It will then be interesting to compare its performance and complexity, both analytically and by simulations, with several sub-optimum PLC-wireless receiver diversity combining schemes including maximum-ratio combining, equal-gain combining, selection combining, switch-and-stay combining, and possibly others [92]. The performance comparison might consider practical effects including: (1) errors in estimating the PLC and wireless channels and the noise correlation, and (2) the effects of synchronization errors and data payload size mismatches between the PLC and wireless transmissions.

## Bibliography

- [1] ZigBee standards overview. <https://www.zigbee.org/Standards/Overview.aspx>. [Online].
- [2] IEEE standard for low-frequency (less than 500 kHz) narrowband power line communications for smart grid applications. <http://standards.ieee.org/findstds/standard/1901.2-2013.html>, 2013. [Online].
- [3] Narrowband orthogonal frequency division multiplexing power line communication transceivers for G3-PLC networks. <http://www.itu.int/rec/T-REC-G.9903>, 2013. [Online].
- [4] Power line communication systems market - global assessment and forecast - (2013 - 2018). <http://www.marketsandmarkets.com/Market-Reports/power-line-communication-plc-market-912.html>, 2013. [Online].
- [5] Powerline related intelligent metering evolution (PRIME). [http://www.prime-alliance.org/?page\\_id=769](http://www.prime-alliance.org/?page_id=769), 2013. [Online].
- [6] The return of the cellular smart grid. <http://www.greentechmedia.com/articles/read/the-return-of-the-cellular-smart-grid>, 2013. [Online].
- [7] IEEE smart grid short course, 22 Oct. 2011, Austin, TX, USA.



- [8] U.S. Energy Information Administration. *International Energy Outlook 2013*. Shaker-Verlag, Aachen, 2013.
- [9] 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. Personal Indoor and Mobile Radio Comm.*, 2010.
- [10] A. Aruzuaga, I. Berganza, A. Sendin, M. Sharma, and B. Varadarajan. PRIME interoperability tests and results from field. In *Proc. IEEE Int. Conf. on Smart Grid Comm.*, pages 126–130, 2010.
- [11] S. Aust, R. V. Prasad, and I. Niemegeers. IEEE 802.11 ah: Advantages in standards and further challenges for sub 1 GHz Wi-Fi. In *Proc. IEEE Int. Conf. on Comm.*, pages 6885–6889, 2012.
- [12] A. Batra and J. Balakrishnan. Improvements to the multi-band OFDM physical layer. In *Proc. IEEE Consumer Comm. and Networking Conf.*, volume 2, pages 701–705, 2006.
- [13] J. M. Bohli, C. Sorge, and O. Ugus. A privacy model for smart metering. In *2010 IEEE Int. Conf. on Comm. Workshops.*, pages 1–5, 2010.
- [14] F. Bonnardot, R. B. Randall, and F. Guillet. Extraction of second-order cyclostationary sources - application to vibration analysis. *Journal of Mechanical System and Sig. Proc.*, 19(6):1230–1244, 2005.

- [15] J. Boutros and E. Viterbo. Signal space diversity: a power-and bandwidth-efficient diversity technique for the Rayleigh fading channel. *IEEE Trans. on Info. Theory*, 44(4):1453–1467, 1998.
- [16] G. Caire, T. Y. Al-Naffouri, and A. K. Narayanan. Impulse noise cancellation in OFDM: an application of compressed sensing. In *Proc. IEEE Int. Symp. Info. Theory*, pages 1293–1297, 2008.
- [17] M. H. L. Chan and R. W. Donaldson. Amplitude, width, and interarrival distributions for noise impulses on intrabuilding power line communication networks. *IEEE Trans. Electromagn. Compat.*, 31(3), 1989.
- [18] K-H. Chang and B. Mason. The IEEE 802.15.4g standard for smart metering utility networks. In *Proc. IEEE Int. Conf. on Smart Grid Comm.*, pages 476–480, 2012.
- [19] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159, 2001.
- [20] J. H. Cho. Joint transmitter and receiver optimization in additive cyclostationary noise. *IEEE Trans. on Info Theory*, 50(12), 2004.
- [21] J. A. Cortés, L. Diez, F. J. Canete, and J. J. Sanchez-Martinez. Analysis of the indoor broadband power-line noise scenario. *IEEE Trans. on Electromagn. Compat.*, 52(4):849–858, 2010.
- [22] A. Dabak, I. H. Kim, and T. Pande. Channel modeling of medium to low voltage links for AMI applications of PLC, 2012.

