DISSERTATION

SENSING, COMMUNICATIONS AND MONITORING FOR THE SMART GRID

Submitted by

Dongliang Duan

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2012

Doctoral Committee:

Advisor: Liuqing Yang

Louis L. Scharf
Jie Luo
Rui Song

ABSTRACT

SENSING, COMMUNICATIONS AND MONITORING FOR THE SMART GRID

With the increasing concern for environmental factors, reliability, and quality of service, power grids in many countries are undergoing revolution towards a more distributed and flexible "smart grid." In the development of the envisioned smart grid, situational awareness takes a fundamental role for a number of crucial advanced operations, such as power flow scheduling, dynamic pricing, energy management, wide area control, wide area protection etc. To fulfill the mission of situational awareness across various entities in the grid, more advanced sensing, communications and monitoring techniques need to be introduced to the existing power grid.

In this research, we will first address the issue of battery power efficiency (BPE) in a wireless sensor network (WSN) which is essential for the sensing system lifetime. We show that the BPE can be improved either by selecting a more battery-power-efficient modulation format or by developing a cooperative communications scheme. Then, to transmit the sensed data over the scarse wireless bandwidth, we adopt cognitive radio as a possible solution. To enable the cognitive radio communication, we aim at improving both the reliability and efficiency of the overall system via cooperative spectrum sensing. With these fundamental communication capabilities available for the sensed data, we then investigate wide area power grid monitoring based on synchronized measurements from newly developed devices such as phasor measurement units (PMUs), mode meters and so on. In addition, an optimal fusion

technique is studied as a good foundation for detection in wireless sensor networks, with application to event detection in the power grid.

ACKNOWLEDGMENTS

First of all, I would like to gratefully and sincerely thank my advisor, Dr. Liuqing Yang, for setting an excellent example as both a good teacher and a good researcher for me. I would like to express my profound gratitude to her for her invaluable support, encouragement and supervision throughout my entire Ph.D. study. Her timely feedback, valuable suggestions and patient guidance have greatly promoted my research progress.

Secondly, I would also like to deeply thank Dr. Louis L. Scharf, who worked closely with me after I transferred to CSU. His scholarly style taught me many invaluable lessons. Furthermore, I have been inspired and obtained fruitful interesting results from our weekly meeting.

Thirdly, I would also like to thank my other two Ph.D. supervisory committee members, Dr. Jie Rocky Luo and Dr. Rui Song, for their time and efforts in serving on my supervisor committee. Valuable discussions with them also helped me very much.

In addition, I also want to thank Dr. Jose C. Principe who helped me as my co-advisor during my study in the University of Florida. My experience within his group, CNEL, has broadened my vision and enriched my research variety a lot.

Last but not least, my thanks go to all members of SCaN group: Dr. Woong Cho, Dr. Huilin Xu, Dr. Rui Cao, Dr. Fengzhong Qu, Dr. Wenshu Zhang, Dr. Ning Wang, Siva Kumar Balaga, Pan Deng, Bo Yu, Julie Cummings, Robert Griffith, Xilin Cheng, Luoyang Fang, Dexin Wang, Jian Dang and Xiaotian Zhou for selflessly

sharing their ideas and inspiring discussions with me. I have benefitted immensely from their friendship and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Smart Grid and Situational Awareness

With the increasing concern for environmental factors, reliability, and quality of service, power grids in many countries are undergoing revolution towards a more economic, stable, distributed and flexible "smart grid." [26] The traditional power system can usually be divided hierarchically into power generation plants, high-voltage transmission network, medium-voltage distribution network and low-voltage customers. The smart grid revolution penetrates all four regimes.

For power generation plants, as the climate is of increasing concern, renewable energy sources are preferable. However, this means increased variety and complexity for the entire generation system. For example, some renewable sources, such as wind power and photovoltaic panels, are innately unstable [3]. In addition, incorporating these various kinds of renewable sources renders the generation much more distributed geographically. The distributed generation structure also unavoidably increases the complexity of the transmission network. Moreover, power deregulation results in different transmission networks operated by different entities and adds further complexity. For the distribution network, even more flexibility is desired with the increasing popularity of an advanced metering infrastructure (AMI) to ensure optimized energy utilization by the customers. On the other hand, the vision to uti-

lize energy storage devices at the customer's side, such as the hybrid electric vehicle (HEV), demands a bi-directional power flow strategy rather than the current uni-directional one. All these ongoing and future smart grid developments make this area an interdisciplinary field of research.

In the development of the envisioned smart grid, situational awareness takes a fundamental role for a number of crucial advanced operations, such as power flow scheduling, dynamic pricing, energy management, wide area control, wide area protection, etc [35]. To fulfill the mission of situational awareness across various entities in the grid, more advanced sensing, communications and monitoring techniques need to be introduced to the existing power grid [23]. A system diagram for obtaining situational awareness for the smart grid is shown in Fig. 1.1.

As illustrated in the figure, sensing is the very first step in obtaining data such as voltage/current instantaneous phasor values, power flow measurements, real-time advanced metering of electric usage, and so on. However, the data will only be meaningful when combined with efficient collection and processing. Therefore, communication of the sensed data is crucial. The quality of service (QoS) indicators for communication within the smart grid measurement system include data rate, delay and so on [66]. When the data is collected in a reliable and timely manner, the power system can be monitored for more advanced operations such as control, protection, pricing and flow scheduling.

Among various possible communication techniques, wireless communications is a promising candidate to fulfill lots of communication tasks in the smart grid due to its

easy implementation and low cost [37]. However, two issues are to be addressed in the application of wireless communications. First, if the sensor nodes, a.k.a. measurement devices, are organized via wireless links into a wireless sensor network (WSN), the energy efficiency is a very important issue since the sensor nodes are usually driven by batteries, which can cause significant environmental hazard. As a result, green communications, a concept of energy-efficient communications, has recently emerged as a hot research area. The second issue is the wireless spectrum. Although many unlicensed techniques such as Wi-Fi, Bluetooth and Zigbees can be directly applied in the smart grid with low cost, more spectrum resources are still required to provide high QoS. With the current exhaustive allocation of the wireless spectrum, assigning dedicated bandwidth for smart grid wireless communications is both unrealistic and costly. Fortunately, most communications within the smart grid in the transmission network are to be deployed mostly in rural areas where plenty of bandwidth is idle. In other words, there is plenty of "spectrum white space" to be exploited. Therefore, the cognitive radio technique can be adopted in such applications to reutilize these spectrum white spaces.

To address these issues in this research, we will first study green communication schemes in WSNs. We show that the battery energy efficiency can be improved either by selecting a more battery-efficient modulation format or by utilizing possible cooperative communication links. Then, to enable cognitive radio functionality for smart grid communications, we aim at improving both the overall system reliability and efficiency via cooperative spectrum sensing. With these fundamental communication

capabilities available for the sensed data, we then investigate wide area power grid monitoring based on synchronized measurements from a newly developed devices such as phasor measurement units (PMUs). In addition, a better sensor fusion technique is developed for potential applications in the smart grid.

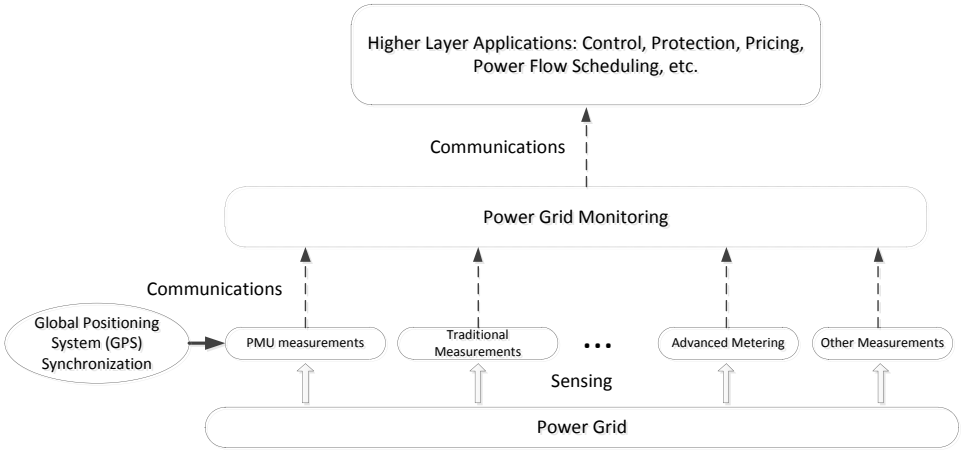


Figure 1.1: System description for gaining situational awareness for the smart grid

1.2 Dissertation Organization

The organization of this dissertation is as follows. We will first discuss green communications for smart grid in Chapter 2. Then, to solve the problem of spectrum scarcity, we will introduce cooperative spectrum sensing techniques in Chapter 4 from a diversity optimization perspective. With these two aspects of wireless communication techniques discussed, we will then study the monitoring of the power grid by state estimation (SE) from the synchrophasor measurement data with possible bad data occurrences in Chapter 5. Then, a general optimal sensor fusion technique is developed in Chapter 6. A summary and discussion of future work will be presented in Chapter 7.

CHAPTER 2

GREEN WIRELESS COMMUNICATIONS FOR THE SMART GRID: OPTIMIZING THE BATTERY EFFICIENCY

2.1 Motivation

To ensure better situational awareness, sensors are being widely installed in the smart grid for better power grid monitoring [20]. However, while the trend of the smart grid is to turn green, the wireless communications among the installed sensors can cause even more environmental damages since the sensor nodes are typically driven by nonrenewable batteries [2]. Battery energy-efficient communication techniques are desired to make wireless communications green. There are quite a few studies of battery efficiency for wireless communications in sensor networks (see e.g. [11, 29]). However, the analyses in most existing literature adopt an ideal model without taking into account the extra power consumption due to circuit operations and battery nonlinearities. In recent years, there has been an increasing trend of incorporating special battery characteristics into network protocol design and optimization (see e.g. [57, 45, 13, 68]) with the introduction of a more realistic sensor node model. To optimize battery energy efficiency, we first select a more battery-efficient modulation scheme for WSNs and then, for further optimization, we introduce the problem of selecting between single-hop direct-link communication and multi-hop cooperative communication utilizing intermediate relaying nodes.

To select a more energy-conserving modulation, the comparisons between energy efficiencies of modulation schemes have been well studied (see e.g. [44]). However, these comparisons do not take into account the circuit operations of the modulation schemes and the extra battery energy loss due to battery nonlinearities. To illustrate the effects of these two factors, we present a case study of the battery energy consumption of two modulation schemes, namely frequency shift keying (FSK) modulation and pulse-position modulation (PPM), which are considered to have identical energy efficiency. Results show that due to the differences of their circuit operation time and pulse energy distribution, these two modulation schemes consume different average battery energy, and the selection criterion depends on internode distances in the network.

To further improve the network battery energy efficiency, it is well known that when relaying is utilized to split the direct transmission from the source to the destination into two or more hops, the total battery energy consumption is expected to be greatly reduced, since the transceiver distances are smaller than those of the direct link and the path loss is thus significantly reduced [28]. However, when circuit energy consumptions are considered, it is shown that relaying does not always save energy (see e.g. [53, 52]). This gives rise to an intriguing relay selection problem. However, considerations of this problem in [53] and [52] use very limiting assumptions such as linear relay node placement, identical and fixed transmission energy at the source, relay nodes without energy allocation optimization, and an ideal linear battery model. Here, we adopt a more general setup by assuming arbitrary relay node placement with

a more realistic nonlinear battery model. With this problem formulation, closed-form expressions are derived for relay selection to maximize battery efficiency. Results show that the relay selection criterion depends on both the transmission distance and the relative position of the relay node with respect to the primary nodes. In addition, numerical results are presented to show the battery power efficiency improvement at different candidate relay locations and for various scenarios.

This chapter is organized as follows: the system model together with the average battery energy consumption analysis for a single pulse transmission are presented in Section 6.2. With this analysis, the battery energy consumption comparison of M-FSK and M-PPM is presented in Section 2.3 and the relay selection problem is solved in Section 2.4. Finally, summarizing remarks are given in Section 2.5.

2.2 System Battery Energy Consumption Model

2.2.1 Node Circuit Operation and Battery Nonlinearity

To capture the actual battery energy consumption at sensor nodes, circuit power consumptions are taken into account with $\mathcal{P}_{ct}$ and $\mathcal{P}_{cr}$ denoting transmitter and receiver circuit power consumption, respectively. Also, the inefficiency of the DC/DC convertor in the node circuit is denoted by a factor $\eta < 1$ and the imperfections of a power amplifier (PA) are described by an extra power loss factor $\alpha > 0$.

Moreover, the real battery discharge process is nonlinear. As introduced in [57], the nonlinear behavior of the battery discharge process can be captured by $\mathcal{P}_0 = \int_{I_{\min}}^{I_{\max}} \frac{Vi}{1-\omega i} f(i)di$, where $\mathcal{P}_0$ is the average power consumption of the battery over a

battery discharge process, V is the battery voltage, $f(i)$ is the density function of the battery discharge current profile during the time period of interest $[t_{\min}, t_{max}]$, $1 - \omega i$ with $0 < \omega \ll 1$ is the *battery efficiency factor* [38] and $I_{\max}$ and $I_{\min}$ are respectively the maximum and minimum affordable discharge currents. To facilitate the ensuing analysis, we define the *instantaneous power consumption* at time t as $\mathcal{P}_0(t) = Vi(t)/(1 - \omega i(t)$. Then, the average power consumption of the battery over the discharge interval $[t_{\min}, t_{\max}]$ can be alternatively expressed as:

$$\mathcal{P}_0 = \int_{t_{\min}}^{t_{\max}} \mathcal{P}_0(t)dt = \int_{t_{\min}}^{t_{\max}} \frac{Vi(t)}{1 - \omega i(t)}dt \ . \tag{2.1}$$

2.2.2 Channel Model

The channel considered here is a path-loss Rayleigh fading channel with additive white Gaussian noise (AWGN). The channel gain factor $G(d)$ depends on the transceiver distance d and is given by [48, Chapter 4]: $G(d) = \mathcal{P}_s/\mathcal{P}_r = M_l G_1 d^K$, where $\mathcal{P}_s$ and $\mathcal{P}_r$ are the transmitted and received power of the signal, and the remaining parameters are defined in Table 2.1. Accordingly, the relationship between the average energy at the transmitter $\mathcal{E}$ and the average energy at the receiver $\mathcal{E}_r$ is:

$$\mathcal{E}/\mathcal{E}_r = \mathcal{P}_s/\mathcal{P}_r = G(d) = M_l G_1 d^K \ . \tag{2.2}$$

2.2.3 The Average Battery Energy Consumption

As a preliminary, we will first analyze the battery energy consumption for a single transmitted pulse.

Table 2.1: Notations

M_l	channel link margin
G_1	gain factor at $d = 1$
K	path-loss exponent
$\mu(i)$	battery efficiency factor $\mu(i) = 1 - \omega i$
η	transfer efficiency of the DC/DC converter
α	extra power loss factor of the PA
$p(t)$	transmitted pulse
γ_p	$\int_0^{T_p} \left(\frac{p(t)}{\int_0^{T_p} \lvert p(t) \rvert dt} \right)^2 dt$
$\mathcal{E}_p$	pulse energy
$\mathcal{E}_0$	average battery energy consumption
$\mathcal{P}_{ct}$	transmitter circuit power
$\mathcal{P}_{cr}$	receiver circuit power

C.1: Transmitter Battery Energy Consumption

With our realistic circuit and battery model, the actual battery energy consumption for transmitting a single pulse can be obtained as the following theorem:

Thoerem 2.1 *The total battery energy consumption for transmitting a single pulse $p(t)$ with duration T_p and energy $\mathcal{E}_p$ is approximately:*

$$\mathcal{E}_{0t} = \frac{\omega \gamma_p (1+\alpha)^2}{V \eta^2} \mathcal{E}_p^2 + \frac{1+\alpha}{\eta} \mathcal{E}_p + \frac{\mathcal{P}_{ct}}{\eta} T_p \ , \tag{2.3}$$

with parameters defined in Table 2.1.

Proof. See Section 2.6 ■

In (2.3), the η and α terms reflect the influence of the inefficiency of the DC/DC converter and the extra PA power loss, respectively. The result in Theorem 2.1 shows that the total battery energy consumption can be decomposed into three parts:

1. The first term in (2.3) refers to the excess power loss due to the nonlinear battery discharge process. This term is proportional to the square of the energy of the transmitted signal. In addition, this term is only affected by the pulse shape through a scaling factor γ_p. Notice that, though γ_p can be interpreted as the "energy" of $p_0(t)$, it is *not* the actual energy consumption of a DC battery with a constant voltage V. Instead, this scaling factor γ_p is the result of normalization of the un-amplified pulse waveform $p_0(t)$ in order to ensure a constant *pure* (without any circuit energy consumption) and *ideal* (without any battery nonlinearity) battery energy consumption independent of the actual shape the pulse takes. Finally, this term also depends on the battery parameter ω, which captures the nonlinear feature of the battery;

2. The second term in (2.3) refers to the energy carried by the transmitted signal. it would be exactly the energy of the transmitted pulse if there were not effects of the DC/DC converter (via η) and the PA (via α);

3. The third term in (2.3) refers to the circuit energy consumption. It depends on the power of the circuit and the pulse duration T_p.

C.2: Receiver Battery Energy Consumption

At the receiving node, there is no PA but a low noise amplifier (LNA) with nearly constant power consumption. Thus, the current $I_r = \mathcal{P}_{cr}/(\eta V)$ where $\mathcal{P}_{cr}$ is the circuit power consumption at the receiver. In general, I_r is very small, so $\mu(I_r) = 1 - \omega I_r \approx \mu_{\max} = 1$ [57]. The circuit of the receiver needs to be turned on for the demodulation

duration T_d. Hence, the total battery energy consumption of the receiving node is

$$\mathcal{E}_{0r} = \frac{\mathcal{P}_{cr}}{\eta} T_d \, . \tag{2.4}$$

Then, the total battery energy consumption can be obtained as the summation of the transmitter and receiver battery energy consumptions.

2.2.4 Performance Criterion

In the analysis of battery energy consumption, average bit-error-rate (BER) is adopted as the performance metric. Thus, in all the comparisons, the average battery energy consumptions to achieve the same target BER performance ($\bar{P}_e$) are compared for both the modulation selection and the communication link selection problems.

2.3 Modulation Selection: A Case Study

In this section, we will establish a framework for choosing the more battery energy efficient modulation in wireless sensor networks, and will present a case study of modulation selection between PPM and FSK to show the effects of circuit power consumption and battery nonlinearities on the battery energy consumption, and obtain the selection criterion expressed as a closed-form function of the internode distance.

With $\triangle f = 1/T_s$, M-FSK and M-PPM have identical bandwidth occupancy and bandwidth efficiency as shown in [57]. However, when the realistic system model is considered, these two schemes are found to have *different* battery energy consumption performances.

For M-FSK, the signal pulse transmitted for symbol $m \in \{0, 1, \ldots, M-1\}$ is $p_m^F(t) = \sin(2\pi(f_c + (2m-1)\triangle f)t), t \in [0, T_s]$, where f_c is the carrier frequency and $\triangle f$ is the frequency spacing.

For M-PPM, all possible signals utilize the same pulse shaper and the information is conveyed by signal pulse position. To compare with M-FSK, pass-band M-PPM is considered. Then, the basic pulse is $p^P(t) = \sin 2\pi f_c t$ with $t \in [0, T_s/M]$, where f_c is the carrier frequency. Correspondingly, the pulse transmitted for symbol m is $p_m^P(t) = p^P\left(t - m\frac{T_s}{M}\right)$, $\frac{mT_s}{M} < t \leq \frac{(m+1)T_s}{M}$.

To calculate the average battery energy consumptions for these two modulation schemes, the pulse shape factors γ_p^F and γ_p^P need to be obtained in the first place. According to Theorem 2.1 and Table 2.1, for M-FSK: $\gamma_{p,m}^F = \frac{\int_0^{T_s}[\sin(2\pi(f_c+(2m-1)\triangle f)t)]^2 dt}{\left[\int_0^{T_s}|\sin[2\pi(f_c+(2m-1)\triangle f)t]|dt\right]^2} = \frac{T_s}{2}/\left(\frac{T_s}{2\pi}\right)^2 = \frac{2\pi^2}{T_s}$. Similarly, for M-PPM, $\gamma_{p,m}^P = \frac{2\pi^2}{T_s/M} = M\frac{2\pi^2}{T_s} = M\gamma_{p,m}^F$.

M-FSK and M-PPM have the same required energy at the receiver, i.e., $\mathcal{E}_{pr}^F = \mathcal{E}_{pr}^P = \mathcal{E}_{sr}$, where $\mathcal{E}_{sr}$ is the required average symbol energy to obtain the target BER performance $\bar{P}_e$. In Rayleigh fading channels, this value can be obtained from the formula given in [57]. Then, according to our path-loss channel model describe in (2.2), the transmitted pulse energies are $\mathcal{E}_p^F = \mathcal{E}_p^P = M_l G_1 \mathcal{E}_{sr} d^K$. At the receiver demodulator, the circuit works the same amount of time to detect received signals. Thus, the receiver circuit energy consumption for both M-FSK and M-PPM is $\mathcal{P}_{cr} T_s$. Substituting all these parameters into (2.3) and (2.4), the difference between the

average battery energy consumptions of M-FSK and M-PPM is obtained as:

$$\begin{aligned}\Delta\mathcal{E}_0^{FP} &= (1-M)\frac{2\pi^2 M_l^2 G_1^2 \omega (1+\alpha)^2}{T_s V \eta^2 \log_2 M}\mathcal{E}_{sr}^2 d^{2K} + \frac{M-1}{M^2}\frac{\mathcal{P}_{ct} T_s}{\eta \log_2 M} \\ &= k_2 d^{2K} + k_0 \ ,\end{aligned} \tag{2.5}$$

where $k_2 = (1-M)\frac{2\pi^2 M_l^2 G_1^2 \omega(1+\alpha)^2}{T_s V \eta^2 \log_2 M}\mathcal{E}_{sr}^2 < 0$ and $k_0 = \frac{M-1}{M^2}\frac{\mathcal{P}_{ct} T_s}{\eta \log_2 M} > 0$. Thus,

Proposition 2.1 *There is a critical distance* $d_c = \left(-\frac{k_0}{k_2}\right)^{\frac{1}{2K}}$ *such that when the internode transmission distance* $d < d_c$*, M-PPM consumes less battery energy than M-FSK and vice versa.*

The critical distance, and the advantageous operation region for both FSK and PPM, is plotted in Fig. 2.1 with system parameters given in Table 2.2.

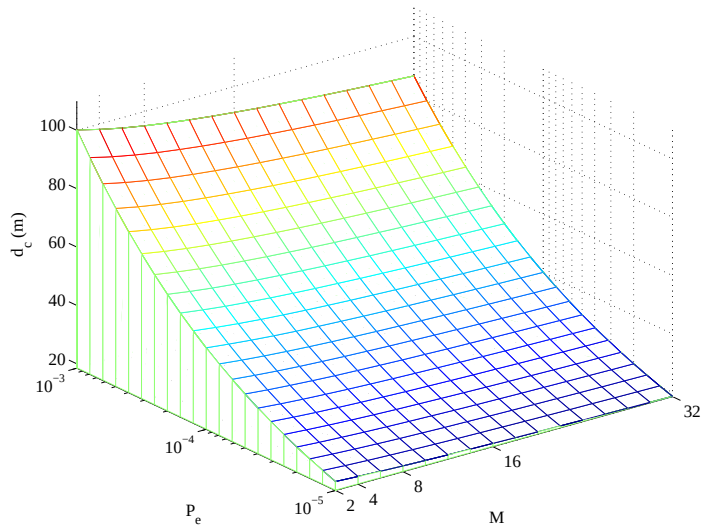


Figure 2.1: Comparison results of M-FSK and M-PPM under Rayleigh fading channel as a function of modulation size M and BER requirement. Below the surface: PPM-advantageous region; above the surface: FSK-advantageous region.

The comparison between M-PPM and M-FSK presented in this section shows that the slight nonlinearity of the battery is actually *not* negligible. If no battery nonlin-

Table 2.2: System Parameters

$\omega = 0.05$	$T_p = 1.33 \times 10^{-4}$s	$\alpha = 0.33$	$\mu_{\min} = 0.5$
$K = 3$	$N_0/2 = -171$dBm/Hz	$G_1 = 27$dB	$M_l = 40$dB
$V = 3.7$V	$\mathcal{P}_{cr} = 52.5$mW	$\mathcal{P}_{ct} = 105.8$mW	$\eta = 0.8$

earity were considered, then we would expect that M-PPM is always preferred since it costs less transmitter circuit power consumption. However, the comparison result shows that M-FSK is preferred when internode transmission distance is sufficiently large since M-FSK has a lower pulse amplitude and thus higher battery discharge efficiency.

2.4 Communication Link Selection: To Relay or Not To Relay?

2.4.1 Single Relay Node Cooperation

In this work, we only consider the case that a single relay node is utilized to assist communication. The primary communication nodes are the source node S_1 and the destination node S_2. Then a relay node R exists as shown in Fig. 2.2. The relay node R can lie anywhere in the two-dimensional space. There are two candidate links for the end-to-end communication from S_1 to S_2, namely the single-hop *direct link* S_1–S_2 and the *two-hop relay link* S_1–R–S_2. We denote the distance of S_1–S_2 as d, the distance of S_1–R as $d_1 = \theta_1 d$, and R–S_2 as $d_2 = \theta_2 d$. Straightforwardly, θ_1 and θ_2 satisfy conditions $\theta_1, \theta_2 > 0$ and $\theta_1 + \theta_2 \geq 1$. The linear relay node placement is a special case with $\theta_1 + \theta_2 = 1$.

