國 立 交 通 大 學

電信工程研究所

博 士 論 文

噪訊增強的無線通訊網路盲蔽式錯誤率估測器之設計與效能分析

Design and analysis of noise-enhanced blind error rates estimation in wireless network

研 究 生：劉人仰

指導教授：蘇育德 博士

中 華 民 國 101 年 7 月

噪訊增強的無線通訊網路盲蔽式錯誤率估測器之設計與效能分析

Design and analysis of noise-enhanced blind error rates estimation in wireless network

研究生：劉人仰　　Student: Jen-Yang Liu

指導教授：蘇育德　博士　　Advisor: Dr. Yu T. Su

國立交通大學

電信工程研究所

博士論文

A Dissertation
Submitted to Institute of Communication Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in
Communication Engineering
Hsinchu, Taiwan

2012 年 7 月

噪訊增強的無線通訊網路盲蔽式錯誤率估測器之設計與效能分析

研究生：劉人仰　　指導教授：蘇育德 博士

國立交通大學電信工程研究所

中文摘要

基於多筆資訊做資料偵測或融合的技術已經廣泛被使用在通訊系統上。本論文考慮一廣義的無線網路框架，在此框架中，我們假設訊號會經過不同的鏈結(link)到達接收器，而鏈結的種類包含了直接鏈結(從傳送端到接收端)以及兩步(two-hop)鏈結(從傳送端經由中繼站到達接收器)。在兩步鏈結中，中繼端會把收到的訊號作解碼並傳送重新編碼過後的訊號到接收端。在此框架中，資料偵測或融合的技術通常需要各鏈結的錯誤率資訊。舉例來說，在二元相位調變(binary phase-shift keying)合作式通訊系統中，最佳資料偵測需要知道遠端鏈結(從傳送端到中繼站的鏈結，縮寫為 SR 鏈結)的錯誤率資訊。同樣地，在資料融合中心(fusion center)，基於多感測測量值的最佳資料融合也是需要各遠端鏈結的錯誤率資訊。

為了在接收端盲蔽式地估測遠端的二進位調變(binary modulation)錯誤率，我們首先把估測問題轉換成解非線性聯立方程式的問題。每一個方程式皆反映成功匹配機率(success matching probability)和兩鏈結的錯誤率關係。其中，我們定義成功匹配機率是給定一傳送訊號下兩鏈結做出相同決策的機率。為了要讓此非線性聯立方程式有解，其基本的要求是要有足夠的鏈結數目。當鏈結數目不夠時，我們需要部分的資訊才能求出解，部分的資訊可能包含了從傳送端到接收端的鏈結(SD 鏈結)錯誤率或是從中繼站到接收端的鏈結(RD 鏈結)錯誤率。而事實上，我們發現當我們沒有從中繼站到接收端的鏈結錯誤率資訊時，即使有再多的鏈結數目亦是無法求出合理解來，這是因為從傳送端經中繼站到接收端的錯誤率會是 SR 鏈結和 RD 鏈結錯誤率的對稱函數。

為了解決上述的問題以及增加收斂速度，我們提出基於蒙地卡羅(Monte-Carlo)的估測器。具體來說，我們在接收端加入噪訊(noise)到接收訊號。在第一種方式中，我們加入噪訊是為了產生虛擬的 SD 鏈結或是 RD 鏈結，藉由這樣的方法，我們可以在即使只有一中繼站情況下亦可以得到有解的非線性聯立方程式問題。此方法我們稱為虛擬鏈結估測器(virtual link aided estimatior)。在第二種方法中，我們加入噪訊使得收到訊號的機率分佈作改變，藉由機率分佈的改變可以使得估測器的效能有所改善。

事實上，第二種方法顯示出隨機震盪（stochastic resonance）的現象，也就是說注入適當的噪訊可以改善均方根估測誤差（mean squared estimation error），此外我們發現到存在最佳的注入噪訊量使得估測效能最佳。針對此現象，我們做了一系列的分析並找出最佳的注入噪訊量為何，並藉由模擬驗證我們分析的正確性。模擬結果也顯示出訊號偵測器配合提出的盲蔽式估測器會和最佳偵測器有差不多的錯誤率效能。