- [23] C. de Fréin, M. Flanagan, and A. Fagan. OFDM narrowband interference estimation using cyclic prefix based algorithm. *RTS*, 6:1, 2006.
- [24] L. Di Bert, 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.*, pages 283–288. IEEE, 2011.
- [25] D. Divsalar and M. K. Simon. Multiple-symbol differential detection of MPSK. *IEEE Trans. on Comm.*, 38(3):300–308, 1990.
- [26] D. Galda and H. Rohling. Narrow band interference reduction in OFDM-based power line communication systems. In *Proc. IEEE Int. Symp. Power Line Comm. and Appl.*, volume 1, pages 4–6, 2001.
- [27] S. Galli and T. Banwell. A novel approach to the modeling of the indoor power line channel-part II: transfer function and its properties. *IEEE Trans. on Power Delivery*, 20(3):1869–1878, 2005.
- [28] 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*, 99(6):998–1027, 2011.
- [29] R. García, L. Díez, J. A. Cortes, and F. J. Canete. Mitigation of cyclic short-time noise in indoor power-line channels. *Proc. IEEE Int. Symp. Power Line Comm. and Its Applications*, pages 396–400, 2007.

- [30] W. A. Gardner and C. M. Spooner. The cumulant theory of cyclostationary time-series. I. foundation. *IEEE Trans. on Sig. Proc.*, 42(12):3387–3408, 1994.
- [31] F. Gianaroli, F. Pancaldi, E. Sironi, M. Vigilante, G.M. Vitetta, and A. Barbieri. Statistical modeling of periodic impulsive noise in indoor power-line channels. *IEEE Trans. on Power Delivery*, 27(3):1276–1283, 2012.
- [32] A. Gomaa, K. M. Zahidul Islam, and Naofal A. Two novel compressed-sensing algorithms for NBI detection in OFDM systems. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Sig. Proc.*, pages 3294–3297. IEEE, 2010.
- [33] I. F. Gorodnitsky and B. D. Rao. Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *IEEE Trans. on Sig. Proc.*, 45(3):600–616, 2002.
- [34] K. Gulati, A. Chopra, R. W. Heath, B. L. Evans, K. R. Tinsley, and X. E. Lin. MIMO receiver design in the presence of radio frequency interference. In *Proc. IEEE Global Comm. Conf.*, pages 1–5, 2008.
- [35] K. Gulati, B. L. Evans, J. G. Andrews, and K. R. Tinsley. Statistics of co-channel interference in a field of poisson and poisson-poisson clustered interferers. *IEEE Trans. on Sig. Proc.*, 58(12):6207–6222, 2010.

- [36] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke. Smart grid technologies: communication technologies and standards. *IEEE Trans. on Industrial Informatics*, 7(4):529–539, 2011.
- [37] S. Güzelgöz, H. B. Celebi, and H. Arslan. Analysis of a multi-channel receiver: Wireless and PLC reception. In *European Sig. Proc. Conf.*, 2010.
- [38] J. Haring. *Error Tolerant Communication over the Compound Channel*. Shaker-Verlag, Aachen, 2002.
- [39] J. Haring and A. J. H. Vinck. Iterative decoding of codes over complex numbers for impulsive noise channels. *IEEE Trans. on Info. Theory*, 49(5):1251–1260, 2003.
- [40] M. Hoch. Comparison of PLC G3 and PRIME. *Proc. IEEE Int. Symp. Power Line Comm. and Its Appl.*, pages 165–169, 2011.
- [41] B. M. Hochwald and T. L. Marzetta. Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading. *IEEE Trans. on Info. Theory*, 46(2):543–564, 2000.
- [42] B. M. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Trans. on Comm.*, 48(12):2041–2052, 2000.
- [43] P. Hoeher and J. Lodge. Turbo DPSK: iterative differential PSK demodulation and channel decoding. *IEEE Trans. on Comm.*, 47(6):837–843, 1999.