Without loss of generality, we adopt a simpler modulation scheme, namely base-band binary phase-shift keying (BPSK) for analysis simplicity. Under the Rayleigh

fading channel with coherent detection at the receiver, the average BER at high signal-to-noise ratio (SNR) is approximated as [59, Chapter 3]:

$$P_e = \frac{N_0}{4\mathcal{E}_r} \, , \tag{2.6}$$

where $\mathcal{E}_r$ is the signal energy at the receiver. In addition, for the relaying link, a Decode and Forward (DF) relaying protocol is considered first. It is shown later that DF and Amplify and Forward (AF) protocols are identical at high SNR.

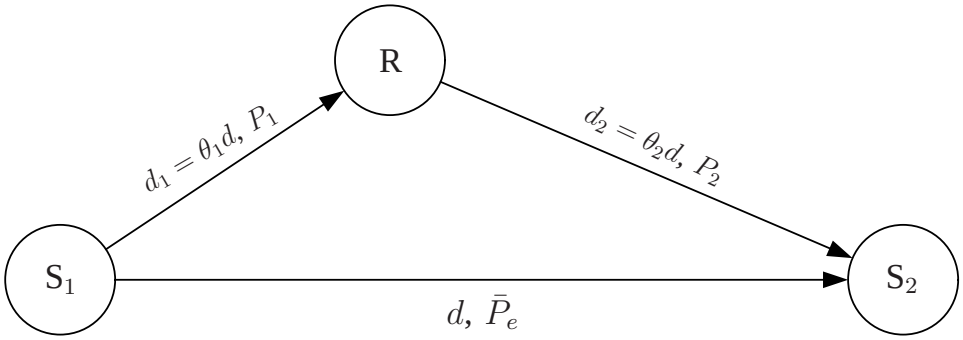


Figure 2.2: Communication with possible single relay node assistance.

2.4.2 Total Battery Energy Consumption for Direct Transmission

Considering baseband BPSK modulation over the Rayleigh fading channel, $\gamma_p^B = \frac{1}{T_p}$ according to Table 2.1. Substituting (2.6) into (2.3) and (2.4), the total battery energy consumption of the direct transmission S_1–S_2, denoted as $\mathcal{E}_D$, can be expressed as an explicit function of the direct-link transmission distance d and the target BER $\bar{P}_e$:

$$\mathcal{E}_D = l_2 \frac{d^{2K}}{\bar{P}_e^2} + l_1 \frac{d^K}{\bar{P}_e} + l_0 \, , \tag{2.7}$$

where $l_2 = \frac{M_l^2 G_1^2 \omega (1+\alpha)^2 N_0^2}{16 T_p V \eta^2}$, $l_1 = \frac{M_l G_1 (1+\alpha) N_0}{4\eta}$, and $l_0 = \frac{(\mathcal{P}_{ct}+\mathcal{P}_{cr})T_p}{\eta}$. Notice that $\mathcal{E}_{\mathrm{D}}$ can be regarded either as a quadratic function of $\bar{P}_e^{-1}$ for a given S_1-S_2 distance d, or as a quadratic function of d^K with some desired $\bar{P}_e$.

2.4.3 Total Battery Energy Consumption for Relaying Transmission

When the relay link is deployed using the DF protocol, the relay node R first demodulates the signal from the source and then forwards the remodulated signal to the destination. Therefore, the relay link S_1–R–S_2 can essentially be separated into two decoupled links S_1–R_{D} and R_{F}–S_2, each having received pulse energy $\mathcal{E}_{pr,i}$ and transmission distance $d_i = \theta_i d$, for $i = 1, 2$. As a result, the total battery energy consumption for the S_1–R–S_2 link can be simply expressed as the summation of the energy consumption of the two decoupled links.

Due to the inverse relationship between the received pulse energy and the average BER given in (2.6), the energy distribution among the source node and the relay node is exclusively determined by the average BER of the two decoupled links, denoted by P_1 and P_2 respectively. The overall average BER for the relay link can be tightly upper bounded using P_1 and P_2 as $\bar{P}_e \leq 1 - (1 - P_1)(1 - P_2) = P_1 + P_2 - P_1 P_2$. In practice, the desired BER is usually small ($\bar{P}_e \leq 10^{-3}$), which means $P_1 + P_2 \gg P_1 P_2$ and therefore $\bar{P}_e \approx P_1 + P_2$. The latter will be considered as the error performance constraint for the relaying transmission in the sequel, and the total battery energy

consumption can be rewritten in terms of P_1 as

$$\mathcal{E}_{\mathrm{R}}(P_1)=l_2\left(\frac{d_1^{2K}}{P_1^2}+\frac{d_2^{2K}}{(\bar{P}_e-P_1)^2}\right)+l_1\left(\frac{d_1^K}{P_1}+\frac{d_2^K}{\bar{P}_e-P_1}\right)+2l_0\ , \tag{2.8}$$

with constraint $0 < P_1 < \bar{P}_e$. Obviously, in this expression it is seen that with relaying, an extra l_0 battery energy consumption term is introduced which is caused by the extra circuit energy consumption of the relay node.

Now the energy optimization problem is introduced: given the relay location d, the relative position of the relay node with respect to the primary nodes, and the BER fixed at $\bar{P}_e$, an optimal energy allocation strategy should be adopted to obtain the S_1–R link BER performance P_1 to minimize the total battery energy consumption $\mathcal{E}_{\mathrm{R}}(P_1)$; that is,

$$\mathcal{E}_{\mathrm{R}}=\min_{P_1}\ \mathcal{E}_{\mathrm{R}}(P_1),\ \text{subject to}\ 0<P_1<\bar{P}_e\ , \tag{2.9}$$

where $\mathcal{E}_{\mathrm{R}}(P_1)$ is given by (2.8).

It turns out that the first-order term in the total battery energy consumption $\mathcal{E}_{\mathrm{R}}(P_1)$ formula for the relay link in (2.8) dominates the total energy consumption within the BER range of $0 < P_1 < \bar{P}_e$. Therefore, second-order terms can be temporarily discarded during the energy allocation optimization process with negligible effect on the optimality.

After removal of the second-order term, the first-order total battery energy consumption becomes

$$\mathcal{E}_{\mathrm{R}}^{1}(P_1) = L_1 \left(\frac{d_1^K}{P_1} + \frac{d_2^K}{\bar{P}_e - P_1} \right) + 2L_0 \ . \tag{2.10}$$

To find the P_1 for minimization, taking the derivative of $\mathcal{E}_{\mathrm{R}}^{1}(P_1)$ with respect to P_1 and setting it to zero, it is obtained that:

$$\left(d_2^K - d_1^K\right) P_1^2 + 2d_1^K \bar{P}_e P_1 - d_1^K \bar{P}_e^2 = 0 \ , \tag{2.11}$$

which is a quadratic equation in P_1. Note that the root of this equation only depends on the distance ratio $d_2/d_1 = \theta_2/\theta_1$ and the overall target BER $\bar{P}_e$. It has nothing to do with the battery consumption formula coefficients L_1 and L_0. The quadratic equation has well-established simple root expressions. Within the range of $0 < P_1 < \bar{P}_e$, (2.11) always has a single positive root, and the suboptimal P_1 can be obtained as:

$$\begin{aligned} P_1^{\mathrm{so}} &= \frac{1}{(\theta_2/\theta_1)^{\frac{K}{2}} + 1} \bar{P}_e \ , \\ \text{and} \quad P_2^{\mathrm{so}} &= \bar{P}_e - P_1^{\mathrm{so}} = \frac{(\theta_2/\theta_1)^{\frac{K}{2}}}{(\theta_2/\theta_1)^{\frac{K}{2}} + 1} \bar{P}_e \ . \end{aligned}$$

It is shown in [68] that the total energy consumption with this suboptimal P_1^{so} approaches that with the theoretical optimal solution to Eq. (2.9). Substituting P_1^{so} back into (2.8), the total battery energy consumption $\mathcal{E}_{\mathrm{R}}$ for the relay link S_1–R–S_2

can be obtained as:

$$\begin{aligned}\mathcal{E}_{\mathrm{R}} \;=\; & \frac{L_2}{\bar{P}_e^2}\left(\theta_1^{\frac{K}{2}}+\theta_2^{\frac{K}{2}}\right)^2\left(\theta_1^K+\theta_2^K\right)d^{2K} \\ & +\frac{L_1}{\bar{P}_e}\left(\theta_1^{\frac{K}{2}}+\theta_2^{\frac{K}{2}}\right)^2 d^K+2L_0\;. \end{aligned} \tag{2.12}$$

With this result, the next question to be addressed is: in order to achieve higher battery energy efficiency, under what conditions, should the relay link S_1–R–S_2 be chosen rather than the direct link S_1–S_2.

2.4.4 Relay Selection Criterion

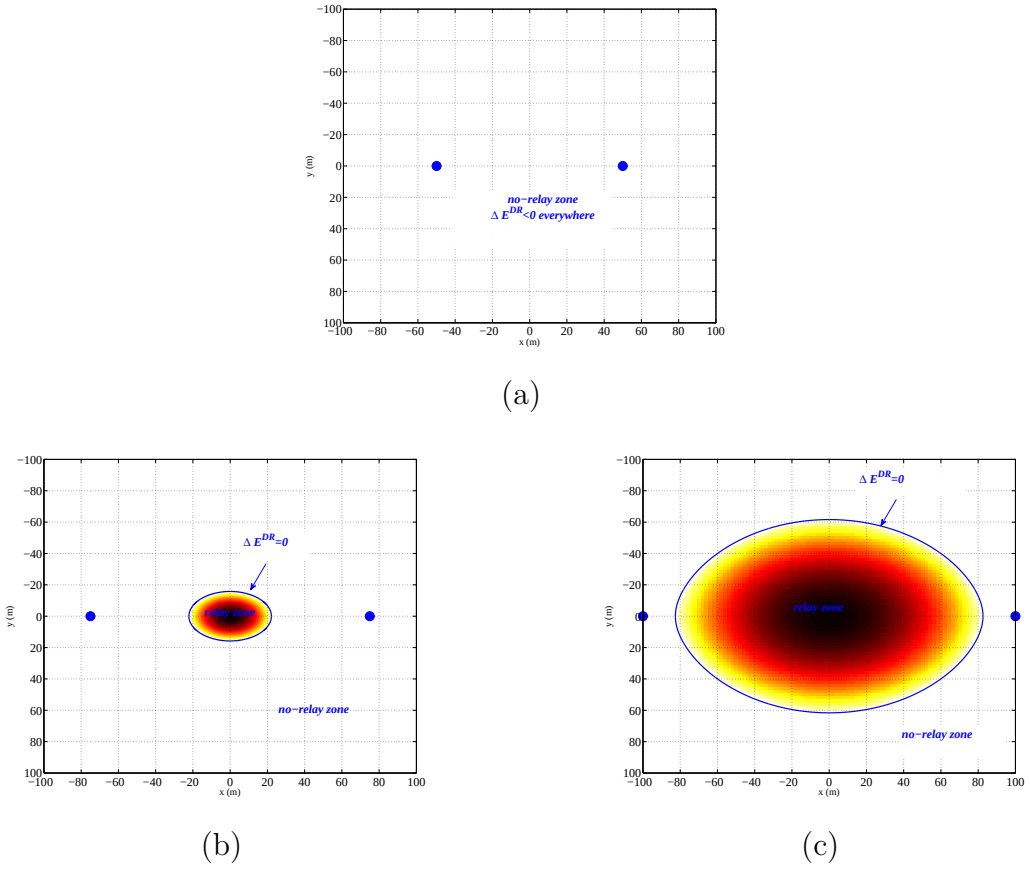
To compare the direct and relaying transmissions in terms of the battery energy efficiency, the difference of the total battery energy consumptions between the direct and the optimal relaying transmissions, denoted as $\Delta\mathcal{E}^{DR}$, is evaluated. The sign of the difference will reveal the more efficient one: direct link outperforms the relay link if the difference is negative, and vice versa. $\Delta\mathcal{E}^{DR}$ can be readily obtained by taking the difference of (2.7) and (2.12):

$$\begin{aligned}\Delta\mathcal{E}^{DR}(\theta_1,\theta_2) \;=\; & \mathcal{E}_{\mathrm{D}}-\mathcal{E}_{\mathrm{R}} \\ \;=\; & \frac{L_2}{\bar{P}_e^2}\left(1-\left(\theta_1^{\frac{K}{2}}+\theta_2^{\frac{K}{2}}\right)^2\left(\theta_1^K+\theta_2^K\right)\right)\left(d^K\right)^2 \\ & +\frac{L_1}{\bar{P}_e}\left(1-\left(\theta_1^{\frac{K}{2}}+\theta_2^{\frac{K}{2}}\right)^2\right)d^K-L_0\;, \end{aligned} \tag{2.13}$$

which is a quadratic function of d^K. The constant term is negative. This results from the fact that the relaying transmission consumes more distance-independent

transceiver circuit energy than the direct transmission. As a result, the relay selection criterion can be obtained as following:

Proposition 2.2 *For a relay node with distances $\theta_1 d$ and $\theta_2 d$ apart from the primary nodes, choose the direct link if $\Delta\mathcal{E}^{DR}(\theta_1, \theta_2) < 0$; otherwise, choose the relay link.*



(a)

(b) (c)

Figure 2.3: The position of the relay node: inside the curve, choose relay-link communication; outside the curve, choose direct-link communication. The solid dots are the primary nodes. (a) $d = 100m$. (b) $d = 150m$. (c) $d = 200m$.

The numerical results for primary nodes located at $(-d/2, 0)$ and $(d/2, 0)$ using the system parameters in Table 2.2 are presented in Fig. 2.3. These figures show the relay and no-relay zones. Also, we can see that the nearer the relay node is to the center of the primary nodes, the more battery energy can be saved (darker as shown in the

figure). Thus, within the curve $\Delta\mathcal{E}^{DR}(\theta_1, \theta_2) = 0$ is the battery-energy saving area. Comparing Figs. 2.3.(a)-(c), we can see that as the distance between the primary nodes increases, the battery-energy saving area enlarges. For example, when $d = 100m$, a relay node cannot reduce the total battery energy consumption; while when $d = 200m$, nearly any relay node between the primary nodes can be utilized for better battery energy efficiency. This is because while the extra circuit consumption caused by relaying is a distance-unrelated constant, relaying can reduce energy consumption for combatting the channel fading in longer-distance communications.

Interestingly, the energy-saving area can be approximated by an ellipse as shown in Fig. 2.4. The expression for the ellipse can be written as: $\frac{x^2}{x_{\max}^2} + \frac{y^2}{y_{\max}^2} = 1$, where the parameters $x_{\max}$ and $y_{\max}$ can be calculated by taking two special points on the curve with $\theta_1 = 1 - \theta_2$ and $\theta_1 = \theta_2$, respectively. Thus, in real applications, one can just check whether there are any relay nodes within the elliptical area to utilize cooperative communication for better battery energy efficiency.

2.4.5 AF and DF Equivalence in Our Scenario

For the AF protocol, the relay amplifies what it receives from the source, including the noise, and forwards it to the destination. With two links S_1–R and R–S_2, the AF relaying transmission follows the same total battery energy consumption formula as the DF case as illustrated in the following: Denote the SNR for the two links as $\rho_1 = \mathcal{E}_{pr1}/N_0$ and $\rho_2 = \mathcal{E}_{pr2}/N_0$. Using the AF protocol, the received signal at the destination essentially involves two parts: useful signal $\mathcal{E}_{pr2} \cdot \rho_1/(\rho_1 + 1)$ and noise $\mathcal{E}_{pr2}/(\rho_1 + 1)$ magnified at the relay node. Therefore, the overall SNR for the relay

link S_1–R–S_2 is $\rho = (\mathcal{E}_{pr2} \cdot \rho_1/(\rho_1+1))/(N_0 + \mathcal{E}_{pr2}/(\rho_1+1))$. Taking the reciprocal of ρ, it is obtained that $1/\rho = (1/\rho_2) \cdot ((\rho_1+1)/\rho_1) + 1/\rho_1 \approx 1/\rho_1 + 1/\rho_2$, where the approximation comes from the fact that the practical desired SNR is small and thus only the medium to high SNR, i.e. $\rho_1 \gg 1$, is considered. Due to the relationship in (2.6), the average BER constraint for the AF relaying transmission is obtained as follows: $\bar{P}_e = P_1 + P_2$, which is identical to that in the DF case. Since the AF relaying transmission shares the same battery energy consumption formula and BER performance constraint with the DF transmission, all the analyses for DF also hold for AF. In other words, the two relaying protocols are equivalent in terms of the relay selection over Rayleigh fading channels at high SNR.

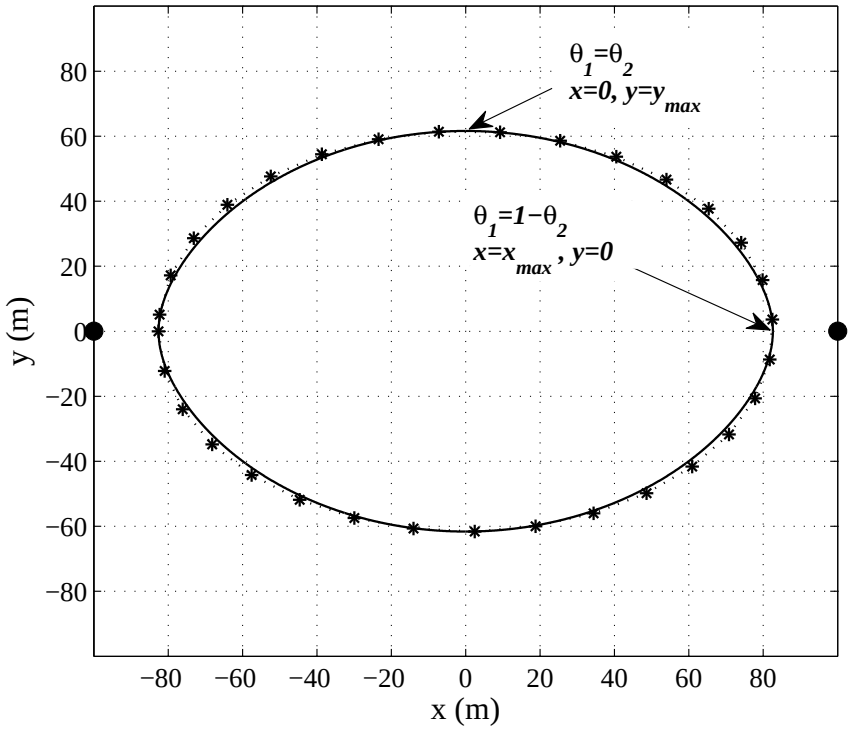


Figure 2.4: The ellipse to approximate the energy-saving area for $d = 200m$. Solid curve: the energy-saving area calculated by numerical computing; dashed curve with stars: the elliptical approximation.

2.5 Conclusions

In this chapter, to ensure green wireless communication in the smart grid, we introduced a realistic node model to optimize the system energy efficiency in wireless sensor networks (WSN), by taking into account both circuit power consumption and battery nonlinearities. Under this model, we first chose a more battery-energy-efficient modulation scheme for the entire network and then, we selected either single-hop direct-link or multi-hop relay-link communication to further reduce the total battery energy consumption. For modulation selection, new understandings on the battery energy efficiency of modulation schemes were facilitated by our case study of M-FSK and M-PPM, which are traditionally considered as identically energy efficient, and it was found that the selection criterion relies on the internode transmission distance of the wireless network. For communication link selection, the multi-hop transmission was shown to be not always more battery energy efficiency and the criterion for selecting single- or multi-hop transmissions depends on both the transmission distance and the relative position of the relay node with respect to the primary nodes.

2.6 Proof of Theorem 2.1

Denote the output current as $p(t)$. Then, at the output stage, the total energy consumption $\mathcal{E}_o$ can be obtained as $\mathcal{E}_o = \int_0^{T_p} V\,|p(t)|\,dt$. Due to the inefficiency of the DC/DC converter and the extra power loss of the power amplifier, the output pulse energy is $\mathcal{E}_p = \eta\mathcal{E}_o/(1+\alpha)$. Define $p_0(t) = p(t)/\int_0^{T_p}|p(t)|\,dt$, then $p(t) = \frac{(1+\alpha)\mathcal{E}_p}{\eta V}p_0(t)$. In addition to the current induced by the transmitted waveform, the circuit power

consumption will also induce a current $I_{ct} = \mathcal{P}_{ct}/(\eta V)$. Thus, the instantaneous current running through the battery is $i(t) = |p(t)| + I_{ct}$. From (2.1), we can obtain the total battery energy consumption in one pulse duration as follows:

$$\begin{aligned}
\mathcal{E}_{0t} &= \int_0^{T_p} \mathcal{P}_0(t)dt = V\int_0^{T_p} \frac{i(t)}{1-\omega i(t)}dt \approx V\int_0^{T_p} i(t)(1+\omega i(t))dt \\
&= V\int_0^{T_p} (|p(t)| + I_{ct})\left[1+\omega(|p(t)|+I_{ct})\right]dt \\
&= V\Bigg[(1+2\omega I_{ct})\int_0^{T_p}|p(t)|\,dt+(1+\omega I_{ct})\int_0^{T_p} I_{ct}dt+\omega\int_0^{T_p}(p(t))^2dt\Bigg] \\
&\approx V\left[\int_0^{T_p}|p(t)|\,dt + \int_0^{T_p} I_{ct}dt + \omega\int_0^{T_p}(p(t))^2dt\right] = V\left[\frac{(1+\alpha)\mathcal{E}_p}{\eta V} + \frac{\mathcal{P}_{ct}}{\eta V}T_p + \omega\frac{(1+\alpha)^2\mathcal{E}_p^2}{\eta^2V^2}\gamma_p\right] \\
&= \frac{\omega\gamma_p(1+\alpha)^2}{\eta^2 V}\mathcal{E}_p^2 + \frac{1+\alpha}{\eta}\mathcal{E}_p + \frac{\mathcal{P}_{ct}}{\eta}T_p\ ,
\end{aligned}$$

where the substitution $\gamma_p := \int_0^{T_p}(p_0(t))^2dt$ is used. In the above derivation, the first approximation comes from the fact that practically $\omega i(t) \ll 1$, and thus $1/(1-\omega i(t)) \approx 1+\omega i(t)$, and the second approximation follows from $\omega I_{ct} \ll 1$.

CHAPTER 3

COGNITIVE RADIO FOR THE SMART GRID 1: COOPERATIVE DIVERSITY OF SPECTRUM SENSING

3.1 Motivation

With wide area measurement, monitoring, protection and control for the power grid, communication loads increase dramatically. Although wireless communications is a very good candidate to fulfill a large part of the tasks, due to its low cost and easy implementation [37], wireless spectrum resources are scarse today due to the explosive growth of wireless communications. Although the existing unlicensed techniques, such as Wi-Fi, Bluetooth and Zigbee, can be utilized, much more bandwidth is still required to fulfill the need for real-time high-volume data communications over the power grid. However, purchasing extra spectrum resources for smart grid applications is both unrealistic and costly. Fortunately although the spectrum is almost fully assigned to various licensed wireless users, its actual utilization is not quite efficient (see e.g., [24]), especially in rural areas where transmission networks are located. To address such inefficiency, cognitive radio systems [33] were proposed as a means of filling the spectrum vacancy in time or space [55] and are also recently proposed and tested for application in the future smart grid [47].

In cognitive radio systems, the unlicensed wireless users (a.k.a. secondary users) take chances to access the spectrum (temporarily or spatially) released by the licensed

users (a.k.a. primary users) so that the spectrum access is dynamic and somewhat opportunistic [69]. To realize this, the first step is to find such opportunities in the primary users' spectrum usage; this is spectrum sensing.

Among existing work on spectrum sensing, some focus on algorithms to improve a single-user's sensing performance by utilizing some side information (see, e.g., [7, 25]). Nonetheless, single-user spectrum sensing is still the system performance bottleneck due to fading and shadowing effects of the wireless channel, as well as the noise uncertainty of the device [56]. To this end, cooperative spectrum sensing by multiple secondary users can significantly improve the sensing performance. Hence, this has become the focus of most ongoing research (see, e.g., [46, 63]). However, while *diversity* has been well acknowledged as the intuitive benefit of cooperative sensing, its rigorous meaning in this setup has remained largely unexplored. In this chapter, we will determine quantitatively the diversity order in various cooperative spectrum sensing schemes.

Diversity has been widely adopted as a fundamental performance indicator in communication systems, where it is defined and quantified in terms of the signal-to-noise ratio (SNR)-dependent behavior of the bit error rate (BER) for symbol detection ([59, Chapter 3]) or the probability of the mutual information ([59, Chapter 5]). This concept was recently extended to the context of cooperative estimation in wireless sensor networks [12]. Therein, diversity refers to the SNR-dependent behavior of the outage probability that the estimation variance exceeds a predefined value. In [39], the missed detection probability is adopted for the determination of diversity in a

cooperative detection scenario while keeping the false alarm probability fixed (SNR independent). In [63], the opposite is considered by emphasizing the false alarm probability while fixing the missed detection probability. However, in the spectrum sensing problem, none of these measures can sufficiently and appropriately quantify diversity.