對於非二進位調變通訊網路，成功匹配機率和鏈結符碼錯誤率(symbol error rate)的關係不再成立，因此，上述提到的方法不再能夠直接使用於此情況。在參考文獻[34]中，此問題被轉換成一個非線性最佳化問題，雖然其問題理論上可以解，但是其複雜度會非常的高如果中繼站數目或是 M-ary 調變的 M 值不小的時候。針對此問題，我們基於二位元表示法提出次佳的接收機以及其對應所需的錯誤率估測器。由於我們可以找到解的數學表示式，其複雜度會遠比上述所提的方法還要低得多。為了進一步改善偵測器的收斂速度，我們亦提出一噪訊增強的估測器。模擬的結果顯示我們提出的偵測器伴隨相對應的估測器會和最佳估測器有相似的效能。在均方估測錯誤(mean square estimation error)效能的評估上，我們亦可以觀察到隨機震盪的現象

Design and analysis of noise-enhanced blind error rates estimation in wireless networks

Student : Jen-Yang Liu Advisor : Y. T. Su

Institute of Communications Engineering

National Chiao Tung University

Abstract

Data detection or fusion based on multiple received copies containing the same information arises in many applications. We consider the scenario that each copy is transmitted from the same source through a different wireless link to the same destination node (DN). These links include single-hop, direct source-to-destination (SD) links and two-hop links that require an intermediate decode-and-forward (DF) node to relay the source signal. Detection or fusion under such a circumstance often need channel side information (CSI) about the link reliability. For example, maximum likelihood (ML) detection of binary modulated signals in a DF based cooperative communication network (CCN), information about the bit error rates (ERs) of the hidden source-relay (SR) links is needed. Similarly, optimal data fusion based on multiple sensor measurements requires that the ERs of various SR links be available at the fusion center.

To estimate multiple ERs blindly at the DN in a binary modulated network, we convert the estimation problem into one of solving a system of nonlinear equations. Each equation arises from the fact that the success matching probability (SMP) that a bit transmitted over two independent links connecting the same source and destination results in identical destination decisions is nonlinearly related to the ERs of the two associated links. However, the number of distinct link pairs must be larger than the

number of ERs to be estimated so that the system is not an underdetermined one. Various degrees of channel side information (CSI) about the ERs of the SD and relay-destination (RD) links is called for to remove the ambiguity arising from the insufficient number of links in the network and from that due to the symmetric nature of a cascaded source-relay-destination (SRD) link's ER as a function of its component SR and RD links' ERs.

We propose novel Monte-Carlo-based estimators that overcome all these shortcomings and accelerate the convergence speed. Our proposals involve injecting noise into the samples received by the DN. The injected noise in the first solution, called the virtual link aided (VLA) estimator, help creating virtual SD and RD links to release all CSI requirements, resolve the symmetric ambiguity and provide estimates for ERs of all component links. Using multitude of VLs, we can enhance the VLA scheme's performance and reduce the number of RNs required. The role the injected noise plays in another solution, called the importance-sampling-inspired (ISI) estimator, is different: it is used to modify link output statistics to improve the VLA estimator's convergence rate.

The latter approach exhibits a stochastic resonance effect, i.e., its mean squared estimation error (MSEE) performance is enhanced by injecting proper noise, and there exists an optimal injected noise power level that achieves the maximum improvement. The stochastic resonance effects are analyzed, and numerical examples are provided to display our estimators' MSEE behaviors, as well as to show that the ER performance of the optimal detector using the proposed estimators is almost as good as that with perfect ER information.

For nonbinary modulation based networks, a relation between the SMP of a link pair and the associated symbol error rates does not exist, hence the nonlinear system based moments approach is not directly applicable. A nonlinear optimization approach which we call LJW blind estimator [34] had been proposed. Unfortunately it requires prohibitively high computational complexity unless M (the modulation order) and the

relay numbers are small. We propose a suboptimal detector based on bit-level representation and a corresponding blind estimator to estimate the error rate of sensor nodes. The complexity of our estimator is much lower than that of LJW as we are able to obtain a closed-form salutation instead of employing an iterative algorithm for solving a nonlinear optimization. To further improve the convergence rate, we propose a noise-enhanced estimator. Simulation results show that the proposed suboptimal detector using the proposed blind estimator render negligible performance loss with respect to that of the optimal detector. A stochastic resonance phenomenon is observed in the estimator's mean square estimation error performance.