- [44] H. Huynh and M. Lecours. Impulsive noise in noncoherent M-ary digital systems. *IEEE Trans. on Comm.*, 23(2):246–252, 1975.
- [45] Texas Instruments. Data sheet for powerline communications analog front-end AFE031. <http://www.ti.com/lit/ds/symlink/afe031.pdf>, 2012. [Online].
- [46] M. Katayama, T. Yamazato, and H. Okada. A mathematical model of noise in narrowband power line communication systems. *IEEE Journal of Selected Areas in Comm.*, 24(7):1267–1276, 2006.
- [47] I. H. Kim, B. Varadarajan, and A. Dabak. Performance analysis and enhancements of narrowband OFDM powerline communication systems. In *Proc. IEEE Int. Conf. on Smart Grid Comm.*, pages 362–367, 2010.
- [48] M. Kuhn, S. Berger, I. Hammerstrom, and A. Wittneben. Power line enhanced cooperative wireless communications. *IEEE Journal on Selected Areas in Comm.*, 24(7):1401–1410, 2006.
- [49] S. W. Lai and G. G. Messier. Using the wireless and PLC channels for diversity. *IEEE Trans. on Comm.*, 60(12):3865–3875, 2012.
- [50] L. Lampe. Bursty impulse noise detection by compressed sensing. *Proc. IEEE Int. Symp. Power Line Comm. and Its Appl.*, pages 29–34, 2011.
- [51] 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 Its Applications*, 2011.

- [52] J. Lin and B. L. Evans. Cyclostationary noise mitigation in narrowband powerline communications. *Proc. APSIPA Annual Summit Conf.*, 2012.
- [53] J. Lin, M. Nassar, and B. L. Evans. Non-parametric impulsive noise mitigation in OFDM systems using sparse Bayesian learning. *Proc. IEEE Global Comm. Conf.*, 2011.
- [54] J. Lin, M. Nassar, and B.L. Evans. Impulsive noise mitigation in powerline communications using sparse Bayesian learning. *IEEE Journal on Selected Areas in Comm.*, (7):1172–1183, 2013.
- [55] 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.*, pages 431–436. IEEE, 2007.
- [56] Y. H. Ma, P. L. So, and E. Gunawan. Performance analysis of OFDM systems for broadband power line communications under impulsive noise and multipath effects. *IEEE Trans. on Power Delivery*, 20(2):674–682, 2005.
- [57] H. Meng, Y. L. Guan, and S. Chen. Modeling and analysis of noise effects on broadband power-line communications. *IEEE Trans. on Power Delivery*, 20(2):630–637, 2005.
- [58] D. Middleton. Non-Gaussian noise models in signal processing for telecommunications: new methods and results for Class A and Class B noise mod-