Unlike traditional detection problems where focusing on either the false alarm or missed detection probabilities while fixing the other is a rather common exercise, doing so in a spectrum sensing problem will risk unbalanced treatment between the system *efficiency* and *reliability*. On the one hand, false alarm probability is of critical importance because the whole purpose of cognitive radio is to maximally utilize the spectrum vacancies, while false alarms lead to undetected spectrum holes and can significantly reduce the efficiency of such usage. On the other hand, missed detections lead to deteriorated "cognition" level and give rise to unexpected interference from the secondary users to the primary ones. In short, false alarm and missed detection probabilities respectively capture the *efficiency* and *reliability* of the overall cognitive system. Hence, a diversity measure of cooperative sensing performance should fairly account for both probabilities. In this work, we consider the false alarm and missed detection probabilities both individually and jointly in terms of the average error probability, which balances system efficiency and reliability.

This new perspective accounting for both efficiency and reliability makes our work unique with respect to existing ones such as [39] and [63]. First, with this perspective, the threshold of the energy detector can be adjusted to improve both performances

simultaneously. Secondly, our study better reflects the nature of the spectrum sensing problem by quantitatively capturing the tradeoff between the efficiency and reliability in closed form. This tradeoff has never been observed or documented before due to the biased emphasis towards specific performance measures. Thirdly, for multi-user sensing with hard information fusion, the local detection strategy and the fusion detection strategy can be jointly optimized based on the tradeoffs established in our analysis.

Our technical contributions are summarized as follows: i) We derive the optimum detection thresholds by minimizing the average error probability in both non-cooperative single-user and cooperative multi-user sensing scenarios. The diversity orders of all three probabilities are then quantified under the optimum thresholds. We also prove that such thresholds lead to the maximum diversity order in both sensing scenarios. ii) We consider two cooperative strategies, namely multi-user sensing with soft information fusion and hard information fusion. The former provides a theoretical bound on the diversity orders and error probability performance in an ideal cooperative sensing setup; whereas the latter leads to practical fusion and decision rules together with their respective quantified diversity orders. iii) We investigate the tradeoff between the system efficiency (via false alarm probability) and reliability (via missed detection probability) and present analytical results to guide practical system designs with differing preference. iv) Depending on whether the secondary users have knowledge of the number of cooperative users, we find that the optimal hard fusion rules are respectively the majority-fusion rule and the OR-fusion rule. v) We verify

the benefit of cooperative sensing and compare the performances of majority-fusion and OR-fusion rules at low SNR.

The signal model, cooperation strategies and the performance metrics with diversity definition will be given in Section 3.2. The diversity orders of the single-user spectrum sensing will be analyzed in Section 3.3, followed by various multi-user cases in Section 3.4. Simulated verifications will be presented in Section 3.5, and concluding remarks will be given in Section 3.6.

Notation: Subscripts 'f', 'md' and 'e' refer to false alarm, missed detection, and average error respectively; subscripts 's' and 'h' refer to fusion with soft information and fusion with hard information respectively. $x \sim \mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian random variable x with mean μ and variance σ^2; $b \sim \text{Bernoulli}(p_0, 1-p_0)$ denotes a Bernoulli random variable b with $p_0 = P[b=0]$; $u \sim \mathcal{U}(a,b)$ denotes a real random variable u uniformly distributed over interval $[a,b]$. $f(\gamma) \sim g(\gamma)$ denotes two functions of γ with $\lim_{\gamma\to+\infty} \frac{f(\gamma)}{g(\gamma)} = k$, where k is a non-zero constant.

3.2 Problem Formulation

In cognitive radio networks, the secondary users need to sense the spectrum usage by the primary users. The performance of spectrum sensing depends heavily on the signal strength at the secondary users. However, the signal strength at a single secondary user can be very low due to channel fading. Thus, cooperation among secondary users can be utilized to improve the sensing performance, as suggested in [15, 24, 25, 46, 63]. In this section, we will introduce the signal model at the spectrum

sensing users, their cooperation strategies and the performance metric of spectrum sensing in terms of the diversity order.

3.2.1 The Signal at Sensing Users

In the spectrum sensing process, the sensing users observe signals under the following two hypotheses:

H_0 : absence of primary user in the spectrum band of interest,

H_1 : presence of primary user in the spectrum band of interest.

We adopt the signal model in [25], where the channels between the primary and the sensing users are Rayleigh fading with additive white Gaussian noise (AWGN). Then the received signal at the sensing user is given by [25]:

$$\begin{aligned} r|H_0 &= n \sim \mathcal{CN}(0, \sigma_n^2), \\ r|H_1 &= hx + n \sim \mathcal{CN}(0, E_x\sigma_h^2 + \sigma_n^2), \end{aligned}$$

where n is AWGN with variance σ_n^2, h is the channel coefficient with variance σ_h^2 and x is the signal from the primary user with energy E_x. Suppose the sensing users know the noise variance, hence, without loss of generality, we normalize the noise variance to 1. Accordingly, the signal at the sensing users becomes:

$$\begin{aligned} r|H_0 &= n \sim \mathcal{CN}(0, 1), \\ r|H_1 &= hx + n \sim \mathcal{CN}(0, \gamma + 1), \end{aligned} \tag{3.1}$$

where $\gamma \triangleq E_x\sigma_h^2/\sigma_n^2$ is the average signal-to-noise ratio (SNR) at the sensing users. With geographically distributed sensing users, it is reasonable to assume that they experience independent fading channels. Thus, the received signals for different sensing users r_is are conditionally independent under each hypothesis.

3.2.2 Cooperative Strategies

Cooperative spectrum sensing requires cooperation among multiple sensing users. In our analysis, a fusion center collects information from all secondary users and facilitates their cooperation. Ideally, the cooperation benefit is maximized if all sensing information from all secondary users reaches the fusion center without any loss. This condition, however, can not always be satisfied due to the limited spectrum resource available to the secondary user system. Hence, we will next consider two types of sensing strategies, namely cooperative multi-user sensing with soft information fusion and cooperative multi-user sensing with hard information fusion.

3.2.2.1 Multi-User Sensing with Soft Information Fusion

In this case, the fusion center can obtain information from the distributed secondary users perfectly. This provides a best case scenario for cooperative sensing among multiple secondary users. Although this may not be practically achievable, it does provide a useful bound on the multi-user sensing performance. Moreover, this is also a good model for the case where multiple independent faded copies of the primary user's signal are collected at a single secondary user. For example, multiple receive antennas with appropriate antenna spacing can provide independent faded copies of

the signal, or in the case of a fast fading scenario, the signals from different time slots are independently distributed. In these cases, the single secondary user can be regarded as the fusion center and the different sources of independent faded signals can be regarded as the multiple spectrum sensing nodes with lossless transmission to the fusion center, leading to soft information fusion.

3.2.2.2 Multi-User Sensing with Hard Information Fusion

In a more practical multi-user setup, each distributed secondary user senses the spectrum usage and then only transmits the one bit sensing decision, '0' for absence of primary users or '1' for presence of primary users, to the fusion center.

3.2.3 Performance Metric and Diversity Order

In traditional signal detection problems, the receiver operating characteristic (ROC) curves (false alarm probability P_f vs. missed detection probability P_{md}) are generally used to graphically illustrate the detection performance [43]. Every ROC curve is plotted for a certain combination of the system parameters such as SNR, number of cooperative users and so on. As a result, they do not provide an explicit quantitative relationship between the system parameters and the system metrics (false alarm, missed detection and average detection error probabilities) [15]. Hence, to better illustrate the effects of the system parameters on the performance, in this chapter, we analyze each single system metric as a function of the system parameter variables.

As introduced in Section 6.1, the performance of spectrum sensing is indicated by the false alarm probability P_f, the missed detection probability P_{md} and the average

error probability P_e. P_f is the probability of deciding on H_1 when H_0 is true (type I error); P_{md} is the probability of deciding on H_0 when H_1 is true (type II error); P_e is the average probability of making a wrong decision. Physically, P_f determines the capability of detecting the available spectrum resource, and thus the *efficiency* of the system; and P_{md} indicates the level of interference that the secondary user system introduces to the primary user system, and thus the *reliability* of the system. As a result, P_e combines the efficiency and reliability considerations. We denote the probability of the absence of the primary user (H_0) as α and thus that of the presence of the primary user (H_1) as $(1-\alpha)$, then the average error probability is $P_e = \alpha P_f + (1-\alpha)P_{md}$.

The concept of diversity was introduced in wireless communications to quantify the effects of independent fading in space, time, frequency or code space on the improvement of the system performance [59]. Quantitatively, the diversity order is defined as:

$$d = -\lim_{\mathrm{SNR}\to+\infty} \frac{\log P}{\log \mathrm{SNR}}$$

where P can be the bit error rate or the outage probability of the communication system and SNR is the average signal-to-noise ratio. In cooperative sensing, the fusion center also receives multiple copies of the original signal under independent fading. Hence, the sensing performance is expected to exhibit a similar behavior. Here we define the diversity order in sensing scenarios as:

$$d_* = -\lim_{\gamma\to+\infty} \frac{\log P_*}{\log \gamma}$$

where $*$ can be f (false alarm), md (missed detection) or e (average error). Accordingly, there will be false alarm diversity d_f, missed detection diversity d_{md} and average error diversity d_e. Obviously, $d_e = \min\{d_f, d_{md}\}$ when $d_f \neq d_{md}$. In the following sections, we will quantify the diversity order of spectrum sensing according to the definitions above for three cases: single-user sensing, multi-user cooperative sensing with soft information fusion and multi-user cooperative sensing with hard information fusion. In addition, we will show that, though the diversity order is defined in the limit when $\gamma \to +\infty$, it actually shows up quite early at low SNR.

3.3 Single-User Sensing

To achieve uniformly most powerful detection performance, we use the Neyman-Pearson (NP) detector [43]. With our signal model, the NP test is the likelihood ratio test:

$$\lambda = |r|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \theta \tag{3.2}$$

where θ is the threshold of the test. According to (3.1), the distribution of the decision statistic is:

$$\begin{aligned} f(\lambda|H_0) &= e^{-\lambda} \quad (\lambda > 0) \\ f(\lambda|H_1) &= \frac{1}{\gamma+1} e^{-\frac{\lambda}{\gamma+1}} \quad (\lambda > 0). \end{aligned} \tag{3.3}$$

Hence the probabilities of false alarm and missed detection are, respectively:

$$P_f = \int_\theta^{+\infty} f(\lambda|H_0) d\lambda = \int_\theta^{+\infty} e^{-\lambda} d\lambda = e^{-\theta} \tag{3.4}$$

and

$$P_{md} = \int_0^{\theta} f(\lambda|H_1)d\lambda = \int_0^{\theta} \frac{1}{\gamma+1} e^{-\frac{\lambda}{\gamma+1}} d\lambda = 1 - e^{-\frac{\theta}{\gamma+1}}. \tag{3.5}$$

3.3.1 Diversity Order when Minimizing P_e

As shown in (3.4) and (3.5), the performance metrics P_f, P_{md} and P_e all rely on the choice of the decision threshold θ. Clearly, one may choose different thresholds when optimizing different performance metrics. Recall that P_f captures the cognitive system's *efficiency* while P_{md} captures its *reliability*. To balance the system efficiency and reliability, we will optimize the threshold θ by minimizing the average error probability $P_e = \alpha P_f + (1-\alpha)P_{md}$. Setting $dP_e/d\theta = 0$ and solving for θ, we obtain the optimum threshold as:

$$\theta^o = \left(1 + \frac{1}{\gamma}\right) \log\left[\frac{\alpha}{1-\alpha}(\gamma+1)\right], \tag{3.6}$$

where log is base-e throughout this chapter unless otherwise specified. Using this threshold, as $\gamma \to +\infty$, we have:

$$\begin{aligned} P_f &= e^{-\theta} = e^{-(1+\frac{1}{\gamma})\log\left[\frac{\alpha}{1-\alpha}(\gamma+1)\right]} \\ &= (\gamma+1)^{-(1+\frac{1}{\gamma})} \left(\frac{\alpha}{1-\alpha}\right)^{-\left(1+\frac{1}{\gamma}\right)} \sim (1+\gamma)^{-1} \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} P_{md} &= 1 - e^{-\frac{\theta}{\gamma+1}} = 1 - e^{-\frac{1}{\gamma}\log\left[\frac{\alpha}{1-\alpha}(\gamma+1)\right]} \\ &\sim \gamma^{-1} \log\left[\frac{\alpha}{1-\alpha}(\gamma+1)\right]. \end{aligned} \tag{3.8}$$

Thus, their respective diversity orders can be obtained as:

$$\begin{aligned} d_f &= -\lim_{\gamma\to+\infty} \frac{\log P_f}{\log \gamma} = 1 \\ d_{md} &= -\lim_{\gamma\to+\infty} \frac{\log P_{md}}{\log \gamma} = 1 \\ d_e &= \min(d_f, d_{md}) = 1. \end{aligned} \tag{3.9}$$

Accordingly, we establish the following result:

Theorem 3.1 *For single-user spectrum sensing, when the threshold θ is chosen to minimize the average error probability P_e as in* (3.6)*, the diversity order of the NP detector is $d_e = d_f = d_{md} = 1$.*

This theorem is quite intuitive since any single sensing user only has one copy of the original signal going through the fading channel and it is well known that the probability of deep fading in this case is proportional to γ^{-1} (see e.g. [59]).

From the analysis above, we see that the a priori probabilities of the hypotheses α and $(1-\alpha)$ do not affect the diversity orders of the performance. Without loss of generality, to simplify the following analyses on the diversity orders, we choose $\alpha = \frac{1}{2}$ for the rest of this chapter.

3.3.2 False Alarm Diversity versus Missed Detection SNR Gain

In Theorem 3.1, we choose θ^o to minimize the average detection error probability P_e, which is the average of efficiency and reliability. However, in some systems, the two features may have different levels of importance. Thus, we will next analyze P_f and P_{md} separately. Our analysis will reveal an interesting tradeoff between the system

efficiency and reliability. This tradeoff can be exploited to achieve the desirable P_f and P_{md} performance and accordingly the preferable spectrum usage efficiency and interference level.

From (3.4) and (3.6), we notice that, if one changes the threshold to $\theta' = d_0\theta^o$, then the false alarm diversity order changes from 1 in (3.9) to:

$$d'_f = -\lim_{\gamma\to+\infty} \frac{\log P'_f}{\log \gamma} = -\lim_{\gamma\to+\infty} \frac{-d_0(1+\frac{1}{\gamma})\log(\gamma+1)}{\log\gamma} = d_0. \tag{3.10}$$

On the other hand, with this new threshold θ', as $\gamma \to +\infty$, we have:

$$P'_{md} = 1 - e^{-\frac{\theta'}{\gamma+1}} = 1 - e^{-\frac{d_0}{\gamma}\log(\gamma+1)} \sim d_0\gamma^{-1}\log(\gamma+1). \tag{3.11}$$

This implies that $d'_{md} = 1$, which is identical to d_{md} in (3.9) with threshold θ^o. In other words, the missed detection diversity order remains unaltered. However, the scalar difference between (3.8) and (3.11) suggests that, to ensure $P'_{md} \approx P_{md}$ as $\gamma \to +\infty$, ones needs $\gamma' = d_0\gamma$. This implies that the missed detection probability P_{md} exhibits a $-10\log_{10} d_0$ dB SNR gain (or equivalently $10\log_{10} d_0$ dB SNR loss) when the threshold is chosen as $\theta = d_0\theta^o$. This interesting phenomenon may result from the fact that the false alarm probability is the right tail of the Rayleigh distribution which decays very rapidly; whereas the missed detection is the left tail of the Rayleigh distribution which decays quite slowly.

We summarize the tradeoff between the false alarm diversity and the missed detection SNR gain in the following corollary:

Collarory 3.1 *For single-user spectrum sensing, when the threshold is set to* $d_0\theta^o$ *with* θ^o *given in* (3.6)*, the false alarm diversity order becomes* $d_f = d_0$*, while the missed detection diversity order remains* $d_{md} = 1$ *but the* P_{md} *curve exhibits* $a - \log_{10} d_0$ *dB SNR gain.*

The tradeoff between the false alarm diversity and the missed detection SNR gain presented above provides system designers with a flexible tool to achieve the desirable tradeoff between the spectrum usage efficiency of the secondary users and the reliability of the primary users. For example, if the primary users in the cognitive system are capable of interference suppression and the spectrum usage efficiency is of major concern for the system designer, then, by the properties above, the secondary users can set the threshold as $d_0\theta^o$ with $d_0 > 1$. This means, the false alarm diversity is increased to $d_0(> 1)$ by sacrificing a $10 \log_{10} d_0$ dB SNR loss for the missed detection probability. On the other hand, if the primary users are vulnerable to interference and the performance of primary users in the cognitive system is of the major concern, then, the secondary users can set the threshold at $d_0\theta^o$ with $0 < d_0 < 1$. As a result, there will be a $-10 \log_{10} d_0$ dB SNR gain for the missed detection probability by reducing the false alarm diversity to $d_0(< 1)$.

As shown later in Section 3.4.2, this flexibility of the false alarm diversity can also be utilized to maximize the diversity order of the multi-user sensing with hard information fusion.

3.4 Multi-User Sensing

3.4.1 Soft Information Fusion

With a soft information fusion strategy, the fusion center receives $r_1, r_2, \ldots, r_N$ from the distributed sensing users, where N is the total number of cooperative sensing users and r_is are conditionally independent identically distributed (i.i.d.) under both H_0 and H_1. Similar to Section 3.3, the NP test is

$$\lambda_s = \sum_{i=1}^{N} |r_i|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \theta_s, \tag{3.12}$$

where the subscript 's' refers to soft information fusion.

Since r_is are conditionally independent, and according to (3.1), we have:

$$\begin{aligned} f(\lambda_s|H_0) &= \lambda_s^{N-1}\frac{e^{-\lambda_s}}{(N-1)!} \quad (\lambda_s > 0) \\ f(\lambda_s|H_1) &= \lambda_s^{N-1}\frac{e^{-\frac{\lambda_s}{\gamma+1}}}{(N-1)!(\gamma+1)^N} \quad (\lambda_s > 0). \end{aligned} \tag{3.13}$$

Hence the probabilities of false alarm and missed detection are, respectively:

$$P_{f,s} = \int_{\theta_s}^{+\infty} f(\lambda_s|H_0)d\lambda_s = \left(\sum_{i=0}^{N-1}\frac{\theta_s^i}{i!}\right)e^{-\theta_s} \tag{3.14}$$

and

$$P_{md,s} = \int_{0}^{\theta_s} f(\lambda_s|H_1)d\lambda_s = \left(\sum_{i=N}^{+\infty}\frac{\theta_s^i}{i!(\gamma+1)^i}\right)e^{-\frac{\theta_s}{\gamma+1}}. \tag{3.15}$$

Accordingly, the average error probability is $P_{e,s} = \frac{1}{2}P_{f,s} + \frac{1}{2}P_{md,s}$. Similar to Section

3.3, minimizing $P_{e,s}$ by taking $dP_{e,s}/d\theta_s = 0$, we obtain the optimum threshold as:

$$\theta_s^o = N\left(1 + \frac{1}{\gamma}\right)\log(\gamma + 1). \tag{3.16}$$

Using this threshold, we establish the following theorem:

Theorem 3.2 *For multi-user sensing with soft information fusion, when the threshold θ_s is chosen as in* (3.16) *to minimize the average error probability $P_{e,s}$, the diversity order of the NP detector is $d_{e,s} = d_{f,s} = d_{md,s} = N$, where N is the number of cooperative users.*

Proof. See Section 3.7. ■

This theorem is also intuitive in that the fusion center has copies of the original received signals from N independently fading channels. Similar to Section 3.3.2, we can also choose the threshold as $\theta_s' = d_0\theta_s^o$, where d_0 can be any positive number and is not necessarily integer, to increase the false alarm diversity to $d_0 N$ while keeping the missed detection diversity unaltered at N. In this case, there is also a tradeoff between the missed detection SNR gain and the false alarm diversity. For the false alarm diversity to be $d_0 N$, the missed detection probability will exhibit a $-10\log_{10} d_0$ dB SNR gain (or equivalently $10\log_{10} d_0$ dB SNR loss).

3.4.2 Hard Information Fusion

With the hard information fusion strategy, each sensing user makes its own local hard decision and then sends the binary decision b_i to the fusion center. For simplicity,

we assume that all distributed sensing users employ the same threshold θ_l for their local decisions where subscript 'l' stands for local. The corresponding local false alarm and missed detection probabilities are denoted as $P_{f,l}$ and $P_{md,l}$, respectively. Clearly, b_i follows conditionally i.i.d. Bernoulli distribution with $(1 - P_{f,l})$ and $P_{md,l}$ as the probabilities of value 0 under H_0 and H_1, respectively; that is:

$$\begin{aligned} b_i|H_0 &\sim \text{Bernoulli}(1 - P_{f,l}, P_{f,l}) \\ b_i|H_1 &\sim \text{Bernoulli}(P_{md,l}, 1 - P_{md,l}) \end{aligned} \tag{3.17}$$

In this case, the NP test becomes:

$$\lambda_h = \sum_{i=1}^{N} b_i \underset{H_0}{\overset{H_1}{\gtrless}} \theta_h, \quad \theta_h = 1, 2, \ldots, N, \tag{3.18}$$

where subscript 'h' refers to hard information fusion. Accordingly, the distribution of λ_h is:

$$\begin{aligned} f(\lambda_h|H_0) &= \binom{N}{\lambda_h} P_{f,l}^{\lambda_h}(1 - P_{f,l})^{N-\lambda_h}, \ \lambda_h = 0, 1, \ldots, N; \\ f(\lambda_h|H_1) &= \binom{N}{\lambda_h}(1 - P_{md,l})^{\lambda_h} P_{md,l}^{N-\lambda_h}, \ \lambda_h = 0, 1, \ldots, N. \end{aligned} \tag{3.19}$$

In hard information fusion, there are two levels of decision making, each level having its own decision performance. For the local decision, there are diversity orders for the local false alarm probabilities ($d_{f,l}$) and local missed detection probabilities ($d_{md,l}$). At the fusion center, there are also corresponding diversity orders for the overall hard-decision false alarm probability ($d_{f,h}$), missed detection probability ($d_{md,h}$) and average error probability ($d_{e,h}$). Here we establish the relationship between the

local decision diversity orders with the overall diversity orders at the fusion center in the following theorem:

Theorem 3.3 *For multi-user sensing with* 1*-bit hard information fusion, and with the fusion center threshold* θ_h *(*$\theta_h = 1, 2, \ldots, N$*), the diversity orders of the NP detector are* $d_{f,h} = \theta_h d_{f,l}$*,* $d_{md,h} = (N-\theta_h+1)d_{md,l}$ *and* $d_{e,h} = \min\{\theta_h d_{f,l}, (N-\theta_h+1)d_{md,l}\}$*, where* N *is the number of cooperative users.*

Proof. See Section 3.8. ■

Similar to the single-user sensing and cooperative sensing with soft information fusion, cooperative sensing with hard information fusion also provides the system designer with the flexibility of balancing between the spectrum efficiency of the secondary users (via $d_{f,h}$) and the reliability of the primary users (via $d_{md,h}$) by the choice of the threshold θ_h. A larger θ_h will improve the false alarm performance, leading to higher spectrum usage efficiency of the secondary users; while a smaller θ_h will improve the missed detection performance, leading to enhanced reliability of the primary users. However, it is worth noting that the tradeoff and flexibility here are very different from what we have discussed in Corollary 3.1 for the single-user sensing and the multi-user soft-fusion cases. In previous cases, the tradeoff was between the false alarm *diversity* and the missed detection *SNR gain*, while here in the case of multi-user hard-fusion, the tradeoff is between the false alarm *diversity* and the missed detection *diversity*.

Note that, though the local threshold θ_l does not appear explicitly in Theorem 3.3, it affects the overall system implicitly via $d_{f,l}$ and $d_{md,l}$. Hence, if one opts to

minimize $P_{e,h} = \frac{1}{2}P_{f,h} + \frac{1}{2}P_{md,h}$, one needs to jointly choose both the optimum local threshold θ_l at the individual sensing users and the optimum hard decision threshold θ_h at the fusion center. However, under this fusion rule, θ_l has a very complicated form in $P_{e,h}$ through parameters $P_{f,l}$ and $P_{md,l}$, rendering optimization intractable. Even with numerical techniques, the optimization over θ_l still requires knowledge of the total number of distributed sensors in the network N, which is not always available to the secondary users in real applications. However, with Theorem 3.3, one can optimize the overall hard decision fusion performance from the diversity perspective with different strategies as detailed in the following two scenarios.

B.1) Number of cooperative users N unknown

In this case, the sensing users can only perform their optimum detection locally. From Theorem 3.1, the local threshold is $\theta_l = \theta^o$ and the local diversities are $d_{f,l} = d_{md,l} = 1$. According to Theorem 3.3, the diversity orders at the fusion center are $d_{f,h} = \theta_h$, $d_{md,h} = N - \theta_h + 1$ and $d_{e,h} = \min\{\theta_h, N - \theta_h + 1\}$. With equal emphasis on false alarm and missed detection performance, by maximizing $\min\{\theta_h, N - \theta_h + 1\}$, we obtain the optimum threshold θ_h^o at the fusion center that maximizes the detection error diversity:

Collarory 3.2 *For multi-user sensing with* 1*-bit hard information fusion, and with each sensing user using the locally optimum threshold* $\theta_l = \theta^o$*, the optimum threshold at the fusion center in the sense of maximizing the detection error diversity is* $\theta_h^o = \lfloor \frac{N+1}{2} \rfloor$ *or* $\theta_h^o = \lceil \frac{N+1}{2} \rceil$ *with* $d_{e,h} = \lfloor \frac{N+1}{2} \rfloor$*, where* N *is the number of cooperative users.*

In this case, the strategy is the so-termed majority-fusion rule. Notice that under this rule, about half of the diversity is lost compared with the soft information fusion. This indicates that a hard decision at each local sensing user leads to considerable information loss of the received signal.