誌 謝

首先感謝我的指導教授 蘇育德 教授，謝謝教授在學術上的諄諄教誨，讓我能夠進一步體會到通訊系統的博大精深之處以及箇中樂趣，也感謝老師在生活上的幫忙，讓我能夠減經家裡的負擔而不需要為生活而憂愁。此外也感謝各口試委員的建議使得本論文能夠更加完善。

當然，我也要深深感謝 811 實驗室歷年來的學長姐和學弟妹。尤其是李昌明 學長，感謝您的分享讓我能夠在這研讀博士的期間留下美好的回憶。也感謝繃帶小孩的陪伴，沒有你的存在，我想我應該會在中途就放棄學位吧。

最後也是最重要的是我要感謝我的家人，感謝你們願意讓我專心研讀博士，包容我這任性的要求，這陣子真的是辛苦你們了。

Contents

Chinese Abstract **i**

English Abstract **iii**

Acknowledgements **vi**

Contents **vii**

List of Figures **x**

List of Tables **xiv**

List of Notations **xv**

1 Introduction **1**

1.1 Stochastic resonance 1

1.2 Motivation and dissertation overview 3

2 Adaptive blind data detection in cooperative communication network **8**

2.1 System model, ML detection and blind ER estimator 8

2.2 Side-Information-Aided Blind Single ER Estimation 11

2.3 Multiple-Relay-Aided Blind Multiple ER Estimation 13

3 Blind Multiple ERs Estimation using Virtual Links **17**

3.1 ER ambiguity in a cascaded link 17

3.2 Virtual link methods 18
3.3 Blind ER Estimation for BFSK and DPSK Signals 22
3.4 Simulation Results 24

4 Noise-Enhanced ER Estimations 31

4.1 Convergence consideration and a simple variance reduction method . . . 31
4.2 A brief introduction to importance sampling 32
4.3 An importance sampling inspired noise-enhanced estimator 35
4.4 Properties and performance analysis of the noise-enhanced estimator . . 39
4.5 Numerical Results 48

5 Data fusion and blind multiple error rate estimation in a non-binary modulation based wireless sensor network 54

5.1 System model and optimal detector 55
5.2 LJW blind ER estimator 56
5.3 Optimal/suboptimal fusion rules for nonbinary signals 57
5.4 Blind symbol/bit ERs estimator 59
5.5 Noise-enhanced ER estimations 64
5.6 Simulation results 70

6 Conclusions and future work 78

Appendix A 80

A.1 Derivation of (2.6) 80
A.2 Derivations of (2.8) and (2.10) 81
A.3 Proof of Lemma 2.1 82
A.4 Proof of Lemma 2.2 83
A.5 Derivation of (2.18) 83

Appendix B **85**

B.1 Derivation of (4.16) 85

B.2 Proof of Lemma 4.1 85

B.3 Proof of (4.20) 86

B.4 Proof of Theorem 4.2 87

B.5 Proof of Lemma 4.3 89

B.6 Proof of Theorem 4.4 90

B.7 Proof of Lemma 4.6 91

Bibliography **95**

List of Figures

1.1 Example of optimal and suboptimal decision boundary 2

1.2 A wireless multiple-relay network. 3

3.1 Ambiguity problem can be solved if the knowledge $e_{rd} = e$ is available. . 18

3.2 Block diagram of the maximum likelihood (ML) detector. 25

3.3 Bit error rate performance of the ML (solid curves), MRC (○), PLO (△) and VLA (∇) detectors. The following system parameter values are used. (i) single-relay system: $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.8$, $\theta_{sr} = 0°$, and $a^v_{sd} = a^v_{rd} = 2$ (single-relay), (ii) 2-relay system: $d_{r_1d} = d_{sr_2} = 7/10$, $\theta_{sr_1} = \theta_{sr_2} = 0°$, $P_s = 0.5P, P_{r_1} = P_{r_2} = 0.25P$, and $\theta = 45°$, (iii) 4-relay system: $d_{sd} = 10, d_{sr_1} = 5, \theta_{sr_1} = 45°, d_{sr_2} = 6, \theta_{sr_2} = 30°, d_{sr_3} = 4, \theta_{sr_3} = 60°, d_{sr_4} = 5, \theta_{sr_1} = 0°, P_p = 0.5P, P_{r_i} = 0.125P$, for $i = 1, 2, 3, 4$ and $\theta = 30°$. 27