- els. *IEEE Trans. on Info. Theory*, 45(4):1129–1149, 1999.
- [59] D. Middleton. Statistical-physical models of electromagnetic interference. *IEEE Trans. Electromagn. Compat.*, (3):106–127, 2007.
- [60] J. Mitra and L. Lampe. Convolutionally coded transmission over markov-gaussian channels: Analysis and decoding metrics. *IEEE Trans. on Comm.*, 58(7):1939–1949, 2010.
- [61] A. Moscatelli. Keynote: Technologies evolution for smart grid applications. In *IEEE Int. Symp. Power Line Comm. and Its Appl.*, 2011.
- [62] M. Nassar, A. Dabak, I. H. Kim, T. Pande, and B. L. Evans. Cyclostationary noise modeling in narrowband powerline communication for smart grid applications. *Proc. IEEE Int. Conf. on Acoustics, Speech and Sig. Proc.*, pages 3089–3092, 2012.
- [63] M. Nassar and B.L. Evans. Low complexity EM-based decoding for OFDM systems with impulsive noise. *Proc. IEEE Asilomar Conf. on Sig., Sys. and Computers*, pages 1943–1947, 2011.
- [64] M. Nassar, K. Gulati, M. DeYoung, B. L. Evans, and K. Tinsley. Mitigating near-field interference in laptop embedded wireless transceivers. *Journal of Signal Proc. Systems*, pages 1–12, 2009.
- [65] M. Nassar, K. Gulati, Y. Mortazavi, and B. L. Evans. Statistical modeling of asynchronous impulsive noise in powerline communication networks. *Proc. IEEE Global Comm. Conf.*, 2011.



- [66] M. Nassar, K. Gulati, A. K. Sujeeth, N. Aghasadeghi, B. L. Evans, and K. R. Tinsley. Mitigating near-field interference in laptop embedded wireless transceivers. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Sig. Proc.*, pages 1405–1408, 2008.
- [67] M. Nassar, J. Lin, Y. Mortazavi, A. Dabak, I. H. Kim, and B. L. Evans. Local utility power line communications in the 3-500 kHz band: Channel impairments, noise, and standards. *IEEE Sig. Proc. Mag.*, 29(5):116–127, 2012.
- [68] M. Nassar, P. Schniter, and B. L. Evans. A factor graph approach to joint OFDM channel estimation and decoding in impulsive noise environments. *IEEE Trans. on Signal Proc.*, 2013.
- [69] S. Nayyef, C. Tsimenidis, A. Al-Dweik, B. Sharif, and A. Hazmi. Time- and frequency-domain impulsive noise spreader for OFDM systems. In *IEEE Int. Conf. on Trust, Security and Privacy in Computing and Comm.*, pages 1856–1861. IEEE, 2012.
- [70] 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.
- [71] R. Nilsson, F. Sjöberg, and J. P. LeBlanc. A rank-reduced LMMSE canceller for narrowband interference suppression in OFDM-based systems. *IEEE Trans. on Comm.*, 51(12):2126–2140, 2003.

- [72] V. Oksman and J. Zhang. G. HNEM: the new ITU-T standard on narrowband PLC technology. *IEEE Comm. Mag.*, 49(12):36–44, 2011.
- [73] P. P. Parikh, M. G. Kanabar, and T. S. Sidhu. Opportunities and challenges of wireless communication technologies for smart grid applications. In *IEEE Power and Energy Society General Meeting*, pages 1–7, 2010.
- [74] V. Pauli, L. Lampe, and R. Schober. “Turbo DPSK” using soft multiple-symbol differential sphere decoding. *IEEE Trans. on Info. Theory*, 52(4):1385–1398, 2006.
- [75] P. Prakash. Data concentrators: The core of energy and data management. <http://www.ti.com/lit/wp/spry248/spry248.pdf>, Sept 2013. [Online].
- [76] R. Prasad and C. R. Murthy. Bayesian learning for joint sparse OFDM channel estimation and data detection. In *Proc. IEEE Global Telecomm. Conf.*, pages 1–6, 2010.
- [77] S. Rangan. Generalized approximate message passing for estimation with random linear mixing. In *Proc. IEEE Int. Symp. on Info. Theory.*, pages 2168–2172, 2011.
- [78] D. Raphaeli. Noncoherent coded modulation. *IEEE Trans. on Comm.*, 44(2):172–183, 1996.
- [79] K. Razazian, M. Umari, A. Kamalizad, V. Loginov, and M. Navid. G3-PLC specification for powerline communication: Overview, system simu-