B.2) Number of cooperative users N known

From Corollary 3.1, we have seen that the false alarm diversity at each local sensing user is flexible. It is shown in the following corollary that this property can be utilized to maximize the average error diversity.

Collarory 3.3 *For multi-user sensing with* 1*-bit hard information fusion, if each distributed sensing user knows the number of cooperative users* N*, the average error diversity order at the fusion center can be maximized by choosing local decision threshold* $\theta_l = N\theta^o$ *with* θ^o *given in* (3.6) *and the fusion decision threshold* $\theta_h = 1$*.*

Proof. See Section 3.9. ■

Notice that the decision strategy turns out to be the so-termed OR-fusion rule. From this corollary, we see that with the number of sensors known at each local sensor, the diversity of the average detection error probability of hard information fusion equals that of soft information fusion ($d_{e,h} = d_{e,s} = N$). In other words, knowledge of N completely compensates for the loss of information by local hard decisions, in terms of the detection error diversity. However, we should also notice that though the diversity performance of the two cases are identical, their average error probability performances are still different. As detailed in Section 3.3.2, there is a $10 \log N$ SNR loss for the missed detection probability $P_{md,l}$ by setting $\theta_l = N\theta^o$.

Remarks: As stated in Section 3.2.2, our results on multi-user sensing with soft information fusion can be readily applied to the case of combining signals from different multiple time slots at a single user, as long as the combined received signals experience independent fading. In this case, if the system also uses cooperative multi-user sensing with hard information fusion among these multiple time slot sensors, our analysis in this section can be readily extended by combining the results in Theorems 3.2 and 3.3. In addition, for the correlated fading case, intuitively, we expect that the diversity orders equal the rank of the correlation matrix of the received signals, which will be justified in the following section.

3.5 Simulation Results

3.5.1 Single-User Sensing

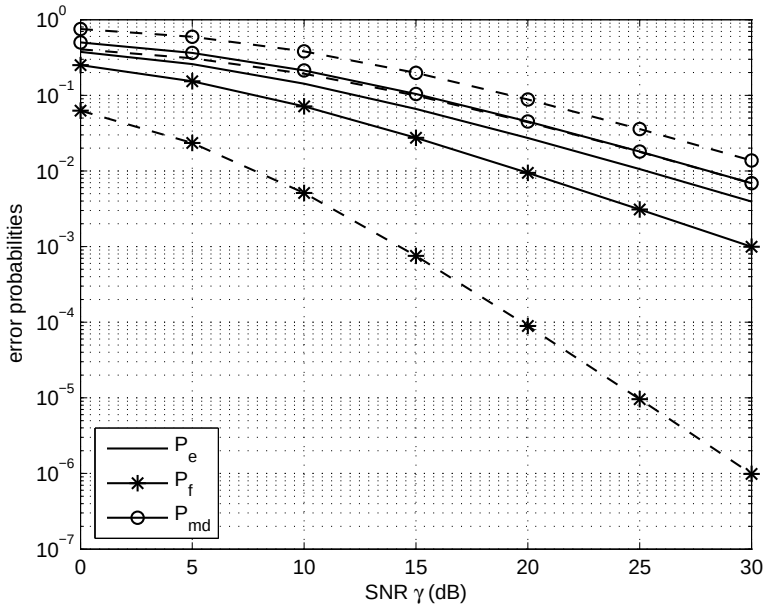


Figure 3.1: Single-user sensing. Solid curves: threshold $\theta = \theta^o$; Dashed curves: threshold $\theta = 2\theta^o$.

For single-user sensing, we use the test given by (3.2) and the threshold of (3.6) to obtain the average error probability, false alarm probability and the missed detection probability. These probabilities are shown as the solid curves in Fig. 3.1. All three curves exhibit the same slope, indicating diversity orders $d_f = d_{md} = d_e = 1$. To illustrate the tradeoff between the false alarm diversity and missed detection SNR gain (loss) discussed in Section 3.3.2, we change the threshold to $\theta' = 2\theta^o$ and obtain the dashed curves in Fig. 3.1. From the figure, we see that $d'_f = 2$ and $d'_{md} = d_{md} = 1$, as predicted Corollary 3.1. In addition, comparing the solid curve with the dashed one, at high SNR, there is an approximately $10 \log_{10} d_0 = 3$ dB SNR loss. This is the price paid for the increase of the diversity for the false alarm probability.

3.5.2 Multi-User Sensing with Soft Information Fusion

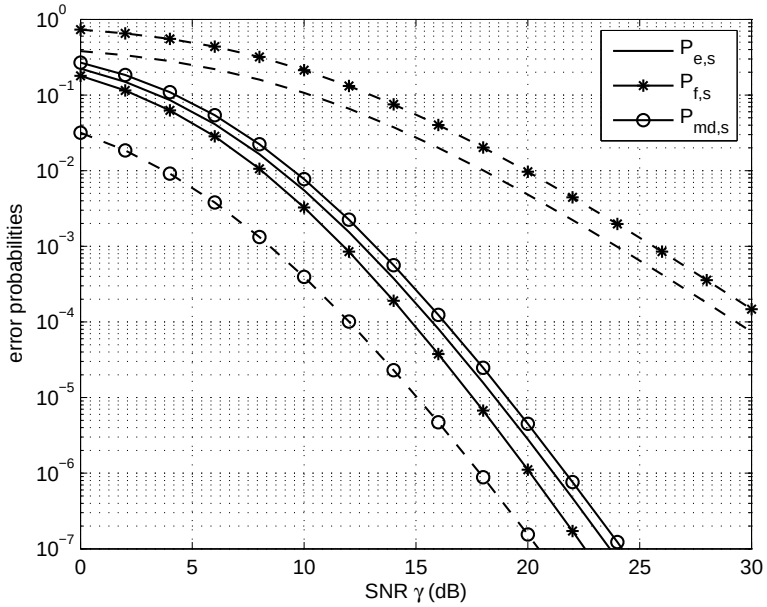


Figure 3.2: $N = 5$ multi-user sensing with soft decision fusion. Solid curves: threshold $\theta = \theta^o_s$; Dashed curves: threshold $\theta = 0.5\theta^o_s$.

Here, we simulate the multi-user sensing with soft information fusion strategy. The total number of cooperating users is $N = 5$. From the analysis in Section 3.4.1, we expect the relative performance of $P_{e,s}$, $P_{f,s}$ and $P_{md,s}$ to be similar to the single-user sensing case, except for the diversity order of 5. This is verified in Fig. 3.2. The solid curves show that with the threshold minimizing the average error probability, the diversity orders are $d_{e,s} = d_{f,s} = d_{md,s} = N = 5$. In multi-user sensing with soft information fusion, we can also set the threshold to $\theta'_s = d_0\theta^o_s$ to make $d_{f,s} = d_0 N$. The dashed curves in Fig. 3.2 show $P_{f,s}$ and $P_{md,s}$ with $d_0 = 0.5$. Similar to the single-user sensing case, comparing the solid curve with the dashed one for $P_{md,s}$, we verify that at high SNR, there is approximately $-10\log_{10} d_0 = 3$ dB SNR gain for the decrease of the false alarm diversity (from 5 to 2.5).

In the simulations above, the fading coefficients at each user are assumed to be independent. However, for multi-time-slot sensing which can be modeled as multi-user sensing with soft information fusion, the fading coefficients can be correlated. To investigate the diversities in this case, we plot the simulation results of the detector under the threshold η^o_s in Fig. 3.3. Notice that in this case, the signals under hypothesis H_0 remains unaltered, thus it suffices to give the missed detection performances only. In this simulation, the number of time slots $N = 5$ and the correlations of the fading coefficients are assumed to be $\mathbb{E}[h^*_i h_j] = r^{|i-j|}$ where i and j are indices for the

time slot. Under this assumption, the correlation matrix is

$$\Lambda_{\mathbf{h}} = \mathbb{E}[\mathbf{h}^H \mathbf{h}] = \begin{bmatrix} 1 & r & r^2 & r^3 & r^4 \\ r & 1 & r & r^2 & r^3 \\ r^2 & r & 1 & r & r^2 \\ r^3 & r^2 & r & 1 & r \\ r^4 & r^3 & r^2 & r & 1 \end{bmatrix}$$

In Fig. 3.3, we see that when the correlation (r) among the fading coefficients increases, the performance degrades. However, as long as the correlation matrix $\Lambda_{\mathbf{h}}$ is full-rank, the diversity results obtained under the independent fading scenario still hold. When the matrix loses rank to rank 1 at $r = 1$, the diversity order reduces to 1.

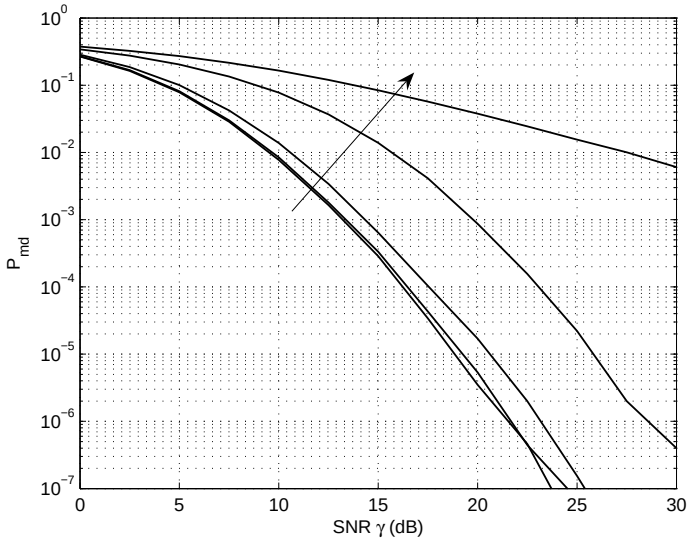


Figure 3.3: $N = 5$ multi-user sensing with soft decision fusion on correlated signals with correlation $\mathbb{E}[h_i^* h_j] = r^{|i-j|}$. In the direction of arrow: $r = 0, r = 0.2, r = 0.5, r = 0.9, r = 1$, respectively.

3.5.3 Multi-User Sensing with Hard Information Fusion

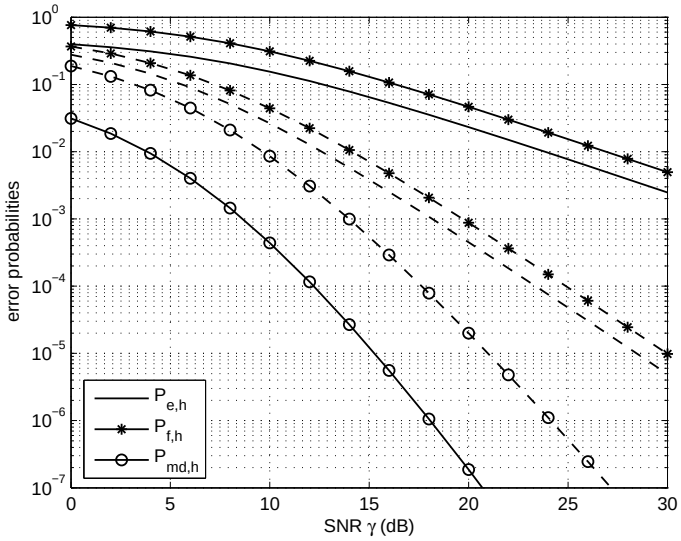


Figure 3.4: $N = 5$ multi-user sensing with hard decision fusion. Solid curves: threshold $\theta_l = \theta^o$, $\theta_h = 1$; Dashed curves: threshold $\theta_l = \theta^o$, $\theta_h = 2$.

For the multi-user sensing with hard information fusion in Section 3.4.2, we first consider the case where the number of cooperative users N is not available to the individual sensing users. Then, each distributed sensing user makes the locally optimum hard decision, i.e., $\theta_l = \theta^o$ as defined in (3.6). In this case, as Theorem 3.3 dictates, the false alarm diversity, the missed detection diversity and the average error diversity are all heavily dependent on the threshold θ_h at the fusion center. Fig. 3.4 shows the behavior of $P_{f,h}$, $P_{md,h}$ and $P_{e,h}$ of the hard information fusion strategy with $\theta_h = 1$ (solid curves) and $\theta_h = 2$ (dashed curves) when the number of cooperative users is $N = 5$. When $\theta_h = 1$, the solid curves show that $d_{f,h} = 1$, $d_{md,h} = 5$ and $d_{e,h} = 1$; when $\theta_h = 2$, the dashed curves show that $d_{f,h} = 2$, $d_{md,h} = 4$ and $d_{e,h} = 2$. These results are consistent with Theorem 3.3 and illustrate the tradeoff between the false alarm diversity and the missed detection diversity.

With locally optimum hard decision ($\theta_l = \theta^o$) at distributed sensing users, the threshold maximizing the average error diversity should be chosen as $\lfloor \frac{N+1}{2} \rfloor$ or $\lceil \frac{N+1}{2} \rceil$ according to Corollary 3.2. This result is shown in Fig. 3.5 when $N = 5$. From this figure, we see that $\theta_h = 3$ gives the maximum average error diversity. Compared with the error performance under soft information fusion, we see here that this locally optimum hard decision strategy suffers from a large loss of diversity.

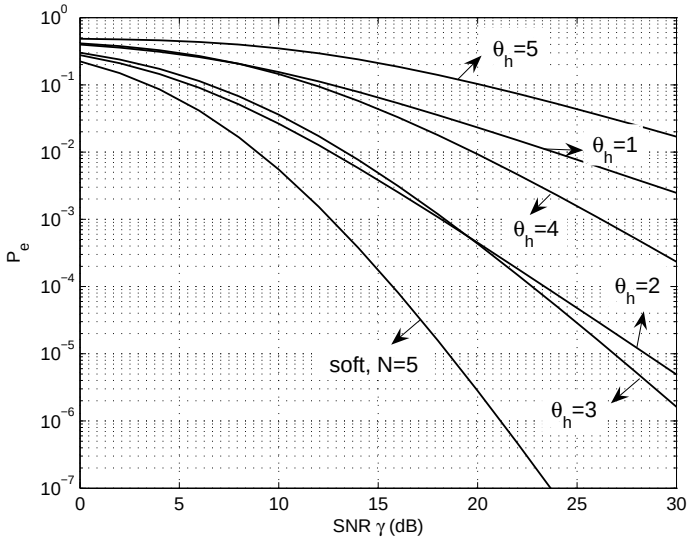


Figure 3.5: $N = 5$ multi-user sensing with soft information fusion and hard information fusion with $\theta_l = \theta^o$ and various θ_hs.

When the distributed sensing users know the number of cooperative users in the network, then the flexibility of the false alarm diversity in Corollary 3.1 can be utilized to maximize the diversities of the hard information fusion. In this case, the local decision threshold is $\theta_l = N\theta^o$ where θ^o is defined in (3.6). Fig. 3.6 shows the performance with this strategy and compares this with that of the soft information fusion strategy. From this figure, we see that in terms of diversity order d_e, the hard information fusion with adjusted local threshold equals the soft information

fusion. However, the soft information fusion has a huge SNR advantage over the hard information fusion on the missed detection and average error probabilities. This is due to the fact that in the local decisions, a $10\log_{10}5 = 7$ dB SNR loss of missed detection probability is introduced to increase the false alarm diversity from $d_{f,l} = 1$ to $d_{f,l} = 5$ by setting $\theta_l = 5\theta^o$. This also explains why the false alarm performance in this hard information fusion strategy with knowledge of the number of cooperative users is better than that in the soft information fusion strategy. In addition, compared with the dotted curves, we see that this strategy recovers the diversity loss introduced by the locally optimum hard decision.

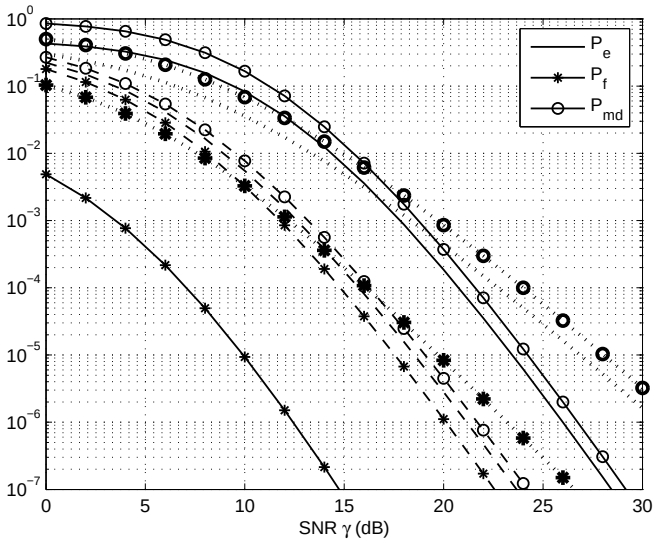


Figure 3.6: $N = 5$ multi-user sensing with OR-fusion rule with $\theta_l = N\theta^o$ (solid curves), majority-fusion rule with $\theta_l = \theta^o$ (dotted curves), and soft information fusion (dashed curves).

3.5.4 Low SNR Performance Comparisons

The analysis in this chapter focuses on the diversity order which is meaningful only at high SNR. However, from the simulation results above, we see that the diversity

shows up quite early in SNR. Actually, most cases with better diversity order also results in better performance as shown at low SNR in Figs. 3.1-3.6.

To further illustrate the relative performance of various decision strategies in the low SNR range, we show in Fig. 3.7 the average detection error of soft information fusion ($d_{e,s} = N$) with $\theta_s = \theta_s^o$, hard information fusion with $\theta_l = \theta^o$ and $\theta_h^o = \lfloor \frac{N+1}{2} \rfloor$ as in Corollary 3.2 ($d_{e,h} = \lfloor \frac{N+1}{2} \rfloor$) and hard information fusion with $\theta_l = N\theta^o$ and $\theta_h = 1$ as in Corollary 3.3 ($d_{e,h} = N$) at $\gamma = 0$ dB. Comparing the solid curve with the other two curves, we see that at low SNR, the soft information fusion is still better than the hard information fusion. This is consistent with our expectation in that the soft information fusion center collects all the information at individual sensing users and the threshold maximizing the diversity in (3.16) is actually optimum for any SNR value γ. Also, the performance of soft information fusion improves as the number of cooperative sensing users N increases. This means that a bigger number of cooperative sensing users not only increases the diversity order at high SNR but also boosts the performance at low SNR.

Comparing the dashed curve with the dotted one, we see that, though with the knowledge of the number of cooperative sensing users, the diversity of the hard information fusion can be maximized as dictated in Corollary 3.3, its low-SNR performance is greatly compromised. This is due to the fact that with the strategy in Corollary 3.3, the performance of local missed detection $P_{md,l}$ is sacrificed for the sake of false alarm diversity. At low SNR, the benefit of false alarm is insignificant while the loss of local missed detection performance takes dominance on the overall decision fusion

performance. Also, it should be noticed that with the strategy in Corollary 3.3, larger number of cooperative users N causes more performance loss of local missed detection $P_{md,l}$, leading to more significantly deteriorated overall performance of the hard information fusion.

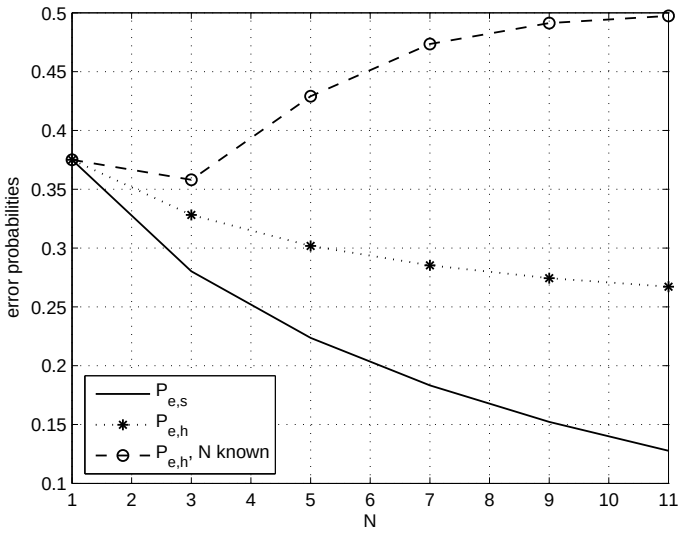


Figure 3.7: Low SNR performance of different decision strategies ($\gamma = 0$ dB). Solid: soft information fusion; dotted: majority fusion with $\theta_l = \theta^o$; dashed: OR fusion with $\theta_l = N\theta^o$.

With this comparison, we also obtain the conclusion that at low SNR, the majority-fusion introduced in Corollary 3.2 is preferred over the OR-fusion introduced in Corollary 3.3 for hard information fusion. This is quite intuitive since the majority-fusion rule is more robust to individual errors.

3.5.5 Simulation with Imperfect SNR Estimate

In all our strategies, one needs the SNR information to determine the decision thresholds θ^o, θ_s^o and θ_h. However, in real applications, the SNR information can only be obtained via estimation and is thus never perfectly known. Next, we simulate

our algorithms with imperfectly estimated SNR values. First, suppose our estimated SNR $\hat{\gamma} = u\gamma$ where $u \sim \mathcal{U}(a, b)$ is the multiplicative noise factor. As discussed before, the performance of the single-user sensing is fundamental to that of the multi-user sensing case. Therefore, here we only present the effect of imperfect SNR estimate on single-user sensing in Fig. 3.8.

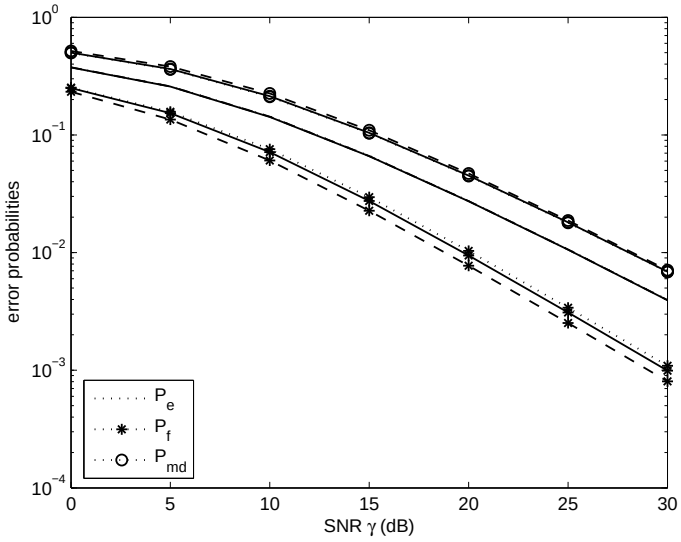


Figure 3.8: Performances of single-user spectrum sensing with imperfect SNR estimate $\hat{\gamma} = u\gamma$. Solid curves: $u = 1$ (perfect); Dotted curves: $u \sim \mathcal{U}(0.5, 1.5)$ (unbiased); Dashed curves: $u \sim \mathcal{U}(1, 1.5)$ (biased).

This figure shows the performances of the single-user sensing with perfect (solid), unbiased (dotted) and biased (dashed) SNR estimate. From this figure we see that there is very little difference among the performances with different SNR estimation quality. This means that our spectrum sensing algorithm is robust against imperfect SNR estimate. This is because though the threshold expression involves the SNR value, the algorithm only requires the threshold to increase or decrease in consistency with SNR. It is also worth noting that with the biased SNR estimate, if the bias is positive as in the dashed curve case, the performance of false alarm (P_f) will actually

get improved because we are choosing a larger threshold. We expect the opposite when the bias is negative.

3.6 Conclusions

In this chapter, we analyzed various cooperative spectrum sensing strategies under different scenarios. By considering both false alarm and missed detection probabilities individually and jointly via the average error probability, we found several tradeoffs between the system efficiency and reliability under three different spectrum sensing strategies. For single-user sensing and multi-user sensing with soft information fusion, the tradeoff is between the false alarm diversity gain and the missed detection SNR loss by altering the detection threshold. For multi-user sensing with hard information fusion without information of cooperative user number, there is a tradeoff between the diversities of false alarm and missed detection. In addition, under hard information fusion, with the knowledge of cooperative user number, the soft decision diversity can be achieved at a given missed detection SNR loss. With these diversity and tradeoff results, we derived the optimum threshold in each cooperative strategy to guide practical system design. Simulations have also been presented to further illustrate the analytical results and compare the various cooperative strategies.

3.7 Proof of Theorem 3.2

With $\theta_s = N(1+\frac{1}{\gamma})\log(\gamma+1)$, as $\gamma \to +\infty$, $e^{-\theta_s} = (1+\gamma)^{-N(1+\frac{1}{\gamma})} \sim (\gamma+1)^{-N}$ and $\sum_{i=0}^{N-1} \theta_s^i/i! = \sum_{i=0}^{N-1} N^i(1+\frac{1}{\gamma})^i(\log(\gamma+1))^i/i! \sim N^{N-1}(\log(\gamma+1))^{N-1}/(N-1)!$.