3.4 Bit error rate performance of the ML, MRC, PLO, and VLA detectors. The following system parameter values are used. Single-relay system: $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.$, $\theta_{sr} = 45°$, and $a^v_{sd} = a^v_{rd} = 2$. 28

3.5 The nonlinear function (2.3) under various ERs. The system parameter values are $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.$, $\theta_{sr} = 45°$ and SNR= 8 dB ($e_{sr} = 0.0049$). 29

3.6 The probability density functions of y_{sd} and y_{rd} given $x = 1$ in single-relay system. The system parameter values are $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.$, $\theta_{sr} = 45°$ and SNR= 8 dB. 30

4.1 The flowchart of the proposed noise-enhanced estimation for data-aided BPSK system in Rayleigh fading channel 37

4.2 Normalized MSEE performance of the noise-enhanced estimator for data-aided point-to-point BPSK communication in Rayleigh fading channel. The channel quality is 30 dB (error rate $e = 2.5 \times 10^{-4}$). The sample size is 10000. 49

4.3 Normalized MSEE performance of the ISI-VLA scheme for (a) various binary modulated 3-link networks ($e_1 = 0.003$, $e_2 = 0.002, e_3 = 0.001$; the injected noise power is such that SH SNR=2 for link 1 and $e_1^{(w)} = e_2^{(w)} = e_3^{(w)}$) and (2) BPSK-based single-relay CCN ($e_{sr} = 0.02922$, $e_{rd} = 0.001988$, $e_{sd} = 0.04356$, $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$, $a_{sd}^{(w)} = 1$ and $a_{rd}^{(w)} = 30$). For 3-link networks, only the performance of $\widehat{e}_1$ is shown. The analytic predictions (solid curves) for these two scenarios are based on (4.23) and (4.42)–(4.44), respectively. 50

4.4 Normalized MSEE performance of VLA, VLA-EM, and EISI-VLA schemes in a BFSK-based single-relay CCN with $e_{sr} = 0.0127, e_{rd} = 5.0711 \times 10^{-5}$ and $e_{sd} = 0.0298$. Other parameter values used are: $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$, $e_{sd}^{(w)} = e_{rd}^{(w)} = 0.05$ and $n_{vl} = 30$. 51

4.5 MSEE reduction ratio (γ) performance of the ISI-VLA estimator with BFSK modulation and $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$. Part (a) is obtained by assuming $d_{sr} = 5$, SH-SNR=25 dB with the path loss exponent = 2 (which leads to $e_{sr} = 0.0016, e_{rd} = 0.0016, e_{sd} = 0.0062$). Part (b) assumes that $d_{sr} = 8$, SH-SNR=18 dB with path loss exponent = 4 so that $e_{sr} = 0.0127, e_{rd} = 5.0711 \times 10^{-5}, e_{sd} = 0.0298$. The MSEE reduction ratio of the RD link is not shown in part (b) as it is relatively small ($\sim O(10^{-3})$). 52

4.6 MSEE reduction ratio behavior of the EISI-VLA estimator for BFSK based CCN with different n_{vl}. Other system parameter values are the same as those of Fig. 4.5(b). 53

5.1 Parallel distributed detection system 55

5.2 A binary symmetric channel with parameter e 58

5.3 A binary nonsymmetric channel with parameter e^0 and e^1 60

5.4 16-QAM with Gray mapping labelling 65

5.5 16QAM error rate in Rayleigh fading with $\alpha = 2$. 67

5.6 Bit (symbol) error rate performance of various detectors with QPSK/16-QAM (MFSK) modulation and Gray mapping labelling. The qualities of the three links for QPSK modulation are denoted by $\left(\frac{E_b}{N_0}, \frac{E_b}{N_0}+1, \frac{E_b}{N_0}+2\right)$ in dB. Similarly, $\left(\frac{E_b}{N_0}, \frac{E_b}{N_0}+2, \frac{E_b}{N_0}+4\right)$ and (SNR,SNR+3,SNR+6) are the link qualities for QAM and 4-FSK, respectively. 72

5.7 Normalized MSEE performance of blind bit-level estimator in a 16-QAM-based three link wireless sensor network. The qualities of these three links are 10, 15, and 20 (dB), respectively. 73

5.8 Normalized MSEE performance of blind symbol ER estimator in a 4-FSK-based three link wireless sensor network. The qualities of these three links are 15, 20, and 25 (dB), respectively. 74