- lation and field trial results. *Proc. IEEE Int. Symp. Power Line Comm. and Its Appl.*, pages 313–318, 2010.
- [80] R. Schober, W. H. Gerstacker, and J. B. Huber. Decision-feedback differential detection of MDPSK for flat Rayleigh fading channels. *IEEE Trans. on Comm.*, 47(7):1025–1035, 1999.
- [81] R. Schober, L. Lampe, W. H. Gerstacker, and S. Pasupathy. Modulation diversity for frequency-selective fading channels. *IEEE Trans. on Info. Theory*, 49(9):2268–2276, 2003.
- [82] C. Studer, P. Kuppinger, G. Pope, and H. Bolcskei. Recovery of sparsely corrupted signals. *IEEE Trans. on Info. Theory*, 58(5):3115–3130, 2012.
- [83] M. E. Tipping. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.
- [84] M. E. Tipping and A. C. Faul. Fast marginal likelihood maximisation for sparse bayesian models. *Proc. Int. Workshop on AI and Stats.*, 2003.
- [85] D. Umehara, M. Kawai, and Y. Morihira. Performance analysis of non-coherent coded modulation for power line communications. *Proc. Int. Symp. Power Line Comm. and Its Appl.*, pages 291–298, 2001.
- [86] B. Varadarajan, I. H. Kim, A. Dabak, D. Rieken, and G. Gregg. Empirical measurements of the low-frequency power-line communications channel in rural North America. *Proc. IEEE Int. Symp. Power Line Comm. and Its Appl.*, pages 463–467, 2011.

- [87] D. P. Wipf and B. D. Rao. Sparse Bayesian learning for basis selection. *IEEE Trans. on Sig. Proc.*, 52(8):2153–2164, 2004.
- [88] D. P. Wipf and B. D. Rao. An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. *IEEE Trans. on Sig. Proc.*, 55(7):3704–3716, 2007.
- [89] Y. G. Yoo and J. H. Cho. Asymptotic analysis of CP-SC-FDE and UW-SC-FDE in additive cyclostationary noise. *Proc. IEEE Int. Conf. Comm.*, pages 1410–1414, 2008.
- [90] S. M. Zabin and H. V. Poor. Efficient estimation of Class A noise parameters via the EM algorithm. *IEEE Trans. on Info. Theory*, 37(1):60–72, 1991.
- [91] Z. Zhang and B. D. Rao. Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning. *IEEE Journal of Selected Topics in Signal Proc.*, 5(5):912–926, 2011.
- [92] L. Zheng and D. N. C. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels. *IEEE Trans. on Info. Theory*, 49(5):1073–1096, 2003.
- [93] S. V. Zhidkov. Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers. *IEEE Trans. on Comm.*, 56(1):5–9, 2008.

- [94] M. Zimmermann and K. Dostert. An analysis of the broadband noise scenario in powerline networks. In *Proc. Int. Symp. Power Line Comm. and Its Applications*, pages 5–7, 2000.
- [95] M. Zimmermann and K. Dostert. An analysis of the broadband noise scenario in powerline networks. *Proc. Int. Symp. Power Line Comm. and Its Applications*, pages 5–7, 2000.
- [96] M. Zimmermann and K. Dostert. Analysis and modeling of impulsive noise in broad-band powerline communications. *IEEE Trans. on Electromagn. Compat.*, 44(1):249–258, 2002.

## Vita

Jing Lin received a B.S. degree in Electrical Engineering in 2008 from Tsinghua University, China. She became a member of the Wireless Networking and Communications Group in Fall 2008 upon joining the University of Texas at Austin, where she obtained an M.S. degree in Electrical and Computer Engineering in 2010. Since then, she has been pursuing a PhD degree at the same university. Her research has been focused on statistical signal processing and machine learning algorithms for noise and interference mitigation in powerline communications.

Permanent address: 3543 Greystone Dr Apt 2070  
Austin, TX 78731

This dissertation was typeset with L<sup>A</sup>T<sub>E</sub>X<sup>†</sup> by the author.

---

<sup>†</sup>L<sup>A</sup>T<sub>E</sub>X is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's T<sub>E</sub>X Program.