Thus, according to (3.14), $P_{f,s} \sim \frac{N^{N-1}}{(N-1)!}(\log(\gamma+1))^{N-1}(\gamma+1)^{-N}$ and

$$d_{f,s} = -\lim_{\gamma\to+\infty} \frac{\log P_{f,s}}{\log \gamma} = N.$$

Also, with θ_s defined above, as $\gamma \to +\infty$, $e^{-\theta_s/(\gamma+1)} = e^{-\frac{N}{\gamma}\log(\gamma+1)} \to 1$ and $\sum_{i=N}^{+\infty} \theta_s^i/(i!(\gamma+1)^i) = \sum_{i=N}^{+\infty} N^i(1+\frac{1}{\gamma})^i(\log(\gamma+1))^i/(i!(\gamma+1)^i) \sim N^N(\log(\gamma+1))^N(\gamma+1)^{-N}/N!$. Thus, according to (3.15), $P_{md,s} \sim \frac{N^N}{N!}(\log(\gamma+1))^N(\gamma+1)^{-N}$ and $d_{md,s} = -\lim_{\gamma\to+\infty} \frac{\log P_{md,s}}{\log \gamma} = N$. Accordingly, $d_{e,s} = \min(d_{f,s}, d_{md,s}) = N$.

3.8 Proof of Theorem 3.3

From (3.18) and (3.19),

$$P_{f,h} = P(\lambda_h \geq \theta_h | H_0) = \sum_{i=\theta_h}^{N} \binom{N}{i} P_{f,l}^i (1-P_{f,l})^{N-i}.$$

The false alarm diversity at the local sensing decision is $d_{f,l}$, so as $\gamma \to +\infty$, $P_{f,l} \sim (\gamma+1)^{-d_{f,l}}$, $1-P_{f,l} \to 1$, thus $P_{f,l}^i(1-P_{f,l})^{N-i} \sim (\gamma+1)^{-id_{f,l}}$. In the summation of $P_{f,h}$, as $\gamma \to +\infty$, the term with lowest power order of $(\gamma+1)^{-1}$ will dominate, thus $P_{f,h} \sim \binom{N}{\theta_h}(\gamma+1)^{-\theta_h d_{f,l}}$, and the false alarm diversity order is

$$d_{f,h} = -\lim_{\gamma\to+\infty} \frac{\log P_{f,h}}{\log \gamma} = \theta_h d_{f,l}, \quad \theta_h = 1, 2, \ldots, N.$$

From (3.18) and (3.19),

$$P_{md,h} = P(\lambda_h < \theta_h | H_1) = \sum_{i=0}^{\theta_h - 1} \binom{N}{i} (1 - P_{md,l})^i P_{md,l}^{N-i}.$$

From the same argument, the missed detection diversity order is

$$d_{md,h} = -\lim_{\gamma \to +\infty} \frac{\log P_{md,h}}{\log \gamma} = (N - \theta_h + 1) d_{md,l}, \ \theta_h = 1, 2, \ldots, N.$$

3.9 Proof of Corollary 3.3

From Corollary 3.1, we know that the local threshold can be chosen as $\theta_l = d_0 \theta^o$ ($d_0 > 0$). In this case, $d_{f,l} = d_0$ and $d_{md,l} = 1$. By Theorem 3.3, if the hard decision threshold is θ_h ($\theta_h \in \{1, 2, \ldots, N\}$), the average error diversity is $d_{e,h} = \min(\theta_h d_0, N - \theta_h + 1)$. To maximize $d_{e,h}$, we need to maximize both $\theta_h d_0$ and $(N - \theta_h + 1)$ simultaneously. By maximizing the latter, we obtain $\theta_h = 1$. Then $d_{e,h} = \min(d_0, N)$. Thus, as long as $d_0 \geq N$, we obtain the maximum diversity $d_{e,h} = N$. However, as stated in Corollary 3.1, higher d_0 will cause higher SNR loss for the missed detection performance. Thus, we choose $d_0 = N$ to minimize the SNR loss for the missed detection performance while achieving the maximum average error diversity.

CHAPTER 4

COGNITIVE RADIO FOR THE SMART GRID 2: COOPERATIVE SPECTRUM SENSING WITH TERNARY LOCAL DECISIONS

4.1 Motivation

In our previous chapter, the gain of cooperation is quantified in terms of the *cooperative diversities* for missed detection, false alarm and average error probabilities. Using diversity as the performance metric, we optimally designed the sensing threshold strategies for cooperative sensing with both soft information fusion (SCoS) and binary information fusion (BCoS). We found that while SCoS can achieve the maximum diversity, BCoS either loses half of the diversity or achieves the full diversity at the price of some signal-to-noise ratio (SNR) loss.

While the performance of SCoS is desirable, it is impractical since it requires infinite bandwidth for the communications between the sensing users and the fusion center. Intuitively, the performance gap between SCoS and BCoS results from the loss of information with the single-bit local decisions in BCoS. It should be possible to improve the performance by allowing the sensing users to provide more information. In this chapter, we investigate a cooperative sensing scheme with local ternary decisions. While developing the optimum strategies is complicated and mathematically intractable, our focus is to show that with local ternary decisions, it is possible to gain in terms of both diversity and SNR. This is in sharp contrary with the inevitable

diversity-SNR tradeoff when binary local decisions are used. Compared with existing work on cooperative spectrum sensing with multi-threshold local decisions such as [8, 27, 32, 31] and [36], our algorithm provides simple and closed-form expressions for both the local thresholds and fusion rule based on the metric of cooperative diversity. Moreover, we also obtain an explicit analytical expression for the performance gain, which has only been illustrated by simulations in the literature.

The problem formulation, together with the preliminaries of binary (BD) and ternary (TD) local decisions will be introduced in Section 4.2. Then, we will determine the detection fusion rules for TD by first finding the relationship between BD and TD in Section 4.3 and then selecting the detection regions for TD in Section 4.4 with simulation results given in Section 4.5. Finally, concluding remarks will be presented in Section 4.6.

Notation: $x \sim \mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian random variable x with mean μ and variance σ^2. $g(\gamma) \sim f(\gamma)$ means $\lim_{\gamma \to +\infty} \frac{g(\gamma)}{f(\gamma)} = c$ where $c > 0$ is a constant. $g(\gamma) \approx f(\gamma)$ means $\lim_{\gamma \to +\infty} \frac{g(\gamma)}{f(\gamma)} = 1$.

4.2 System Model

4.2.1 Signal Model and Performance Metrics

In the spectrum sensing process, the sensing users observe signals under the following two hypotheses:

$$H_0 : \text{absence of primary user},$$

$$H_1 : \text{presence of primary user}.$$

We assume that the channels between the primary and the sensing users are Rayleigh fading with additive white Gaussian noise (AWGN). Then after normalization, the signal at each sensing users becomes (see e.g. [14, 25]):

$$\begin{aligned} r_i|H_0 &= n_i \sim \mathcal{CN}(0,1), \\ r_i|H_1 &= h_i x + n_i \sim \mathcal{CN}(0,\gamma+1)\ , \end{aligned} \tag{4.1}$$

where γ is the average SNR at the sensing users. With geographically distributed sensing users, it is reasonable to assume that they experience independent fading channels. With this assumption, the received signals for different sensing users r_is are conditionally independent identically distributed (i.i.d.) under each hypothesis. Under this model, the Neyman-Pearson (NP) detector at each secondary user is the energy detector with $\|r_i\|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \theta_l$, where θ_l is the local decision threshold.

In our previous chapter, it has already been shown that the a priori probabilities of the hypotheses do not affect the diversity gains. Therefore, without loss of generality,

we assume that $P(H_0 = 0) = P(H_1 = 1) = \frac{1}{2}$ in this chapter. For the detection problem introduced in Eq. (4.1), there are three performance measures, namely false alarm (P_f), missed detection (P_{md}) and average error (P_e) probabilities. As in Chapter 3, we will use the diversity defined as $d_* = -\lim_{\gamma \to +\infty} \frac{\log P_*}{\log \gamma}$ for each of them.

By definition, diversity only captures the performance at high SNR. Hence, in our analyses, we will aim at achieving better low-SNR performance while maintaining the same diversity gain.

4.2.2 Binary Local Decision (BD) and BCoS-k_0

For BCoS introduced in Chapter 3, the secondary users make local binary decisions $D_i \in \{0, 1\}$ and a fusion center will collect all decisions and make a global decision. The local decisions are:

$$D_i = \begin{cases} 0 & \text{if } 0 \leq \|r_i\|^2 < \theta_{l,B} \\ 1 & \text{if } \|r_i\|^2 > \theta_{l,B} \ . \end{cases} \tag{4.2}$$

If the local decision threshold is $\theta_{l,B} = k_0 \theta^o$, where $\theta^o = (1 + \frac{1}{\gamma}) \log(1 + \gamma)$ is the local optimum threshold [14]. Then, as $\gamma \to +\infty$, $P_{f,l} = e^{-\theta_{l,B}} \approx \gamma^{-k_0}$ and $P_{md,l} = e^{\frac{\theta_{l,B}}{\gamma+1}} \approx k_0 \gamma^{-1}$. With the NP detector $\sum_{i=1}^{N} D_i \underset{<}{\overset{>}{H_0}}{}^{H_1} \theta_{f,B}$, the diversities are $d_{f,B} = k_0 \theta_{f,B}$ and $d_{md,B} = N - \theta_{f,B} - 1$, where N is the total number of secondary users. To jointly optimize both diversities, the fusion threshold can be selected as[1]: $\arg\max_{\theta_{f,B}}(\min(k_0 \theta_{f,B}, N - \theta_{f,B} - 1)) = \frac{N+1}{k_0+1}$. The optimized diversities can be deter-

[1]It should be noticed that at the fusion center, $\sum D_i$'s can only take integer values. However, to simplify the notation, the integer restrictions are neglected without affecting the analysis.

mined accordingly as $d_e = d_f = d_{md} = \frac{k_0}{k_0+1}(N+1)$. This indicates that with larger k_0, BCoS-k_0 can achieve higher diversities by setting larger local threshold. However, in this case, the local decisions have missed detection probabilities $P_{md,B} \approx k_0\gamma^{-1}$ with $-10\log_{10} k_0$ SNR loss and smaller false alarm probabilities $P_{f,B} \approx \gamma^{-k_0}$. The larger missed detection probability P_{md} will dominate the overall average error probabilities performance P_e. Furthermore, instead of the diversity gain, the SNR losses for the missed detection probabilities will dominate the overall average error probability P_e at low-to-medium SNR.

4.2.3 Ternary Local Decision (TD)

The local decisions for TD are:

$$D_i = \begin{cases} 0 & \text{if } 0 \le \|r_i\|^2 < \theta_{l,1} \\ \spadesuit & \text{if } \theta_{l,1} \le \|r_i\|^2 \le \theta_{l,2} \\ 1 & \text{if } \|r_i\|^2 > \theta_{l,2}\ , \end{cases} \tag{4.3}$$

where $\theta_{l,2} > \theta_{l,1}$ are two local decision thresholds and $\spadesuit$ means "not sure". Then, the "0" or "1" decisions[2]are sent to the fusion center for the global decision $D \in \{0, 1\}$.

[2]Note that the sensor will remain silent when the local decision is $\spadesuit$.

Under this local decisions, the conditional probabilities under each hypothesis are:

$$\begin{aligned}
P(D_i = 0|H_0) &= \alpha_1 = 1 - e^{-\theta_{l,1}} \ , \\
P(D_i = \spadesuit|H_0) &= \alpha_2 = e^{-\theta_{l,1}} - e^{-\theta_{l,2}} \ , \\
P(D_i = 1|H_0) &= \alpha_3 = e^{-\theta_{l,2}} \ , \\
P(D_i = 0|H_1) &= \beta_1 = 1 - e^{-\frac{\theta_{l,1}}{\gamma+1}} \ , \\
P(D_i = \spadesuit|H_1) &= \beta_2 = e^{-\frac{\theta_{l,1}}{\gamma+1}} - e^{-\frac{\theta_{l,2}}{\gamma+1}} \ , \\
P(D_i = 1|H_1) &= \beta_3 = e^{-\frac{\theta_{l,2}}{\gamma+1}} \ .
\end{aligned} \tag{4.4}$$

At the fusion center, D_is follow the trinomial distribution as:

$$\begin{aligned}
P(D_1, D_2, \ldots, D_N|H_0) &= \alpha_1^{n_0} \alpha_2^{N-n_0-n_1} \alpha_3^{n_1} \ , \\
P(D_1, D_2, \ldots, D_N|H_1) &= \beta_1^{n_0} \beta_2^{N-n_0-n_1} \beta_3^{n_1} \ ,
\end{aligned} \tag{4.5}$$

where $n_0 = \{\text{the number of } D_i = 0\}$, $n_1 = \{\text{the number of } D_i = 1\}$ and N is the total number of cooperating local detectors. Accordingly, the sufficient statistics is (n_0, n_1). Denoting $\boldsymbol{\mathcal{R}}_1$ as the set of (n_0, n_1) to make global decision $D = 1$ and $\boldsymbol{\mathcal{R}}_0$ vice versa, we have:

$$\begin{aligned}
P_f &= \sum_{(n_0,n_1)\in\boldsymbol{\mathcal{R}}_1} P(n_0, n_1|H_0) \ , \\
P_{md} &= \sum_{(n_0,n_1)\in\boldsymbol{\mathcal{R}}_0} P(n_0, n_1|H_1) \ .
\end{aligned} \tag{4.6}$$

Based on Eqs. (4.4), (4.5) and (4.6), the optimum fusion rule can be obtained by jointly optimizing $P_e = \frac{1}{2}(P_f + P_{md})$ over $\theta_{l,1}$, $\theta_{l,2}$ and $\boldsymbol{\mathcal{R}}_1$. However, not only that this is mathematically intractable, the solution also does not provide any clear insights

on the diversity-SNR tradeoff. As an alternative, we will first find the relationship between fusions with TD and BD and then develop the fusion rule for cooperative sensing with ternary local decisions (TCoS).

4.3 The Link between Fusions with BD and TD

It is worth noting that at the fusion center, BD has a one-dimensional sufficient statistics set with $n_0+n_1=N$ while TD has a two-dimensional set with $n_0+n_1\leq N$. We find that when the fusion center with TD makes a global decision based on only one of n_0 and n_1, it is equivalent to the fusion with BD as the following:

Theorem 4.1 *For cooperative sensing based on local ternary decisions with thresholds $\theta_{l,1}$ and $\theta_{l,2}$:*

1. *If $\boldsymbol{\mathcal{R}}_1=\{(n_0,n_1):n_1\geq\theta_t\}$, this TD fusion is equivalent to BD fusion with local threshold $\theta_{l,B}=\theta_{l,2}$ and fusion threshold $\theta_{f,B}=\theta_t$;*

2. *If $\boldsymbol{\mathcal{R}}_0=\{(n_0,n_1):n_0\geq N-\theta_t+1\}$, this TD fusion is equivalent to BD fusion with local threshold $\theta_{l,B}=\theta_{l,1}$ and fusion threshold $\theta_{f,B}=\theta_t$;*

Proof. If $\boldsymbol{\mathcal{R}}_1=\{(n_0,n_1):n_1\geq\eta_t\}$, then:

$$\begin{aligned}P_{f,t}=&\sum_{n_1=\eta_t}^{N}\sum_{n_0=0}^{N-n_1}\frac{N!}{n_0!(N-n_0-n_1)!n_1!}\alpha_1^{n_0}\alpha_2^{N-n_0-n_1}\alpha_3^{n_1}\\=&\sum_{n_1=\eta_t}^{N}\frac{N}{n_1!(N-n_1)!}(1-\alpha_3)^{N-n_1}\alpha_3^{n_1}\ ,\end{aligned}$$

and

$$P_{md,t}=\sum_{n_1=0}^{\eta_t-1}\sum_{n_0=0}^{N-n_1}\frac{N!}{n_0!(N-n_0-n_1)!n_1!}\beta_1^{n_0}\beta_2^{N-n_0-n_1}\beta_3^{n_1}$$
$$=\sum_{n_1=0}^{\eta_t-1}\frac{N}{n_1!(N-n_1)!}(1-\beta_3)^{N-n_1}\beta_3^{n_1}\ .$$

This is equivalent to BT with $\eta_{l,B}=\eta_{l,2}$ and $\eta_{f,B}=\eta_t$.

If $\boldsymbol{\mathcal{R}}_0=\{(n_0,n_1):n_0\geq N+1-\eta_t\}$, then:

$$P_{f,t}=\sum_{n_0=0}^{N-\eta_t+1}\sum_{n_1=0}^{N-n_0}\frac{N!}{n_0!(N-n_0-n_1)!n_1!}\alpha_1^{n_0}\alpha_2^{N-n_0-n_1}\alpha_3^{n_1}$$
$$=\sum_{n_0=0}^{N-\eta_t+1}\frac{N}{n_0!(N-n_0)!}\alpha_1^{n_0}(1-\alpha_1)^{N-n_0}\ ,$$

and

$$P_{md,t}=\sum_{n_0=N-\eta_t+1}^{N}\sum_{n_1=0}^{N-n_1}\frac{N!}{n_0!(N-n_0-n_1)!n_1!}\beta_1^{n_0}\beta_2^{N-n_0-n_1}\beta_3^{n_1}$$
$$=\sum_{n_0=N-\eta_t+1}^{N}\frac{N}{n_0!(N-n_0)!}\beta_1^{n_0}(1-\beta_1)^{N-n_0}\ .$$

This is equivalent to BT with $\eta_{l,B}=\eta_{l,1}$ and $\eta_{f,B}=\eta_t$. ■

4.4 TCoS Fusion Rule

With the relationship between BD and TD established in Theorem 4.1, we will next develop the fusion rule for TCoS. In particular, with local sensing thresholds $\theta_{l,1}=k_1\theta^o$ and $\theta_{l,2}=k_2\theta^o$ and $k_1<k_2$, the corresponding sensing strategy is termed as TCoS-k_1-k_2.

Recall that from Section 4.2.2, BCoS-k_1 has a smaller diversity order. On the other hand, BCoS-k_2 achieves larger diversity but suffers from the SNR loss with the missed

detection probability. Therefore, here we try to improve BCoS-k_2 missed detection performance by moving part of the decision region $\boldsymbol{\mathcal{R}}_0$ to $\boldsymbol{\mathcal{R}}_1$ while maintaining its larger false alarm and hence the overall diversity.

With TCoS-k_1-k_2, the probabilities for local decisions are: $\alpha_1 \approx 1-\gamma^{-k_1}, \alpha_2 \approx \gamma^{-k_1}, \alpha_3 \approx \gamma^{-k_2}$ and $\beta_1 \approx k_1\gamma^{-1}, \beta_2 \approx (k_2-k_1)\gamma^{-1}, \beta_3 \approx 1-k_2\gamma^{-1}$. Accordingly,

$$P(n_0,n_1|H_0) \approx \alpha_2^{N-n_0-n_1}\alpha_3^{n_1} \sim \gamma^{-(k_1N-k_1n_0+(k_2-k_1)n_1)}, \tag{4.7}$$

$$P(n_0,n_1|H_1) \approx \beta_1^{n_0}\beta_2^{N-n_0-n_1} \sim \gamma^{-(N-k_1)}. \tag{4.8}$$

The false alarm probability can be then calculated as:

$$\begin{aligned} P_f &= \sum_{(n_0,n_1)\in\boldsymbol{\mathcal{R}}_1} P(n_0,n_1|H_0) \\ &\sim \sum_{(n_0,n_1)\in\boldsymbol{\mathcal{R}}_1} \gamma^{-(k_1N-k_1n_0+(k_2-k_1)n_1)}. \end{aligned} \tag{4.9}$$

By Theorem 4.1, the decision region corresponding to BCoS-k_2 is $\boldsymbol{\mathcal{R}}_1 = \{(n_0,n_1) : n_1 \geq \frac{N+1}{k_2+1}\}$, $\boldsymbol{\mathcal{R}}_0 = \{(n_0,n_1) : n_1 < \frac{N+1}{k_2+1}\}$ and $d_{f,\mathrm{BCoS}-k_2} = \frac{k_2}{k_2+1}(N+1)$. In order to reduce the missed detection probability which dominates the average error performance, we want to increase the decision region of $\boldsymbol{\mathcal{R}}_1$, or equivalently decrease $\boldsymbol{\mathcal{R}}_0$. At the same time, however, the false alarm diversity should be preserved. From Eq. (4.7), if $k_1N-k_1n_0+(k_2-k_1)n_1 \geq \frac{k_2}{k_2+1}(N+1)$, or equivalently $n_0 \leq (k_2-k_1)n_1 + \frac{k_1(k_2+1)N-k_2(N+1)}{k_2+1}$, $P(n_0,n_1|H_0)$ will have a larger exponent of γ^{-1} than $d_{f,\mathrm{BCoS}-k_2}$. Therefore, all points in $\boldsymbol{\mathcal{R}}_0$ satisfying $n_0 \leq (k_2-k_1)n_1 + \frac{k_1(k_2+1)N-k_2(N+1)}{k_2+1}$ can be moved into $\boldsymbol{\mathcal{R}}_1$ without affecting the false alarm diversity according to (4.9).

As a result, the boundary between $\boldsymbol{\mathcal{R}}_1$ and $\boldsymbol{\mathcal{R}}_0$ of the resultant fusion rule is the line $n_0 = (\frac{k_2}{k_1}-1)n_1 + \frac{k_1(k_2+1)N-k_2(N+1)}{k_1(k_2+1)}$, which starts from $(n_0, n_1) = \left(\frac{k_1(k_2+1)N-k_2(N+1)}{k_1(k_2+1)}, 0\right)$ and ends at $(n_0, n_1) = \left(\frac{k_2N-1}{k_2+1}, \frac{N+1}{k_2+1}\right)$ where $\frac{k_2N-1}{k_2+1} + \frac{N+1}{k_2+1} = N$.

The decision region for TCoS-k_1-k_2 is illustrated in Fig. 4.1. The bold line is the boundary between $\boldsymbol{\mathcal{R}}_0$ and $\boldsymbol{\mathcal{R}}_1$ for TCoS-k_1-k_2 while the dashed line is the boundary corresponding to BCoS-k_2.

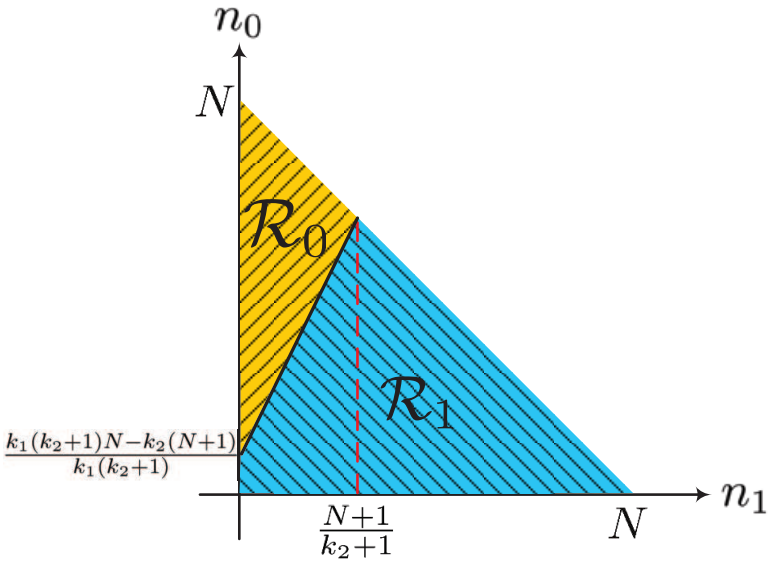


Figure 4.1: The decision region for TCoS-k_1-k_2 with the points at the boundary belonging to $\boldsymbol{\mathcal{R}}_1$.

Compared with BCoS-k_2, TCoS-k_1-k_2 will provide the same overall diversities with $d_e = d_{md} = d_f = \frac{N+1}{k_2+1}$. Denoting $\Delta\mathcal{R} = \{(n_0, n_1) : 0 \leq n_1 < \frac{N+1}{k_2+1}, 0 \leq n_0 \leq (k_2 - k_1)n_1 + \frac{k_1(k_2+1)N-k_2(N+1)}{k_2+1}\}$, the difference between the missed detection probabilities of TCoS-k_1-k_2 and BCoS-k_2 is

$$\Delta P_{md} = P_{md,\text{TCoS}-k_1-k_2} - P_{md,\text{BCoS}-k_2} = -\sum_{\Delta\mathcal{R}} \beta_1^{n_0} \beta_2^{N-n_0-n_1} \beta_3^{n_1} \ , \tag{4.10}$$

and the difference between the false alarm probabilities of TCoS-k_1-k_2 and BCoS-k_2

is

$$\Delta P_f = P_{f,\text{TCoS}-k_1-k_2} - P_{f,\text{BCoS}-k_2} = \sum_{\Delta\mathcal{R}} \alpha_1^{n_0} \alpha_2^{N-n_0-n_1} \alpha_3^{n_1} \ . \tag{4.11}$$

It should be noticed that $\Delta P_e = P_{e,\text{TCoS}-k_1-k_2} - P_{e,\text{BCoS}-k_2} = \frac{1}{2}(\Delta P_{md} + \Delta P_f) < 0$. Therefore, we obtain overall performance gain over BCoS-k_2. This once again confirms that TCoS not only keeps the diversity gain that captures high SNR performance but also improves the SNR gain that characterizes the low-to-medium.

4.5 Simulations

Although it is possible to optimize k_1 for any given k_2 to obtain the maximum performance gain of TCoS by the analytical expressions given in Section 4.4, the complexity is high. Therefore, we opt to verify and demonstrate the performance gain of TCoS over BCoS using some simple k_1 and k_2 values.