5.9 MSEE reduction ratio behavior of the noise-enhanced estimator in three links sensor network with 16-QAM modulation and Rayleigh fading channel. The qualities $\left(\frac{E_b}{N_0}\right)$ of these three links are 30, 35, and 40 (dB). Noise are injected into the three links such that all links have the same $\frac{E_b}{N_0}^{(w)}$. . 75

5.10 MSEE reduction ratio behavior of the noise-enhanced estimator in three links sensor network with 16-QAM modulation and Rayleigh fading channel. The qualities $\left(\frac{E_b}{N_0}\right)$ of these three links are 30, 35, and 40 (dB). We inject noise into the first link $\left(\left(\frac{E_b}{N_0}\right) = 30 \text{ (dB)}\right)$ such that the quality of the link is $\frac{E_b}{N_0}^{(w)}$. We keep the difference of the quality of these three links. That is, the other two links with noise injection have the quality $\frac{E_b}{N_0}^{(w)} + 5$ and $\frac{E_b}{N_0}^{(w)} + 10$ (dB). 76

5.11 MSEE reduction ratio behavior of the noise-enhanced estimator in three links sensor network with 4-FSK modulation and Rayleigh fading channel. The qualities $\left(\frac{E_b}{N_0}\right)$ of these three links are 30, 35, and 40 (dB). We inject noise into the first link $\left(\left(\frac{E_b}{N_0}\right) = 30 \text{ (dB)}\right)$ such that the quality of the link is $\frac{E_b}{N_0}^{(w)}$. We keep the difference of the quality of these three links. That is, the other two links with noise injection have the quality $\frac{E_b}{N_0}^{(w)} + 5$ and $\frac{E_b}{N_0}^{(w)} + 10$ (dB). 77

List of Tables

2.1 A blind ER estimation algorithm for a multiple-relay CCN with knowledge of $e_{r_k d}$. 15

3.1 A blind ER estimation algorithm for BPSK modulation. 22

3.2 The required CSI and the solutions of nonlinear systems under various modulations. 24

3.3 A blind ER estimation algorithm for BFSK/DPSK modulation. 24

4.1 A noise-enhanced ER estimation algorithm for data-aided BPSK system in Rayleigh fading channel. 38

4.2 A unified blind noise-enhanced ER estimation algorithm. 40

5.1 A blind ER estimation algorithm for high order modulation in sensor network. 62

5.2 The bisection method for finding $\eta = \frac{E_b}{N_0}$. 68

5.3 A noise-enhanced blind ER estimation algorithm for high order modulation in sensor network. 69

List of Notations

Symbol	Meaning
e_j^{ik}	Probability that fuscion center makes decision H_k based on the reporting of the jth sensor given H_i is true
e_{sd}	ER of the link from SN to DN
$\widehat{e}_{sd}$	Estimate of e_{sd}
$e_{sd}^{(v)}$	ER of the link from virtual SN to DN
$e_{sd}^{(w)}$	ER of the noise-enhanced link from SN to DN
e_{sr_k}	ER of the link from SN to the kth RN
$\widehat{e}_{sr_k}$	Estimate of e_{sr_k}
e_{r_kd}	ER of the link from the kth RN to DN
$e_{r_kd}^{(v)}$	ER of the link from the virtual kth RN to DN
$e_{rd}^{(w)}$	ER of the noise-enhanced link from RN to DN
$\widehat{e}_{r_kd}$	Estimate of e_{r_kd}
e_{srd}	End-to-end ER through a RN
$f_T(z;\varepsilon)$	Nonlinear function of the optimal detector in CCN
h_{sd}	Channel coefficient of the SD link
h_{sr_k}	Channel coefficient of the link from SN to the kth RN
h_{r_kd}	Channel coefficient of the link from the kth RN to DN
L	Number of sensor/relay nodes
N	Number of samples
n_{vl}	Number of virtual links

Symbol	Meaning
P_b^{psk}	ER for BPSK modulation
P_b^{dpsk}	ER for DPSK modulation
P_b^{bfsk}	ER for BFSK modulation
P_s	Transmission power of SN
P_{r_k}	Transmission power of the kth RN
p_{sk}	Success (matching) probability of $\widehat{y}_{sd}$ and $\widehat{y}_{r_k d}$
$\widehat{p}_{sk}$	Estimate of p_{sk}
p_{kl}	Success (matching) probability of $\widehat{y}_{r_k d}$ and $\widehat{y}_{r_l d}$
$\widehat{p}_{kl}$	Estimate of p_{kl}
$\widehat{p}_{(vs)r}$	Estimate of $\mathrm{P}_r(\widehat{y}_{sd}^{(v)} = \widehat{y}_{rd})$
$\widehat{p}_{s(vr)}$	Estimate of $\mathrm{P}_r(\widehat{y}_{sd} = \widehat{y}_{rd}^{(v)})$
$\widehat{p}_{(vs)(vr)}$	Estimate of $\mathrm{P}_r(\widehat{y}_{sd}^{(v)} = \widehat{y}_{rd}^{(v)})$
$q_0(\cdot),\ q_k(\cdot)$	Weight function of the optimal detector in CCN
Q_k	End-to-end ER through the kth RN
$Q_k^{(v)}$	End-to-end ER through the virtual kth RN
$\widehat{Q}_k$	Estimate of Q_k
$Q(x)$	Gaussian Q-function defined by $\int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-y^2/2\right) dy$
$Q(x, y, \rho)$	Bivariate Gaussian distribution function defined by $\frac{1/2\pi}{\sqrt{1-\rho^2}} \int_x^\infty \int_y^\infty \exp\left[-\frac{x_1^2+y_1^2-2\rho x_1 y_1}{2(1-\rho^2)}\right] dy_1 dx_1$
y_{sd}	Signal from SN to DN
$\widehat{y}_{sd}$	Hard decision of y_{sd}
y_{sr_k}	Signal from SN to the kth RN
$y_{r_k d}$	Signal from the kth RN to DN
$y_{r_k d}^{(v)}$	Virtual signal from the kth RN to DN
$\widehat{y}_{r_k d}$	Hard decision of $y_{r_k d}$

Chapter 1

Introduction

1.1 Stochastic resonance

In signal processing and communication systems, noise is usually the main factor of degrading the system performance. Hence, we always remove the noise either by filter or by some signal processing algorithms. For example, we can improve the quality of a image if we use some denoising algorithm (such as the median filter [1]). However, noise can do improve a system performance in some situations [2]-[17] and we call this phenomenon as stochastic resonance or noise benefit. This phenomenon does not only appear in signal processing or communication systems. It also occurs in the field of sensory neurons [8], circuits and measurement [7]. For the detail and other applications, the reader can refer to [9] and [10].

To realize the stochastic resonance, we consider a very simple example discussed in [11]. Consider an equal priori binary hypothesis testing problem, where x is drawn according to the zero mean Gaussian distribution with variance 1 under H_0. Under H_1, x obeys the Gaussian distribution with mean 1 and variance 1. It is easy to show that the optimal decision boundary is $x = 0.5$. However, we consider the suboptimal decision boundary: $x = 0$, as shown in Fig. 1.1. Without changing the detector structure, a simple way to improve the performance is to add a constant signal -1/2 to the received signals. This is equivalent to shift the suboptimal decision boundary to the optimal one and we have the best performance. This procedure can be viewed as a transformation

of the received signals. In [11], a transformation $g(x)$ is also proposed to achieve the performance of the optimal detector, where

$$g(x) = \begin{cases} x & \text{for } x \geq 1/2 \text{ and } x < 0 \\ -x & \text{for } 0 \leq x < 1/2 \end{cases}$$

This simple example illustrates two concepts. First, adding a proper noise (it is -1/2 in this example) can improve the performance of suboptimal detector (stochastic resonance effect). Second, a (nonlinear) transformation can also benefit and can be thought as a generalization of stochastic resonance. Hence, it may exist several ways to improve the performance. In this dissertation, we focus on the first type: adding noise.