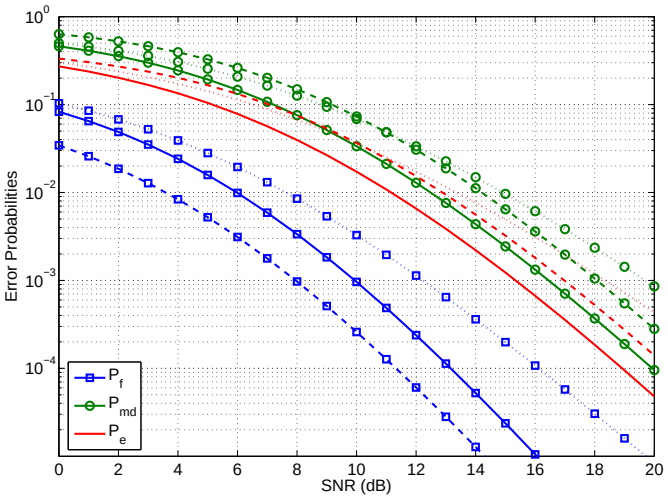


Figure 4.2: TCoS-1-2 (solid) vs. BCoS-1 (dotted) and BCoS-2 (dashed) with $N = 5$.

To illustrate the performance gain of TCoS-k_1-k_2 over both BCoS-k_1 and BCoS-k_2, we simulate TCoS-1-2 with 5 cooperating users and compare its performance with

BCoS-1 and BCoS-2 in Fig. 4.2. We see that BCoS-2 exhibit a higher diversity than BCoS-1, but a worse performance at low SNR. Our proposed TCoS-1-2 not only retains the higher diversity of BCoS-2 but also has better performance at low SNR.

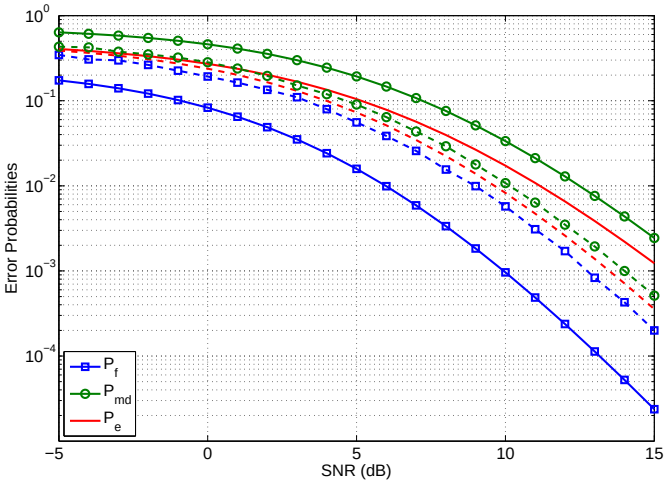


Figure 4.3: TCoS-1-2 (solid) vs. optimal TCoS by exhaustive search (dashed) with $N = 5$.

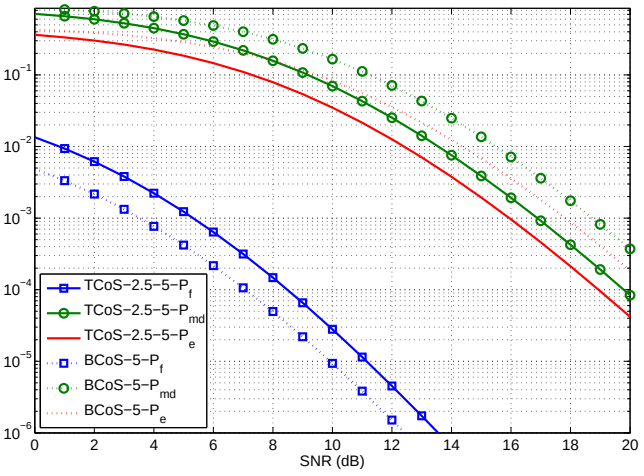


Figure 4.4: TCoS-2.5-5 vs. BCoS-5 with $N = 5$.

In Fig. 4.3, the performance of our proposed TCoS-1-2 is compared with the optimal TCoS with exhaustive search. It can be seen that our TCoS-1-2 only sacrifices a little performance (≈ 1.5 dB) in exchange for the low-complexity closed-form local

and global decision rules. In addition, we know that BCoS-N achieves the maximum diversity but suffers from considerable SNR loss. Here, we compare the performance of TCoS-2.5-5 and BCoS-5 in Fig. 4.4. It can be observed that TCoS-2.5-5 also achieves the same full diversity ($d_e = 5$), but has about 2dB SNR gain. Together with Fig. 4.2, it is confirmed that the overall SNR gain is obtained without losing any false alarm diversity.

4.6 Concluding Remarks

In this chapter, we proposed cooperative sensing with ternary local decisions (TCoS) to improve upon binary hard decisions (BCoS) by gaining SNR while maintaining the same diversity. The link between the fusion with BD and TD has been established and further used to determine the fusion rule for TCoS. The algorithm developed in this chapter provides simple and closed-form expressions for both the local decision thresholds and the fusion rule. The performance gain was also derived analytically. Furthermore, simulations confirmed that, as the middle ground between BCoS and SCoS, TCoS provides a practical yet effective solution for the inevitable diversity-SNR tradeoff encountered by BCoS.

CHAPTER 5

SMART GRID MONITORING USING SYNCHROPHASOR MEASUREMENTS: STATE ESTIMATION WITH BAD DATA

5.1 Motivation

With synchronization from Global Positioning System (GPS) satellite signals, direct, accurate and synchronized measurement of the voltage and current phasors in the power system becomes feasible with phasor measurement units (PMU) [40]. PMUs are widely installed in current power systems to form wide-area measurement systems (WAMS) for better power grid monitoring. Among various objectives for WAMS, wide-area state estimation (SE) is a very important component of the supervisory control and data acquisition (SCADA) system and the energy management system (EMS). The synchronized phasor measurements provided by PMUs can be beneficial to the system state estimation process [16]. In conventional power systems, only the real and reactive power measurements are available and the state estimate can only be obtained from an iterative method [67]. With PMU measurements, the state estimate can be obtained linearly [41].

While PMU measurements are expected to significantly improve SE accuracy, there is little literature on SE from PMU measurements with bad data. Many just use the largest residual removal (LRR) method in conventional SE (see e.g. [42, Chapter 7]). However, with the linear signal model enabled by PMU measurements, more

sophisticated algorithms become feasible. In this chapter, we develop and analyze several algorithms by explicitly incorporating the bad data into our system model. To deal with the bad data, one can either estimate the locations first, or estimate their locations and values simultaneously. With the former approach, the determined bad data locations can be directly used to simply remove the contaminated measurements, or they can facilitate the estimation of the bad data values and subtract them from the measurements. We show that these will result in identical state estimators. Thus, bad data removal is actually a simplified form of bad data subtraction with separate location and value estimations.

Before proposing our algorithms for SE with bad data, we first show that the traditional LRR method has no performance assurance when the data redundancy is not enough. Then, we rewrite the joint bad data and state estimation problem into a sparsity-regularized minimization problem (SRM), which is further transformed into a compressive sensing problem facilitating application of various algorithms in the literature [49]. This algorithm is trying to estimate the locations and values of bad data simultaneously. Next, we propose the projection and minimization (PM) algorithm to estimate the bad data location first and then remove those bad data. Simulations are presented to compare these algorithms. Results show that our PM algorithm provides the best performance among all, and that it perfectly locates the bad data in the case of a single occurrence.

Notations: Bold capital letters denote matrices; bold lowercase letters denote column vectors; $\boldsymbol{A}(i,:)$ and $\boldsymbol{A}(:,i)$ denote the i-th row and column of matrix $\boldsymbol{A}$, respectively;

$\boldsymbol{A}^*$, $\boldsymbol{A}'$ and $\boldsymbol{A}^{-1}$ denotes the conjugate, conjugate transpose and inverse of $\boldsymbol{A}$, respectively; $\boldsymbol{I}_{p\times p}$ denotes a $p \times p$ identity matrix; $\|\boldsymbol{a}\|_n$ denotes the L_n-norm of vector $\boldsymbol{a}$; $\boldsymbol{A}(i,:) = [\ ]$ indicates removal of the i-th row of matrix $\boldsymbol{A}$.

5.2 System Model

With PMU measurements, the signal model is given by:

$$\boldsymbol{m} = \begin{bmatrix} \boldsymbol{m}_v \\ \\ \boldsymbol{m}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} \\ \\ \boldsymbol{Y} \end{bmatrix} \boldsymbol{s} + \boldsymbol{e} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{b} + \boldsymbol{\eta} \tag{5.1}$$

where $\boldsymbol{m}_v$ and $\boldsymbol{m}_i$ are p voltage and $q-p$ current synchronized phasor measurements, respectively; $\boldsymbol{s}$ is the state of the power grid which contains the p bus voltage phasors; $\boldsymbol{Y}$ is the admittance matrix which is determined by the power grid structure and the transmission line parameters [1, Chapter 1]; $\boldsymbol{e}$ is the measurement error vector composed of the measurement device noise $\boldsymbol{\eta}$ and possible high-magnitude bad data $\boldsymbol{b}$ due to communication errors or equipment failures. Since communication errors or equipment failures are rare, we expect $\boldsymbol{b}$ to be a sparse vector, i.e., $\|\boldsymbol{b}\|_0 \ll p$.

Without loss of generality, we assume entries of $\boldsymbol{\eta}$ follow the i.i.d. proper complex Gaussian distribution with variance σ^2. Then, assuming absence of bad data $\boldsymbol{b}$, the maximum likelihood estimator of the system state is the least-squares (LS) estimator [41]: $\hat{\boldsymbol{s}}^{ML} = \hat{\boldsymbol{s}}^{LS} = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m}$. With knowledge of the noise variance σ^2, bad data presence can be readily detected by the χ^2-test [42, Chapter 7], which declares bad data occurrence when $\|\boldsymbol{H}\hat{\boldsymbol{s}}^{LS} - \boldsymbol{m}\|_2^2/\sigma^2 > \chi^2_{1-\alpha,2(q-p)}$, where $\chi^2_{1-\alpha,2(q-p)}$ is the tail value

of $1-\alpha$ for χ^2-distribution with $2(q-p)$ degrees of freedom and α is the detection confidence level.

Once the bad data presence is detected, there are two ways to deal with them. One can either start by determining locations of the bad data, or estimating their locations and values simultaneously. With the former approach, the determined bad data locations can be directly used to simply remove the contaminated measurements, or they can facilitate the estimation of the bad data values which can be then subtracted from the measurements. In the next section, we will first analyze and compare these two options.

5.3 Removal or Subtraction?

Let $\mathcal{I}_b$ denote the bad data locations with $|\mathcal{I}_b| = k$. The bad data vector becomes $\boldsymbol{b} = \boldsymbol{\Theta}\boldsymbol{b}_k$ where vector $\boldsymbol{b}_k$ contains all k non-zero bad data entries from $\boldsymbol{b}$ and $\boldsymbol{\Theta} = [\boldsymbol{I}(:, i_1), \ldots, \boldsymbol{I}(:, i_k)]$ is the $q \times k$ selection matrix satisfying $\boldsymbol{\Theta}'\boldsymbol{\Theta} = \boldsymbol{I}_k$. Then, the signal model in Eq. (5.1) becomes:

$$\boldsymbol{m} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{\Theta}\boldsymbol{b}_k + \boldsymbol{\eta} \tag{5.2}$$

To remove the k contaminated measurements, pre-multiply $\boldsymbol{m}$ with the $(q-k) \times q$ bad data removal matrix $\boldsymbol{\Theta}_\perp = (\text{null}(\boldsymbol{\Theta}^T))^T$, which evidently satisfies $\boldsymbol{\Theta}_\perp\boldsymbol{\Theta} = \boldsymbol{0}$ and $\boldsymbol{\Theta}\boldsymbol{\Theta}' + \boldsymbol{\Theta}'_\perp\boldsymbol{\Theta}_\perp = \boldsymbol{I}_{q\times q}$. Accordingly, the state estimate based on bad data removal is $\hat{\boldsymbol{s}}_R = \arg\min_{\boldsymbol{s}} \|\boldsymbol{\Theta}_\perp\boldsymbol{m} - \boldsymbol{\Theta}_\perp\boldsymbol{H}\boldsymbol{s}\|_2^2 = (\boldsymbol{H}'\boldsymbol{\Theta}'_\perp\boldsymbol{\Theta}_\perp\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{\Theta}'_\perp\boldsymbol{\Theta}_\perp\boldsymbol{m}$. To formulate the state estimate based on bad data subtraction, one starts with a conditional estimate of

the bad data $\hat{\boldsymbol{b}}(\boldsymbol{s}) = \boldsymbol{\Theta}'(\boldsymbol{m}-\boldsymbol{H}\boldsymbol{s})$ which comes from $\boldsymbol{\Theta}'\boldsymbol{m} = \boldsymbol{\Theta}'\boldsymbol{H}\boldsymbol{s}+\boldsymbol{b}_k+\boldsymbol{\Theta}'\eta$. Hence, the state estimate after bad data subtraction can be obtained as: $\hat{\boldsymbol{s}}_S = \arg\min_{\boldsymbol{s}}\|\boldsymbol{m}-\boldsymbol{\Theta}\hat{\boldsymbol{b}}_k(\boldsymbol{s})-\boldsymbol{H}\boldsymbol{s}\|_2^2$. The objective function can be rewritten as: $\|(\boldsymbol{I}-\boldsymbol{\Theta}\boldsymbol{\Theta}')(\boldsymbol{m}-\boldsymbol{H}\boldsymbol{s})\|_2^2 = \|\boldsymbol{\Theta}_\perp'\boldsymbol{\Theta}_\perp(\boldsymbol{m}-\boldsymbol{H}\boldsymbol{s})\|_2^2 = \|\boldsymbol{\Theta}_\perp(\boldsymbol{m}-\boldsymbol{H}\boldsymbol{s})\|_2^2$. As a result, $\hat{\boldsymbol{s}}_S = \hat{\boldsymbol{s}}_R$. Interestingly, these estimators are also equivalent to the joint state and bad data estimator given in [6, Eq. (32)]. The reason is that the oblique pseudo-inverse $(\boldsymbol{H}'\boldsymbol{\Theta}_\perp'\boldsymbol{\Theta}_\perp\boldsymbol{H})^{-1} = (\boldsymbol{H}'\boldsymbol{P}_{\boldsymbol{\Theta}_\perp}\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{P}_{\boldsymbol{\Theta}_\perp}$ solves the LS problem and $\boldsymbol{P}_{\boldsymbol{\Theta}_\perp}$ is bad measurement removal.

The preceding analyses assume perfect knowledge of the bad data locations. However, when only the estimate $\hat{\boldsymbol{\Theta}}$ is available, the removal-subtraction equivalence still holds since the objective functions of both estimators will remain the same except that $\boldsymbol{\Theta}$ is replaced by $\hat{\boldsymbol{\Theta}}$.

Due to this equivalence, we will not distinguish the bad data removal and subtraction algorithms in our ensuing analyses and comparisons.

5.4 State Estimation with Bad Data Presence

5.4.1 Largest Residual Removal (LRR)

Traditionally, the measurement with the largest residual is usually considered bad and removed in power grid SE as described in Algorithm 1 [42]. However, as we will show next, the largest residual location is not always that of the bad data even for a single bad data occurence. Considering only a single bad data and the LS state

estimator, the measurement residual is:

$$\begin{aligned}\boldsymbol{r} &= \boldsymbol{m} - \boldsymbol{H}\hat{\boldsymbol{s}} = \boldsymbol{m} - \boldsymbol{H}(\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m} \\ &= (\boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}')\boldsymbol{m} \\ &= \boldsymbol{P}_H^{\perp}\boldsymbol{m} = \boldsymbol{P}_H^{\perp}(\boldsymbol{H}\boldsymbol{s} + \boldsymbol{e}) = \boldsymbol{P}_H^{\perp}\boldsymbol{e}\end{aligned}$$

where $\boldsymbol{P}_H^{\perp} = \boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'$. In the absence of additive noise ($\boldsymbol{\eta} = \boldsymbol{0}$), we have $\boldsymbol{r} = \boldsymbol{P}_H^{\perp}\boldsymbol{b} = \boldsymbol{P}_H^{\perp}\boldsymbol{\Theta}b_i = b_i\boldsymbol{P}_H^{\perp}(:, i)$, where i is the bad data location. Then, with the LRR algorithm, the bad data location will be determined as $\hat{i} = \arg\max_j \|\boldsymbol{P}_H^{\perp}(j, i)\|_2$. The relationship between $\hat{i}$ and i relies on the property of $\boldsymbol{P}_H^{\perp}$.

By singular-value decomposition (SVD), $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $q \times q$ and $p \times p$ orthogonal matrices, respectively and $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Lambda}_{p\times p} & \boldsymbol{0}_{p\times(q-p)} \\ \boldsymbol{0}_{(q-p)\times p} & \boldsymbol{0}_{(q-p)\times(q-p)} \end{bmatrix}$ with $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ and λ_i being the singular values of $\boldsymbol{H}$. Since $\mathrm{rank}(\boldsymbol{H}) = p$, we know that $\forall i, \lambda_i \neq 0$. Then:

$$\begin{aligned}\boldsymbol{P}_H^{\perp} &= \boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}' \\ &= \boldsymbol{I} - (\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V})(\boldsymbol{V}'\boldsymbol{\Sigma}'\boldsymbol{U}'\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V})^{-1}\boldsymbol{V}'\boldsymbol{\Sigma}'\boldsymbol{U}' \\ &= \boldsymbol{I} - \boldsymbol{U}\begin{bmatrix} \boldsymbol{I}_{p\times p} & \boldsymbol{0}_{p\times(q-p)} \\ \boldsymbol{0}_{(q-p)\times p} & \boldsymbol{0}_{(q-p)\times(q-p)} \end{bmatrix}\boldsymbol{U}' \\ &= \boldsymbol{U}\begin{bmatrix} \boldsymbol{0}_{p\times p} & \boldsymbol{0}_{p\times(q-p)} \\ \boldsymbol{0}_{(q-p)\times p} & -\boldsymbol{I}_{(q-p)\times(q-p)} \end{bmatrix}\boldsymbol{U}'\end{aligned}$$

Accordingly, $\boldsymbol{P}_H^{\perp}(j,i) = \sum_{k=p+1}^{q} \boldsymbol{U}(j,k)\boldsymbol{U}^*(k,i)$. Notice that:

$$\sum_{k=1}^{q} \boldsymbol{U}(j,k)\boldsymbol{U}^*(k,i) = \begin{cases} 1, \text{ for } j = i \\ 0, \text{ for } j \neq i \end{cases}$$

due to the orthogonality of $\boldsymbol{U}$. Then,

$$\boldsymbol{P}_H^{\perp}(i,j) = \begin{cases} -1 + \sum_{k=1}^{p} \boldsymbol{U}(j,k)\boldsymbol{U}^*(k,i), \text{ for } j = i \\ \sum_{k=1}^{p} \boldsymbol{U}(j,k)\boldsymbol{U}^*(k,i), \text{ for } i \neq j \end{cases} .$$

When $q \gg p$, we expect $\sum_{k=1}^{p} \boldsymbol{U}(j,k)\boldsymbol{U}^*(k,i)$ to be small. Then, $\|\boldsymbol{P}_H^{\perp}(i,i)\|_2 \approx 1$ and $\|\boldsymbol{P}_H^{\perp}(j,i)\|_2 \approx 0$ for $j \neq i$. In this case, $\hat{i} = \arg\max_j \|\boldsymbol{P}_H^{\perp}(j,i)\|_2 = i$. However, when q is not large enough, this argument obviously does not hold. Thus, we conclude that LRR can not guarantee the performance unless the measurements are sufficiently redundant.

Algorithm 1: Largest Residual Removal (LRR) [42]

Input: The measurement matrix $\boldsymbol{H}$ with size $q \times p$ and $q > p$, the PMU measurements $\boldsymbol{m}$ with size $q \times 1$, the noise variance σ^2, the confidence level α

Result: The state estimate $\hat{\boldsymbol{s}}$

Initialization: $\hat{\boldsymbol{s}} = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m}$;

while $\|\boldsymbol{H}\hat{\boldsymbol{s}} - \boldsymbol{m}\|_2^2/\sigma^2 > \chi^2_{1-\alpha,2q}$ **do**

  Calculate the residual: $\boldsymbol{r} = \boldsymbol{m} - \boldsymbol{H}\hat{\boldsymbol{s}}$;

  Find the largest residue index: $\hat{i} = \arg\max_i \|\boldsymbol{r}(i)\|_2$;

  Bad date removal and parameter update:

  $\boldsymbol{H}(\hat{i},:) = [\ ]$, $\boldsymbol{m}(\hat{i}) = [\ ]$ and $q = q - 1$;

  if $r(\boldsymbol{H}) = p$ **then** *Update* $\hat{\boldsymbol{s}} = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m}$;

  else break;

end

5.4.2 Sparsity Regularized Minimization (SRM)

Accounting for the sparsity of bad data $\boldsymbol{b}$, the joint estimation of bad data and state can be formulated as a sparsity regularized minimization problem as follows:

$$(\hat{\boldsymbol{s}}, \hat{\boldsymbol{b}}) = \arg\min_{(\boldsymbol{s},\boldsymbol{b})} \left(\|\boldsymbol{m} - \boldsymbol{b} - \boldsymbol{H}\boldsymbol{s}\|_2^2 + \lambda \cdot \mathrm{spar}(\boldsymbol{b}) \right) \tag{5.3}$$

where $\mathrm{spar}(\boldsymbol{b}) = \|\boldsymbol{b}\|_0$, but can be replaced with any other approximating sparsity measures such as $\|\boldsymbol{b}\|_p$ with $p \leq 1$.

With the conditional estimate $\hat{\boldsymbol{s}}(\boldsymbol{b}) = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'(\boldsymbol{m} - \boldsymbol{b})$, Eq. (5.3) can be rewritten as:

$$\begin{aligned} \hat{\boldsymbol{b}} &= \arg\min_{\boldsymbol{b}} \left(\|\boldsymbol{m} - \boldsymbol{b} - \boldsymbol{H}(\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'(\boldsymbol{m} - \boldsymbol{b})\|_2^2 + \lambda\, \mathrm{spar}(\boldsymbol{b}) \right) \\ &= \arg\min_{\boldsymbol{b}} \left(\|\boldsymbol{P}_H^{\perp}(\boldsymbol{b} - \boldsymbol{m})\|_2^2 + \lambda\, \mathrm{spar}(\boldsymbol{b}) \right) \end{aligned} \tag{5.4}$$

where $\boldsymbol{P}_H^{\perp}\boldsymbol{m}$ is the residual defined in Section 5.4.1.

However, Eq. (5.4) does not facilitate direct adoption of existing sparsity regularized algorithm because $\boldsymbol{P}_H^{\perp}$ is a square matrix. To deal with this, recall that $\mathrm{rank}(\boldsymbol{P}_H^{\perp}) = q - p$. Hence, by QR decomposition [22], we can obtain $\boldsymbol{P}_H^{\perp} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{H}_{\perp(q-p)\times q} \\ \\ \boldsymbol{0}_{p\times p} \end{bmatrix}$, where $\boldsymbol{Q}$ is an orthogonal matrix and $\boldsymbol{H}_{\perp}\boldsymbol{H} = \boldsymbol{0}$ since $\boldsymbol{P}_H^{\perp}\boldsymbol{H} = \boldsymbol{0}$. Then, we have $\left\|\boldsymbol{P}_H^{\perp}(\boldsymbol{b} - \boldsymbol{m})\right\|_2^2 = \left\|\boldsymbol{Q}'\left(\boldsymbol{P}_H^{\perp}(\boldsymbol{b} - \boldsymbol{m})\right)\right\|_2^2 = \left\| \begin{bmatrix} \boldsymbol{H}_{\perp}\boldsymbol{b} \\ \\ \boldsymbol{0} \end{bmatrix} - \begin{bmatrix} \boldsymbol{H}_{\perp}\boldsymbol{m} \\ \\ \boldsymbol{0} \end{bmatrix} \right\|_2^2 = \|\boldsymbol{H}_{\perp}(\boldsymbol{b} - \boldsymbol{m})\|_2^2$. Thus, the

problem defined in Eq. (5.4) can be reformulated as:

$$\hat{\boldsymbol{b}} = \arg\min_{\boldsymbol{b}} \left(\|\boldsymbol{H}_{\perp}(\boldsymbol{b}-\boldsymbol{m})\|_2^2 + \lambda \, \mathrm{spar}(\boldsymbol{b}) \right) \tag{5.5}$$

where $\boldsymbol{H}_{\perp}\boldsymbol{m}$=$\boldsymbol{z}$ can be interpreted as the syndrome in the channel decoding problem [51]. Notice that Eq. (5.5) is essentially the compressive sensing formulation [5]. It is well known that with l_0-norm regularization, it is an NP hard problem to obtain the optimal solution. However, plenty of algorithms are proposed in the literature to find suboptimal solutions with reasonable complexity by using different sparsity measures of $\boldsymbol{b}$. Alternative forms of this problem also exist, which include the sparsity constrained minimization and sparsity minimization with squared error constraint [49].