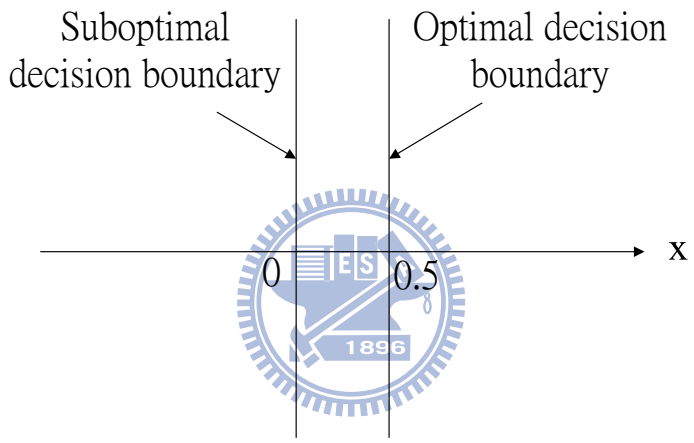


Figure 1.1: Example of optimal and suboptimal decision boundary

Several different performance measurement are used to express the stochastic resonance. For example, the performance measurement is error rate in the above example. It is also possible to use signal-to-noise ratio (SNR) [12]. Other possible performance measurements are Cramér-Rao lower bound (CRLB) [13], mutual information [14], Fisher information [15] and correlation [16]. In fact, a good measurement in a system may not be a good one in another system. [17] indicates that it is more useful to measure variations instead of SNR in biology, for instance. It seems that there is no unified performance measurement for all systems to investigate the stochastic resonance. Here, the performance measurement is the mean square estimation error (MSEE), which is

widely used in estimation problem.

1.2 Motivation and dissertation overview

We consider the basic scenario illustrated in Fig. 1.2 where the destination node (DN) d receives sequences originated from the same source node (SN) s via multiple (L) flat-fading links. These links may include a direct single-hop (SH) source-destination (SD) link and indirect two-hop links, each connecting the SN and DN with the help of an intermediate relay node (RN), say r_k. Such a scenario occurs, for example, in a cooperative communication network (CCN) in which the SD communication is aided by single or multiple relays which act as virtual antennas to allow resource sharing and provide spatial diversity gains [18]. Another popular example is the so-called CEO (Central Estimating Officer) problem associated with a wireless sensor network where each sensor sends its measurement to the CEO which often does not have direct access to the SN [19]. It is the CEO's responsibility to reliably recover the source information based on the data it has received from various sensors [20].

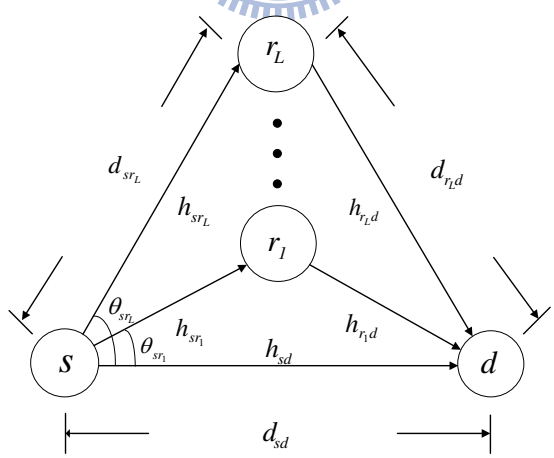


Figure 1.2: A wireless multiple-relay network.

For convenience of subsequent discourse, we define a single-relay CCN as one which consists of a source, a relay (or sensor) and a destination nodes only. We refer to the

associated SD, source-relay (SR) and relay-destination (RD) links as component links and the indirect source-relay-destination (SRD) link as cascaded link. Although many sensing-relay schemes have been proposed, we only consider the Decode-and-Forward (DF) scheme [18]-[26] for which a RN (sensor) demodulates/decodes the received signal from the SN and re-modulates/re-encodes the decoded bit stream before re-transmitting.

Since a sensor or cooperative RN may erroneously detect or sense its received signal, conventional maximum ratio combing (MRC) or similar fusion rule is no longer optimal for the DN. In fact, data fusion of various kinds in the presence of imperfect DF relays [24]-[26] and relay selection in a DF-based CCN [27] all require some forms of channel state information (CSI). Depending on the modulation used, the required CSI includes short-term CSI (ST-CSI) like instantaneous link gains and signal-to-noise ratios (SNRs) and long-term CSI (LT-CSI) such as 2average link gains and error rates (ERs) of the component links. The former has been intensively studied in terms of channel estimation, gain control and carrier recovery loops while the LT-CSI receives much less attention.

Pilot-aided ER estimators are obtainable at the cost of increasing the RNs' computing load and result in bandwidth and power efficiency reductions. The overhead and delay become significant if the true ER is small, the packet size is small or if the number of RNs is large. It is therefore desired that a DN performs all ER estimation tasks blindly.