Algorithm 2: Bad Data Estimation with Sparsity Regularized Minimization (SRM)

Input: The measurement matrix $\boldsymbol{H}$ with size $q \times p$ and $q > p$, the PMU measurements $\boldsymbol{m}$ with size $q \times 1$, the noise variance σ^2, the confidence level α

Result: The state estimate $\hat{\boldsymbol{s}}$

Initialization: $\hat{\boldsymbol{s}} = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m}$;

if $\|\boldsymbol{H}\hat{\boldsymbol{s}} - \boldsymbol{m}\|_2^2 > \chi^2_{1-\alpha,2q}$ **then**

$\boldsymbol{H}_{\perp} = (\mathrm{null}(\boldsymbol{H}^T))^T$; syndrome: $\boldsymbol{z} = \boldsymbol{H}_{\perp}\boldsymbol{m}$

Find: $\hat{\boldsymbol{b}} = \arg\min_{\boldsymbol{b}} \left(\|\boldsymbol{H}_{\perp}\boldsymbol{b} - \boldsymbol{z}\|_2^2 + \lambda \, \mathit{spar}(\boldsymbol{b}) \right)$

using algorithms for spectrum sensing problems;

Update the measurement: $\boldsymbol{m} = \boldsymbol{m} - \hat{\boldsymbol{b}}$;

Update the state estimation: $\hat{\boldsymbol{s}} = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m}$;

finish

else finish;

5.4.3 Projection and Minimization

Algorithm 2 simultaneously estimates both the locations and values of the bad data. However, one can also estimate the bad data location and values separately and we have shown that with this strategy the location estimate alone dictates the performance. Thus, we will develop an algorithm only to determine bad data locations and remove them recursively based on the objective function in Eq. (5.5).

To locate the bad data one by one, we restrict $\|\boldsymbol{b}\|_0 = 1$ at each step, i.e., $\boldsymbol{b} = b\boldsymbol{I}(:,i)$, then $(\hat{b},\hat{i}) = \text{argmin}_{(b,i)}\|b\boldsymbol{H}_\perp \boldsymbol{I}(:,i) - \boldsymbol{z}\|_2^2 = \text{argmin}_{(b,i)}\|b\boldsymbol{H}_\perp(:,i) - \boldsymbol{z}\|_2^2$,where $\hat{b}(i)$ can be found by projecting vector $\boldsymbol{z}$ onto $\boldsymbol{H}_\perp(:,i)$ and the process is described in Algorithm 3. It is worth noticing that unlike existing LRR, with single bad data occurance, this algorithm guarantees the correct bad data removal when noise is absent, since it is equivalent to exhaustive search over all bad data positions in Eq. (5.2).

Algorithm 3: Projection and Minimization (PM)

Input: The measurement matrix $\boldsymbol{H}$ with size $q \times p$ and $q > p$, the PMU measurements $\boldsymbol{m}$ with size $q \times 1$, the noise variance σ^2, the confidence level α

Result: The state estimate $\hat{\boldsymbol{s}}$

Initialization: $\hat{\boldsymbol{s}} = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m}$; ;

while $\|\boldsymbol{H}\hat{\boldsymbol{s}} - \boldsymbol{m}\|_2^2 > \chi^2_{1-\alpha,2q}$ **do**

  Estimate the bad data index:

  $\boldsymbol{H}_\perp = (\text{null}(\boldsymbol{H}^T))^T$; syndrome: $\boldsymbol{z} = \boldsymbol{H}_\perp \boldsymbol{m}$

  for $i = 1$ *to* q **do**

    Calculate: $\hat{b}(i) = (\boldsymbol{H}_\perp(:,i))'\boldsymbol{z}/\|\boldsymbol{H}_\perp(:,i)'\|_2^2$;

    $f_i = \|\hat{b}(i)\boldsymbol{H}_\perp(:,i) - \boldsymbol{z}\|_2^2$

  end

  $\hat{i} = \arg\min_i f_i$;

  Bad data removal and parameter update:

  $\boldsymbol{H}(\hat{i},:) = [\ ]$, $\boldsymbol{m}(\hat{i}) = [\ ]$, and $q = q - 1$;

  if $r(\boldsymbol{H}) = p$ **then** *Update:* $\hat{\boldsymbol{s}} = (\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{m}$;

  else break;

end

5.5 Simulations

We use the IEEE 14-bus system in [62] for our simulations and the diagram is drawn in Fig. 5.1. The noise variance is kept constant with per sample SNR of 20dB. Among various sparsity measures and algorithms, here we use l_1-norm with LASSO algorithm [54] and l_0-norm minimization algorithm [34]. Genie-aided solutions with known bad data locations are also included as a reference.

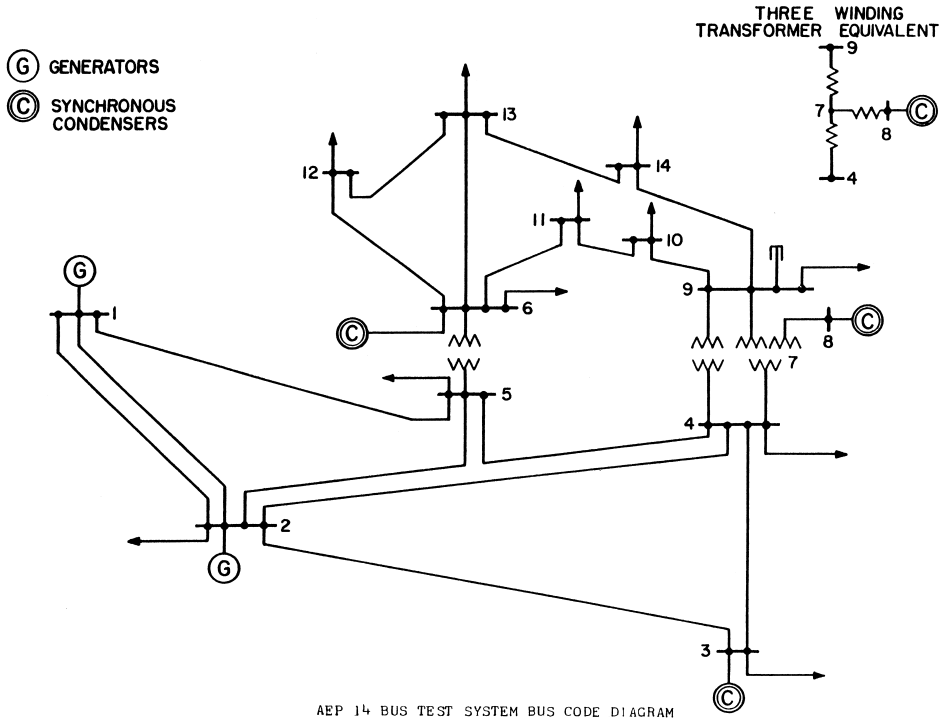


Figure 5.1: IEEE 14-bus test system.

In Figs. 5.2 and 5.3, we simulate the case with 14 bus voltage and 14 injection current measurements (q=2p) and the occurrence of a single bad measurement and 3 bad measurements, respectively. Clearly, the measurements are not redundant enough for LRR to guarantee a good performance. In Fig. 5.2, PM is best and perfectly identifies the single bad data as expected. Moreover, none of other algorithms is

good even in this single bad data occurrence case. In Fig. 5.3 with 3 bad data, it can be seen that the performance is ranked as Genie-Aid>PM>l_0-SRM>LASSO-SRM.

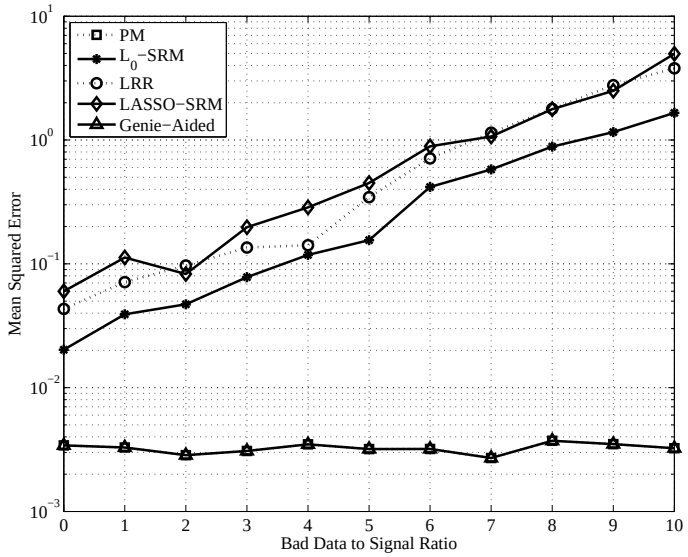


Figure 5.2: The state estimation performance for measuring the bus voltages and injection currents with 1 bad data.

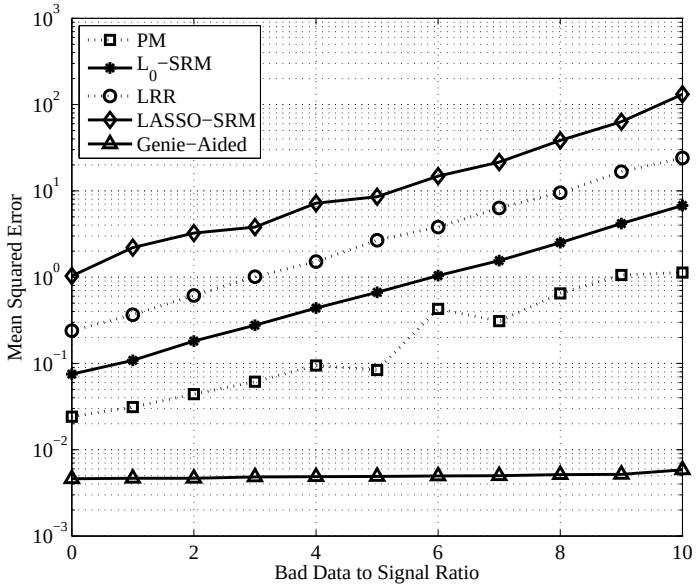


Figure 5.3: The state estimation performance for measuring the bus voltages and injection currents with 3 bad data.

In Fig. 5.4, we increase the data redundancy by including all possible 68 measurements, i.e., 14 bus voltage, 14 injection current and 40 line current measurements ($q > 4p$). Out of these, 4 bad measurements are randomly generated. The results show that even with increased bad measurements, performance of all but LASSO is significantly better than those in Figs. 5.2 and 5.3. This implies that in order to improve the estimator performance, it is desirable to include as many measurements as possible. However, LASSO has worse performance than in former cases, which may be caused by the non-fatness of the 54×68 regression matrix $\boldsymbol{H}_{\perp}$. On the other hand, it is worth noticing that in this case, the measurement redundancy resulted in pretty good performance with LRR.

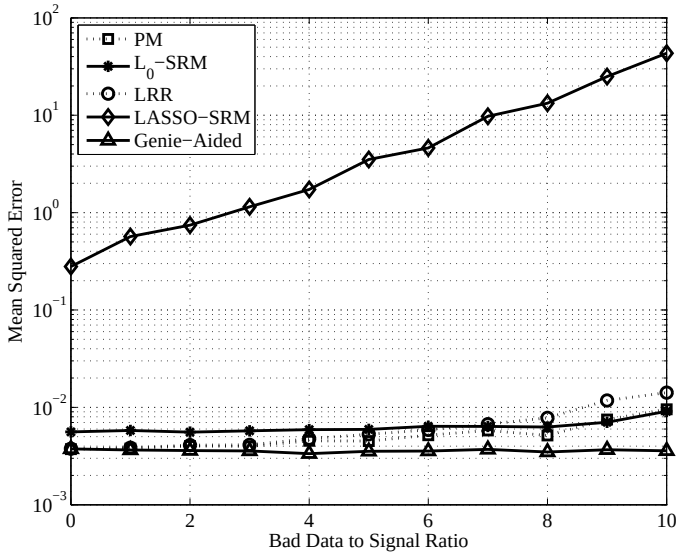


Figure 5.4: The state estimation performance for measuring the bus voltages, injection currents and all line currents with 4 bad data.

5.6 Conclusions

In this chapter, power grid state estimation algorithms using PMU measurements with bad measurements were proposed. With bad data explicitly incorporated into the estimation model, we showed that the conventional LRR does not have performance assurance and developed several more sophisticated algorithms. To deal with the bad data, one can estimate the bad data location and values either separately or jointly. For the former approach, we established the equivalence between bad data removal and subtraction which is a consequence of the fact that the oblique pseudo-inverse [6] is a subtraction algorithm for a selection matrix. Then we developed the projection and minimization (PM) algorithm. For the joint bad data location and value estimation, we formulated a sparsity regularization minimization (SRM) problem and transformed it into a compressive sensing problem. Simulations on the IEEE 14-bus test system with different levels of measurement redundancy and bad data occurrence are provided. Results showed that our PM algorithm has not only the lowest complexity but also the best performance among existing competitors.

CHAPTER 6

OPTIMAL LOCAL DETECTION FOR SENSOR FUSION BY LARGE DEVIATION ANALYSIS

6.1 Motivation

Signal detection is a classical problem in applications including radar, wireless sensor networks, wireless communication systems, cognitive radio spectrum sensing, and so on. It will also have important applications in future power system for fault diagnosis, event classification, and so on. To enhance performance, a fusion center collects information from multiple local detectors and makes a global decision. Due to the bandwidth constraint, the local detectors often make decisions first and transmit the one bit decisions to the fusion center. Accordingly, the entire process is called detection fusion or decision fusion [64].

In the pioneering work of Tsitsiklis on this problem [60], it has been shown that while the fusion strategy can be easily obtained by the Neyman-Pearson (NP) lemma, the selection of a local decision rule to optimize the global performance is intractable. In the current literature, some work fixes the fusion rule and then obtains the optimal local decision rule [58]; whereas others compare the fusion detection performance for various local decision rules, including the locally optimal minimized average error probability [14], the maximum decision output entropy [8] and the largest divergence between the statistical distribution under different hypotheses [30, 21]. None of these

detectors is optimum. Recently, some asymptotic analyses for detection fusion have been reported in the literature. For example, [18] develops a fusion rule for channel distorted decisions using a Chernoff exponent bound analysis. Similar analysis is followed in [19] to obtain an asymptotically optimum fusion rule for an M-hypothesis testing problem, and in [17] for non-centralized distributed fusion. However, these papers focus only on designing the *fusion* rule, while the optimum local decision strategy remains an open problem. In this chapter, our goal is to find an optimal *local* decision strategy that optimizes the asymptotic global performance.

We will deal with a parallel fusion structure [65] and work with a binary hypothesis testing problem. By large deviation analysis, we will optimize the local thresholds to obtain the best global performance, asymptotically in the number of local detectors. Compared with existing work in the literature, our method has a lower complexity and guarantees the global optimal performance, asymptotically. Some interesting properties of the optimal strategy will also be discussed. Then, with a specific example of cooperative energy sensing, we will demonstrate the optimality of our proposed algorithm.

This chapter is organized as follows: we first present the general signal model for detection fusion in Section 6.2 and formulate the joint optimization problem in Section 6.3. Then, we will present the error exponent expressions in Section 6.4 and develop the asymptotically optimized local detection strategy accordingly in Section 6.4. Finally, we present a case study to compare performance under various local decision strategies in Section 6.6 and give concluding remarks in Section 6.7. Throughout

the chapter, $X \sim \mathcal{CN}(\mu, \sigma^2)$ denotes a random variable X following a proper complex Gaussian distribution with mean μ and variance σ^2; $d \sim \text{Ber}(p)$ denotes a Bernoulli random variable; $X \sim \text{Bin}(N, p_1)$ denotes a random variable X following a binomial distribution; $f(x) \sim g(x)$ means that $\lim_{x \to +\infty} \frac{f(x)}{g(x)} = c$ where c is a constant.

6.2 System Model

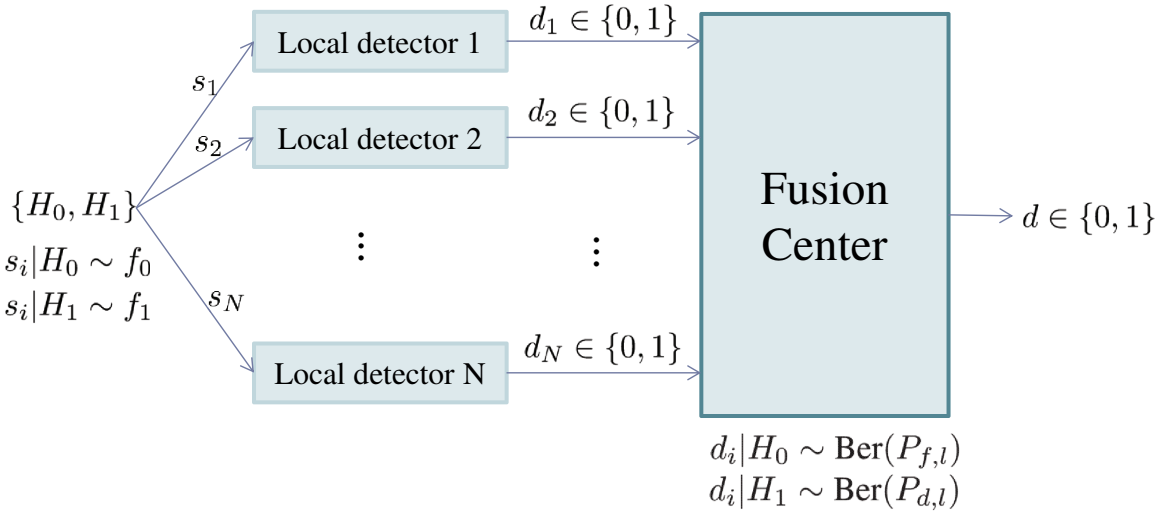


Figure 6.1: System diagram for detection fusion.

The diagram for a detection fusion system is shown in Fig. 6.1. As depicted in this figure, there is a common random signal source which follows either distribution f_0 under hypothesis H_0, or distribution f_1 under hypothesis H_1, where $P(H_0 \text{ true}) = \pi_0$ and $P(H_1 \text{ is true}) = \pi_1$ are the a priori probabilities of the hypotheses. Each local detector will make its own local decision $d_i \in \{0, 1\}$ based on its own observed signal s_i. Then, a fusion center will collect all local decisions d_is and make a global decision $d \in \{0, 1\}$ accordingly. It has been shown that in the case that the signals at local decisions are dependent, the solution for optimal detection fusion is non-deterministic

polynomial-time hard [61]. Therefore, in our analysis, we assume that the signals at different detectors are independent, which is true in many real applications. Then, in [60], it is proved that to obtain asymptotically optimal performance, all local detectors should follow the same decision rule. Under this strategy, the d_is are independent and identically distributed.

To describe the distributions of the d_is at the fusion center, we denote $P_{f,l} = P(d_i = 1|H_0)$ as the local false alarm probability and $P_{d,l} = P(d_i = 1|H_1)$ as the local detection probability. Then, $(P_{f,l}, P_{d,l}) \in [0,1] \times [0,1]$ is called the receiver operating characteristic (ROC) curve. The local decision d_i follows a Bernoulli distribution with $P_{f,l}$ and $P_{d,l}$ under hypothesis H_0 and H_1, respectively. At the fusion center:

$$\begin{aligned} P(d_1, d_2, \ldots, d_N|H_0) &= P_{f,l}^{\sum_{i=1}^N d_i} \left(1 - P_{f,l}\right)^{N - \sum_{i=1}^N d_i} , \\ P(d_1, d_2, \ldots, d_N|H_1) &= P_{d,l}^{\sum_{i=1}^N d_i} \left(1 - P_{d,l}\right)^{N - \sum_{i=1}^N d_i} . \end{aligned} \tag{6.1}$$

Accordingly, $d_s = \sum_{i=1}^{N} d_i$ is the sufficient statistics and it follows a binomial distribution under each hypothesis.

6.3 Optimum Local and Fusion Decisions

In this chapter, we adopt the global average error probability as the performance metric, i.e., $P_e = \pi_0 P(d = 1|H_0) + \pi_1 P(d = 0|H_1)$. To obtain the best performance, we want to find a local threshold and a corresponding fusion rule that minimizes P_e.

The Bayesian detector will minimize P_e by implementing the likelihood ratio test [50]

$$\frac{\pi_1 P_{d,l}^{d_s}(1-P_{d,l})^{N-d_s}}{\pi_0 P_{f,l}^{d_s}(1-P_{f,l})^{N-d_s}} \underset{H_0}{\overset{H_1}{\gtrless}} 1 \ , \tag{6.2}$$

and the corresponding minimized P_e can be calculated. Notice that as long as $(P_{f,l}, P_{d,l})$ is known to the fusion center, the optimal fusion rule can be easily obtained according to Eq. (6.2).

From Eq. (6.2), it is easy to verify that for any given local false alarm probability $P_{f,l}$, the larger the local detection probability $P_{d,l}$ is, the smaller the global average error probability P_e will be. Therefore, at local detectors, the NP detector or equivalently the maximum likelihood (ML) detector [50] should be adopted to achieve the best performance:

$$\frac{f_1(s_i)}{f_0(s_i)} \underset{H_0}{\overset{H_1}{\gtrless}} L \ . \tag{6.3}$$

However, this will only give an ROC curve $(P_{f,l}, P_{d,l})$ for the local detectors. How to select the optimal point $(P_{f,l}^o, P_{d,l}^o)$ on the ROC of the NP detector according to Eqs. (6.2) and (6.3) is usually a non-convex and mathematically intractable problem. In addition, the optimization process involves the number of local detectors N, which is not always available to local detectors.

In this chapter, we will use large deviation analysis to obtain the optimal local decision strategy, i.e. $(P_{f,l}^o, P_{d,l}^o)$ to minimize the global average error probability P_e, asymptotically in N.

6.4 Error Exponent Expressions

As introduced in Section 6.2, the sufficient statistic at the fusion center $d_s = \sum_{i=1}^{N} d_i$ follows a binomial distribution:

$$\begin{aligned} H_0 &: d_s \sim \text{Bin}(N, P_{f,l}) \ , \\ H_1 &: d_s \sim \text{Bin}(N, P_{d,l}) \ . \end{aligned} \tag{6.4}$$

Let the fusion threshold be $P_{f,l}N < \eta_f = \theta_F N < P_{d,l}N$. Then by large deviation analysis, the global error probabilities are asymptotically [4]:

$$\begin{aligned} P_f &= P(d_s \geq \theta_F N | H_0) \ &\sim e^{-NE_0} \ , \\ P_{md} &= P(d_s < \theta_F N | H_1) \ &\sim e^{-NE_1} \ , \end{aligned} \tag{6.5}$$

where

$$\begin{aligned} E_0 &= \theta_F \log \frac{\theta_F}{P_{f,l}} + (1-\theta_F) \log \frac{1-\theta_F}{1-P_{f,l}} = D_{KL}(\theta_F || P_{f,l}) \ , \\ E_1 &= \theta_F \log \frac{\theta_F}{P_{d,l}} + (1-\theta_F) \log \frac{1-\theta_F}{1-P_{d,l}} = D_{KL}(\theta_F || P_{d,l}) \ , \end{aligned} \tag{6.6}$$

and $D_{KL}(\cdot)$ denotes the Kullback-Leibler divergence [10]. Accordingly, the overall probability of error is given by:

$$P_e = \pi_0 P_f + \pi_1 P_{md} \sim \pi_0 e^{-NE_0} + \pi_1 e^{-NE_1} \sim e^{-N \min(E_0, E_1)} \tag{6.7}$$

6.5 Asymptotically Optimal Local Decision

To minimize this global average error probability asymptotically, we need to maximize $\min(E_0, E_1)$. Hence, the problem becomes:

$$\max_{P_{f,l}, P_{d,l}, \theta_F} \min(E_0, E_1) \tag{6.8}$$

It should be noticed that when $P_{f,l} < \theta_F < P_{d,l}$, $E_0(\theta_F)$ is an increasing function of θ_F and $E_1(\theta_F)$ is a decreasing function of θ_F. As a result, the maximum value of $\min(E_0(\theta_F), E_1(\theta_F))$ is achieved when $E_0(\theta_F) = E_1(\theta_F)$. According to Eq. (6.6),

$$\theta_F^o = \frac{\log \frac{1-P_{d,l}}{1-P_{f,l}}}{\log \frac{P_{f,l}}{P_d,l} + \log \frac{1-P_{d,l}}{1-P_{f,l}}} \, . \tag{6.9}$$

According to Eqs. (6.8) and (6.9), the optimal local decision rule can be obtained as follows:

$$(P_{f,l}^o, P_{d,l}^o) = \arg \max_{(P_{f,l}, P_{d,l})} D_{KL}(\theta_F^o || P_{f,l}) \, , \tag{6.10}$$

where θ_f^o is parameterized by $(P_{f,l}, P_{d,l})$ according to Eq. (6.9).

Recall that for local detectors, we already have an NP detector ROC curve which constrains $(P_{f,l}, P_{d,l})$ to lie on the ROC curve, at a point determined by the local threshold. So, Eq. (6.10) can be interpreted as a search over the ROC curve to find a point which leads to the maximum error exponent. Although $(P_{f,l}, P_{d,l})$ is two-dimensional, it only has a one-dimensional degree of freedom, namely the local threshold. This renders the optimization a one dimensional problem. In fact, under

many signal models, the NP local detectors are in the form of a scalar sufficient statistic compared to a single threshold and in this case $P_{f,l}$ and $P_{d,l}$ can often be represented by this threshold analytically in closed form. Therefore, the global average error exponent can be rewritten as a single-variable function. The objective function $D_{KL}(\theta_F^o||P_{f,l})$ is uni-modal in many scenarios and hence can be easily optimized by line search techniques such as those in [9, Chapter 7].

Note that although Eq. (6.9) gives an asymptotically optimal fusion threshold, the fusion center always uses an NP detector according to Eq. (6.2) to obtain the best fusion performance.

Remarks:

1. Asymptotically, the optimal local decision strategy is *independent* of the total number of sensors N, but only dependent on the signal model s_i under the original hypotheses. This enables the global optimization even when the local distributed detectors do not know the network size N. In fact, if the sensors have sufficient computing resources, the local thresholds could be periodically recomputed locally if the distribution of s_i changes over time.

2. Asymptotically, the optimal local decision strategy is independent of the a priori probabilities. This is due to the fact that when N approaches infinity, the π_0 and π_1 terms in Eq. (6.2) will contribute very little to the likelihood ratio.

6.6 Example: Energy Sensing

To illustrate our solution for the asymptotically optimum detection fusion, we adopt the specific signal model for a cooperative energy sensing problem as an example and show the performance comparisons.

6.6.1 Signal Model

In the energy sensing problem, the task is to determine whether there is a signal transmitted over a certain channel (H_1) or not (H_0). Under Rayleigh fading and additive white Gaussian noise, the normalized signal model for local detectors is [14]:

$$\begin{aligned} s_i|H_0 = n &\sim \mathcal{CN}(0,1) \\ s_i|H_1 = hx + n &\sim \mathcal{CN}(0,\gamma+1) \end{aligned} \tag{6.11}$$

where n is white Gaussian noise, h is a Rayleigh fading channel, x is the transmitted signal and γ is the average signal to noise ratio (SNR). Under this signal model, the NP detector is the energy detector:

$$\|s_i\|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta \ . \tag{6.12}$$

Correspondingly, the local false alarm and detection probabilities are

$$\begin{aligned} P_{f,l} &= e^{-\eta} \ , \\ P_{d,l} &= e^{-\frac{\eta}{\gamma+1}} \ . \end{aligned} \tag{6.13}$$

6.6.2 Numerical Results

To gain a better understanding of the detection fusion optimization problem, we first plot in Fig. 6.2 the performance (P_e) surface vs. the local and fusion thresholds. In this figure, the number of local detectors is $N = 20$. Evidently, there are 4 local minima. This verifies our discussions of the non-convexity in Section 6.3. In addition, the number of local minima will increase with N.

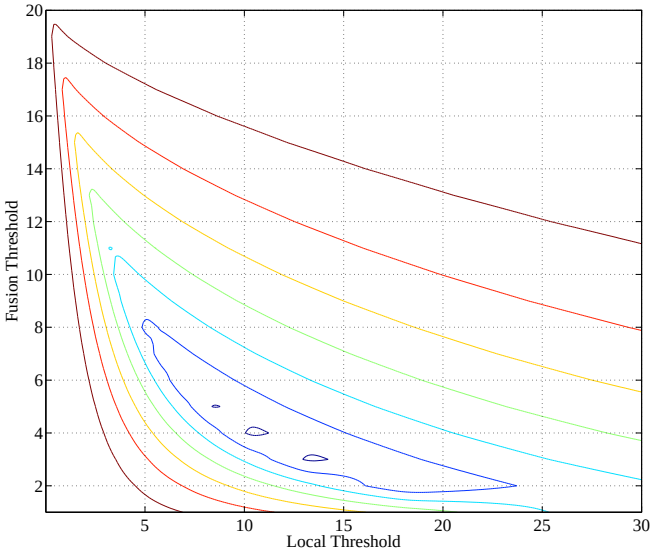


Figure 6.2: The performance surface under local and fusion thresholds with $N = 20$.

For large deviation analysis, we plot the error exponent under different local decision thresholds in Fig. 6.3. In this figure, it can be observed that with the energy sensing signal model, the error exponent is a uni-modal function of the local threshold. Therefore, the optimal local threshold can be easily found using a one-dimensional line search algorithm.

The local thresholds for joint optimization by exhaustive search, and the thresholds by large deviation analysis for several different per sensor SNR values are plotted

vs. N in Fig. 6.4. It can be verified that as the number of local detectors increases, the local thresholds obtained by the joint optimization will converge to the threshold given in our large deviation analysis.

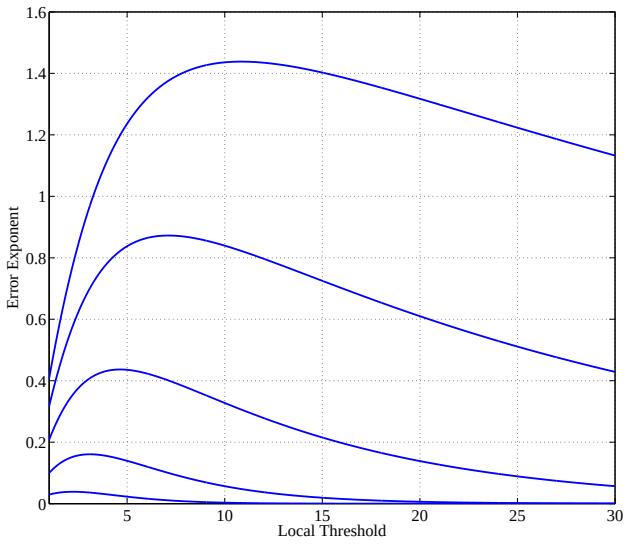


Figure 6.3: Error exponent under different local thresholds. From bottom to top, the per sensor SNR is $\gamma = 0, 5, 10, 15, 20$ dB.

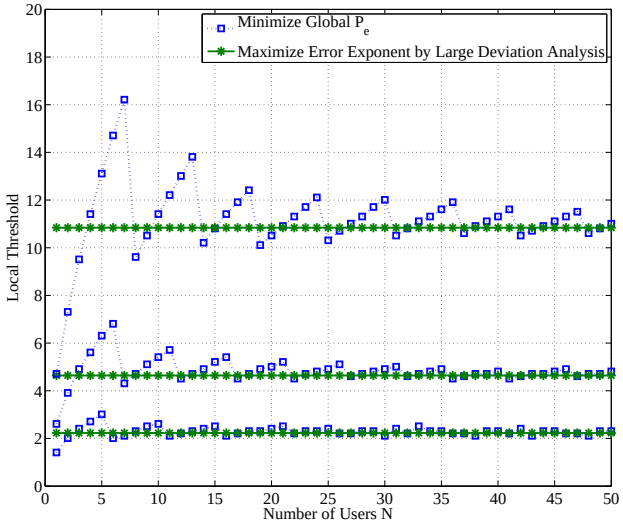


Figure 6.4: Local decision thresholds under joint optimization by exhaustive search and large deviation analysis. From bottom to top, the per sensor SNR is $\gamma = 0, 10, 20$ dB.

In Fig. 6.5, we compare the performance of the large deviation solution with existing ones, including the local average error probability minimization $\min(P_{f,l}+1-P_{d,l})$ [14], the decision output entropy maximization [8] with $P_{f,l} = P_{d,l}$ and the mutual information maximization between decision and hypothesis with $\max(I(H, d_i))$ [21]. We also present the performance limit by optimizing the local thresholds via exhaustive search. Note that in all cases, the fusion threshold is obtained according to Eq. (6.2).

In Fig. 6.5, we plot the global average error probability at per sensor SNR $\gamma = 15$ dB as a function of the number of local detectors N. It can be observed that the average error probability does decay exponentially with N as the large deviation analysis indicates. In addition, our proposed method approaches the optimized detection fusion by exhaustive search very well and actually does not require N to be very large to approach the optimal performance.

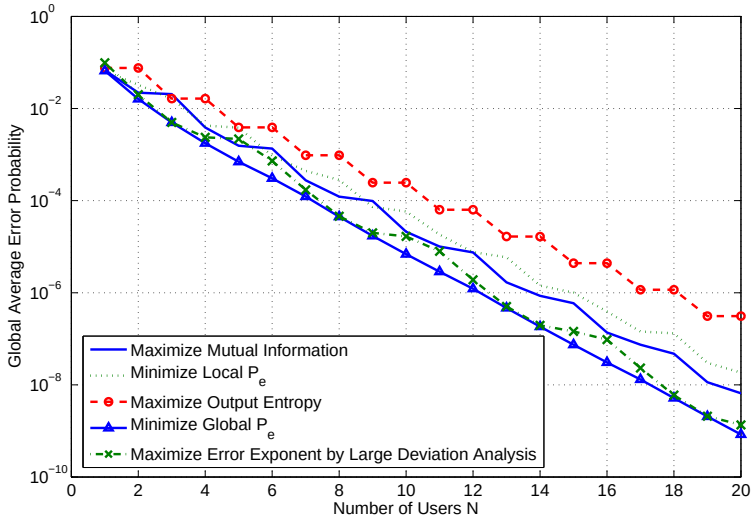


Figure 6.5: Global average error probability under different local decision criteria at per detector SNR $\gamma = 15$ dB as a function of the number of local detectors N.

In Fig. 6.6, we plot the global average error probability performance at $N = 8$ as a function of per sensor SNR. It can be seen that even with a quite small N in this case, our proposed fusion rule has nearly the same performance with the exhaustively optimized fusion rule. In addition, the performance has a faster decaying rate than all other alternatives when the per sensor SNR increases.

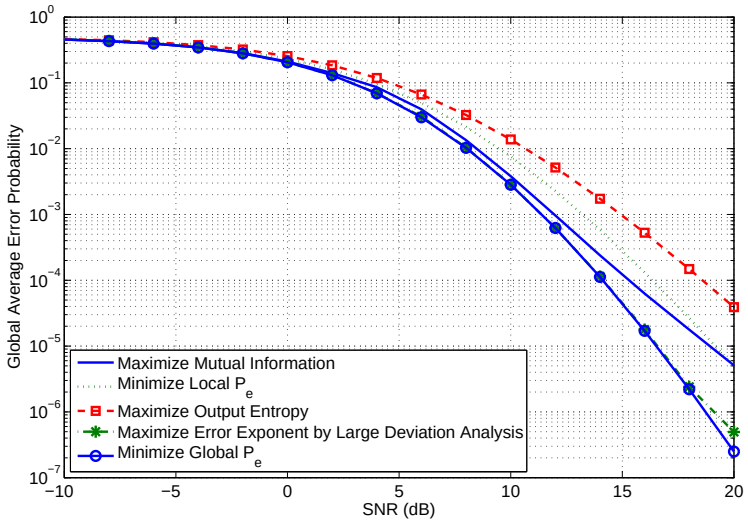


Figure 6.6: Global average error probability P_e under different local decision criteria at $N = 8$ as a function of the per detector SNR.

6.7 Conclusions

In this chapter, large deviation analysis is used to derive the asymptotically optimal local detection strategy for detection fusion. Asymptotically, the joint optimization problem was simplified to a simple line search on an ROC curve. It was observed that the asymptotically optimal local decision rule is independent of the number of local detectors N and the a priori probabilities of the hypotheses. A cooperative energy sensing problem was considered to demonstrate our proposed approach. Numerical

results verify that our proposed method approaches the optimal local detection strategy obtained by exhaustive search and has demonstrated better performance than all other reported local decision alternatives at small to moderate N values, with no additional information required at the local detectors.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

In this research, we first introduced the smart grid and highlighted the importance of situational awareness. To gain better situational awareness, sensing, communications and monitoring of the smart grid are involved. To fulfill the various communication requirements of the smart grid, wireless communication is a very good candidate for its simplicity and low cost [20]. However, wireless communication does have several drawbacks. Firstly, the devices in wireless communications are usually driven by batteries which can cause damage to the environment which works against the green trend of the smart grid. Thus, we looked into green communication strategies by selecting more battery efficient modulation schemes and utilizing possible cooperative communication links. Secondly, wireless communications nowadays are running out of spectrum resources. Investment in extra spectrum resources for smart grid communications would counteract the low cost advantage of wireless communication. Fortunately, cognitive radio, which accesses unused bands for existing licensed users is a promising technique to solve this problem. In our work, we developed cooperative spectrum sensing schemes from a diversity perspective to improve the cognitive radio system performance and advance its availability for real applications. Then, when basic sensing and communications are enabled in the smart grid, we looked into

the monitoring of the power grid states utilizing the synchronized measurement from PMUs. We developed and compared several algorithms to address the problem of possible bad data occurrence. Higher layer activities in smart grid, such as power grid control, protection, electric pricing, power flow schedule and so on, can be benefited from our works. In addition, optimal sensor fusion technique under a general system setup was developed in our work, which can be applied to specific smart grid scenarios.

7.2 Future Works

7.2.1 State Estimation with Parameter and Topology Errors

For the monitoring of the power grid, we studied state estimation (SE), which is a crucial component in the supervisory control and data acquisition (SCADA) and energy management systems (EMS). We utilized the recently developed synchronized measurement device, namely the phasor measurement unit (PMU) to make SE a linear process which can be carried out much more efficiently than SE based on traditional measurements. With this convenient linear model, several sophisticated bad data processing techniques were developed. Our proposed projection and minimization (PM) bad data removal scheme is fast, simple and has good performance. However, besides bad data in the measurement, possible parameter and topology errors may also occur [1, Chapters 7 and 8]. In our future work, we will develop algorithms to detect and identify parameter and topology errors in the state estimation process.

The possible coexistence of bad data in the measurements, parameter errors and topology errors renders state estimation a very complicated process. We propose to utilize the structure and properties of the data to distinguish different types of errors. Traditionally, in the detection process, we only evaluate the residual $\boldsymbol{r}$ for the entire model. Actually, finer detection can be made to verify the correctness of the data or the model parameters and the structure. Assuming that the full redundancy measurements are available, then the following finer detections can be made:

1. **The agreement between voltage and current measurement:** With full redundancy measurements, the disagreement between the voltage and current measurement can be detected. For any current measurements, $\tilde{\boldsymbol{m}_I} = \boldsymbol{m}_I - \boldsymbol{Y}\boldsymbol{m}_V$. An hypothesis test of $\tilde{\boldsymbol{m}_I}$ indicates the agreement/disagreement between voltage and current measurements. If disagreement occurs, we can locate suspicious bad measurements, parameter and topology errors.

2. **KCL agreement:** KCL states that the overall current of a nodes should sum up to zero, i.e., $\sum_j \boldsymbol{m}_I(ij) = 0$. If disagreement occurs among current measurements for KCL, then it should be caused by the bad measurements and the corresponding suspicious bad measurements can be located.

3. **KVL agreement:** KVL states that the overall voltage drop within a loop should sum up to zero, i.e., $\sum_{l \in \{\text{Loop } L\}} \boldsymbol{Y}_l \boldsymbol{m}_{I,l}$. If disagreement occurs among current measurements for KVL, one can locate the suspicious bad measurements, parameter and topology errors.

By combining all these locations of suspicious bad measurements, parameter and topology errors from the finer local tests, it is possible to determine the exact type of error and develop the corresponding error correction mechanism.

7.2.2 Optimal Sensor Fusion with More Complicated Setup

In Chapter 6, we introduced the asymptotical optimal local detection for sensor fusion by large deviation analysis. However, we were assuming a quite simple setup. There are several ways to incorporate more complex setups:

1. The local detectors make multiple-bit decisions. In this case, there would be multiple thresholds for local detectors, and the resultant statistics at the fusion center would be multi-nomail. How to select the local thresholds would be a challenge and interesting problem.

2. In our current work, we only deal with binary hypothesis testing problems. In many applications, there might be multiple hypotheses under test. To find the optimal local detection strategy in this scenario is also a possible future work.

3. In our current setup, we assume that the signal models for all individual sensors are the same. However, in many applications, they observe signals with different distributions even under the same hypothesis. For example, they might have different distances from the signal sources, hence have different SNRs in their received signals. How to make sensor fusion in this case is another interesting and challenging problem.

LIST OF REFERENCES

[1] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementaion*. CRC Press, 2004.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, August 2002.

[3] N. Armaroli and V. Balzani, "The future of energy supply: challenges and opportunities," *Angewandte Chemie Internation Edition*, vol. 46, no. 1-2, pp. 52–66, December 2006.

[4] R. Arratia and L. Gordon, "Tutorial on large deviations for the binomial distribution," *Bulletin of Mathematical Biology*, vol. 51, no. 1, pp. 125–131, January 1989.

[5] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, July 2007.

[6] R. T. Behrens and L. L. Scharf, "Signal processing applications of oblique projection operators," *IEEE Trans. on Signal Processing*, vol. 42, no. 6, pp. 1413–1424, June 1994.

[7] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Experimental study of spectrum sensing based on energy detection and network cooperation," in *Proc. ACM 1st Int. Workshop on Technology and Policy for Accessing Spectrum (TAPAS)*, Boston, MA, August 2006, pp. 1–8.

[8] S. Chaudhari, J. Lunden, V. Koivunen, and H. V. Poor, "Cooperative sensing with imperfect reporting channels: Hard decisions or soft decisions?" *IEEE Trans. on Signal Processing*, vol. 60, no. 1, pp. 18–28, January 2010.

[9] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*, 3rd ed. Hoboken, NJ: Wiley-Interscience, 2008.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, August 1991.

[11] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, September 2005.

[12] S. Cui, J. Xiao, A. J. Goldsmith, Z. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, September 2007.

[13] D. Duan, F. Qu, L. Yang, A. Swami, and J. C. Principe, "Modulation selection from a battery power efficiency perspective," *IEEE Trans. on Communications*, vol. 58, no. 7, pp. 1907–1911, July 2010.

[14] D. Duan, L. Yang, and J. Principe, "Cooperative diversity of spectrum sensing for cognitive radio systems," *IEEE Trans. on Signal Processing*, vol. 58, no. 6, pp. 3218–3227, June 2010.

[15] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Baltimore, MD, November 8-11, 2005, pp. 131–136.

[16] A. Gomez-Exposito, A. Abur, A. de la Villa Jaen, and C. Gomez-Quiles, "A multilevel state estimation paradigm for smart grids," *Proceeding of the IEEE*, vol. 99, no. 6, pp. 952–976, June 2011.

[17] J. A. Gubner, E. K. P. Chong, and L. L. Scharf, "Aggregation and compression of distributed binary decisions in a wireless sensor network," in *Proc. of the 48th IEEE CDC 2009*, Shanghai, China, Decemember 15-18, 2009, pp. 909–913.

[18] J. A. Gubner, L. L. Scharf, and E. K. P. Chong, "Exponential error bounds for binary detection using arbitrary binary sensors and an all-purpose fusion rule in wireless sensor network," in *Proc. of Intl. Conf. on Acoustics, Speech, and Signal Processing*, Taipei, Taiwan, April 19-24, 2009, pp. 2781–2784.

[19] ——, "Optimization of exponential error rates for a suboptimum fusion rule in wireless sensor networks," in *Proc. of Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, November 6-9, 2011, pp. 1–4.

[20] V. C. Gungor, B. Lu, and G. P. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Trans. on Industrial Electronics*, vol. 57, no. 10, pp. 3557–3564, October 2010.

[21] Y. I. Han and T. Kim, "Mutual and conditional mutual informations for optimizing distributed Bayes detectors," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 37, no. 1, pp. 147–157, January 2001.

[22] D. A. Harville, *Matrix Algebra From a Statistican's Perspective*. Springer, 2008.

[23] C. H. Hauser, D. E. Bakken, and A. Bose, "A failure to communication: next generation communication requirement, technologies, and architecture for the electrical power grid," *IEEE Power and Energy Magazine*, vol. 3, no. 2, pp. 47–55, March 2005.

[24] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, February 2005.

[25] S. Hong, M. H. Vu, and V. Tarokh, "Cognitive sensing based on side information," in *IEEE Sarnoff Symposium*, Princeton, NJ, April 28-30, 2008, pp. 1–6.

[26] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE Power and Energy Magazine*, vol. 7, no. 2, pp. 52 –62, March-April 2009.

[27] P. Kaewprapha, J. Li, and Y. Yu, "Cooperative spectrum sensing with tri-state probabilistic inference," in *Proc. of Military Communications Conference*, San Jose, CA, October 31-November 3, 2010, pp. 318–323.

[28] S. Lakkavalli, A. Negi, and S.Singh, "Stretchable architectures for next generation cellular networks," in *Proc. of the Intl. Symp. on Advanced Radio Tech.*, 2003, pp. 59–65.

[29] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocals and outage behavior," *IEEE Trans. on Information Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.

[30] C.-C. Lee and J.-J. Chao, "Optimum local decision space partitioning for distributed detection," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 25, no. 4, pp. 536–544, July 1989.

[31] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. on Wireless Communications*, vol. 7, no. 11, pp. 4502–4507, November 2008.

[32] S. M. Mishra, R. Tandra, and A. Sahai, "The case for multiband sensing," in *Proc. of the Forty-Fifth Annual Allerton Conference on Communication, Control, and Computing.* The Allerton House, University of Illinois, September 2007. [Online]. Available: http:///php/pubs/pubs.php/114.html

[33] J. Mitola III and G. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, August 1999.

[34] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed L-0 norm," *IEEE Trans. on Signal Processing*, vol. 57, no. 1, pp. 289–301, January 2009.

[35] K. Moslehi and R. Kumar, "Smart grid: a reliability perspective," in *Proc. of Innovative Smart Grid Technologies (ISGT) 2010*, Gaithersburg, MD, January 19-21, 2010.

[36] M. Mustonen, M. Matinmikko, and A. Mämmalä, "Cooperative spectrum sensing using quantized soft decision combining," in *Proc. of the Fourth International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Hannover, Germany, June 22-24, 2009, pp. 1–5.

[37] 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*, Minneapolis, MN, July 25-29, 2010, pp. 1–7.

[38] M. Pedram and Q. Wu, "Battery-powered digital CMOS design," *IEEE Trans. on VLSI Systems*, vol. 10, no. 5, pp. 607–607, October 2002.

[39] A. Pezeshki, R. Calderbank, and S. D. Howard, "Diversity order for detection with distributed sensors," *IEEE Trans. on Signal Processing*, Submitted November 2007.

[40] A. G. Phadke, "Synchronized phasor measurements in power systems," *IEEE Computer Applications in Power*, vol. 6, no. 2, pp. 10–15, April 1993.

[41] A. G. Phadke, J. S. Thorp, R. F. Nuqui, and M. Zhou, "Recent developments in state estimation with phasor measurements," in *Proc. of Power Systems Conference and Exposition*, Seattle, WA, March 15-18, 2009.

[42] A. G. Phadke and J. Thorp, *Sychronized phasor measurements and their applications*. New York: Springer Science, 2010.

[43] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Springer-Verlag, 1994.

[44] J. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, New York, February 2001.

[45] F. Qu, D. Duan, L. Yang, and A. Swami, "Signaling with imperfect channel state information: A battery power efficiency comparison," *IEEE Trans. on Signal Processing*, vol. 56, no. 9, pp. 4486–4495, September 2008.

[46] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE Journal of Selected Topics in Singal Processing*, vol. 2, no. 1, pp. 28–40, February 2008.

[47] R. C. Qui, Z. Hu, N. Guo, R. Ranganathan, S. Hou, and G. Zheng, "Cognitive radio network for the smart grid: experimental system architecture, control algorithms, security and micro grid testbed," *IEEE Trans. on Smart Grid*, vol. 2, no. 4, pp. 724–740, December 2011.

[48] T. S. Rappaport, *Wireless communications–principles and practice*, 2nd ed. Prentice Hall, 2002.

[49] Rice University, *Compressive Sensing Resources*. [Online]. Available: http://dsp.rice.edu/cs

[50] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

[51] L. L. Scharf, P. Mathys, and R. T. Behrens, "Rank reduction for decoding linear block-codes over the complex field," in *Proceedings of the 21st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1987.

[52] K. Schwieger and G. Fettweis, "Power and energy consumption for multi-hop protocols: A sensor network point of view," in *International Workshop on Wireless Ad-hoc Network, IWWAN 2005*, 2005.

[53] J. Song, H. Lee, and D. Cho, "Power consumption reduction by multi-hop transmission in cellular networks," in *Proc. of Vehicular Tech. Conf.*, vol. 5, Sep 26-29, 2004, pp. 3120–3124.

[54] Stanford University, *SparseLab: Seeking Sparse Solutions to Linear Systems of Equations*. [Online]. Available: http://sparselab.stanford.edu/

[55] G. Staple and K. Werbach, "The end of spectrum scarcity," *IEEE Spectrum*, vol. 41, no. 3, pp. 48–52, March 2004.

[56] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 4–17, February 2008.

[57] Q. Tang, L. Yang, G. B. Giannakis, and T. Qin, "Battery power efficiency of PPM and FSK in wireless sensor networks," *IEEE Trans. on Communications*, vol. 6, no. 4, pp. 1308–1319, April 2007.

[58] R. R. Tenney and N. R. Sandell, JR., "Detection with distributed sensors," *IEEE Trans. on Aerospace and Electronics Systems*, vol. AES-17, no. 4, pp. 501–510, July 1981.

[59] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

[60] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals and Systems*, vol. 1, no. 2, pp. 167–182, 1988.

[61] J. N. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Transcations on Automatic Control*, vol. 30, no. 5, pp. 440–446, May 1985.

[62] University of Washington, *Power System Test Case Archive*. [Online]. Available: http://www.ee.washington.edu/research/pstca/

[63] J. Unnikrishnan and V. V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *IEEE Journal of Selected Topics in Singal Processing*, vol. 2, no. 1, pp. 18–27, February 2008.

[64] P. K. Varshney, *Distributed Detection and Data Fusion*. New York, NY: Springer-Verlag, 1996.

[65] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I– fundamentals," *Proceedings of The IEEE*, vol. 85, no. 1, pp. 54–63, January 1997.

[66] W. Wang, Y. Xu, and M. Khanna, "A survey on the communication architectures in smart gird," *Computer Networks*, vol. 55, no. 15, pp. 3604–3629, October 2011.

[67] F. F. Wu, "Power system state estimation: A survey," *International Journal of Electrical Power & Energy Systems*, vol. 12, no. 2, pp. 80–87, April 1990.

[68] W. Zhang, D. Duan, and L. Yang, "Relay selection from a battery energy efficiency perspective," *IEEE Trans. on Communications*, vol. 59, no. 6, pp. 1525–1529, June 2011.

[69] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.