In addition, pilot-aided ER estimators is not feasible in a sensor network due to the overhead and the property of source node. In a wireless sensor network, the sensor nodes usually are battery-limited devices Therefore, one prefers a blind estimator to reduce the overhead issues. In addition to the overhead issues, the source possibly could not transmit the pilot signals in a wireless sensor. It occurs in a main application of wireless sensor network: the environmental monitoring. [28]-[30]. In this case, one wants to detect some phenomenon, such as fire detection [31]. These two reasons enforce us to focus on the design of blind estimations.

For multiple-relay networks, the blind ER estimation problem can be transformed

into one of solving a system of nonlinear equations. Each equation describes a relation among the ERs of a pair of links and the probability that the same bit transmitted through these two independent links is decoded with identical decision. Using all available link pairs and assuming no hidden SR links, Dixit, *et al.* [32] converted the problem into a structured eigenvalue task and proposed a modified power method to find the solution. Delmas and Meurisse [33] suggested an EM-based blind ER estimator that outperforms Dixit's estimator by using the method of moments based solution of the nonlinear system as the initial estimate. These novel approaches, however, suffer from some drawbacks. First, the nonlinear system is underdetermined unless we have sufficient relays so that the number of distinct link pair combinations is no smaller than the unknown ERs. Second, even if there are enough RNs, it is not possible to simultaneously estimate all (SR, SD, and RD links) ERs and LT-CSI is needed to resolve the ambiguity resulted from the fact that the ER of a cascaded SRD link is a symmetric function of the corresponding component links' ERs. Finally, the convergence rate is slow whence it often takes a long period to obtain a reliable estimate.

It is the purpose of this dissertation to present novel blind ER estimation schemes that overcome all the above shortcomings. To simply our presentation, we focus mostly on the CCN scenario with the understanding that the proposed schemes can be readily applied in other similar scenarios. As a prelude, we briefly review a unified system model for a multiple-relay wireless network and describe the corresponding maximum likelihood (ML) detector and ER estimator structures in Chapter 2. We begin our discussion with the simplest case of a binary phase-shift keying (BPSK) based single-relay CCN, assuming the required ST- and LT-CSI's of RD and SD links are all available to the DN, i.e., the only unknown CSI which needs to be estimated is the average ER of the SR link. Even for this case, we show that blind ML ER estimation based on the DN's matched filter outputs requires high computational complexity and storage cost. A simple CSI-aided average count based estimator is thus given. We then extend

the approach to multiple-relay CCNs with less LT-CSI and obtain the basic nonlinear system (set of equations) for a 3-link (two relays plus a direct SD link) CCN and its solution. Some properties of the proposed estimator are given in the same section as well.

The main results are presented in Chapters 3-5. In Chapter 3, we discuss the ER ambiguity in a cascaded link and propose a novel approach which creates virtual SD/RD links by either rotating or injecting noise into link output samples to resolve the ambiguity and estimate all ERs without the help of multiple RNs. We show in the same chapter that the same concept can be applied to binary frequency-shift keying (BFSK) and differential phase-shift keying (DPSK) based systems. In Chapter 4, we first address the convergence rate issue and suggest a simple scheme to improve the virtual link aided (VLA) approach, using a multitude of VLs to obtain what we call the enhanced VLA (EVLA) estimator. We then proceed to propose a more subtle approach which is conceptually similar to the importance sampling (IS) based simulations and is therefore referred to as the IS-inspired (ISI) estimator. The ISI estimator also needs to inject noise into link outputs but the purpose of noise-injecting is not for building VLs but for modifying the output statistic and producing more importance events. To help the reader understand the concept of ISI estimator, we briefly introduce the importance sampling technique. A toy example is provided to express the main concept of designing the ISI estimator. Important properties of the proposed estimators and the associated MSEE performance analysis are also given in Chapter 4.

In addition to a CCN with binary modulation, we also consider a wireless sensor network with high order modulation (QAM) in Chapter 5. In this chapter, we first review the optimal detector shown in and a blind ER estimation proposed by Liu et al [34] in 2011. We will show that the proposed ER estimation has high computational complexity issue and it is almost impossible for a sensor node to implement when 16QAM or higher order modulation is applied. To solve this issue, we first approximate the

optimal detector by utilizing the bit level representation. Then, the proposed novel bit-level detector requires only the knowledge of bit error rate (BER). Hence, the unknown parameters can be reduced significantly. To improve the rate of convergence or MMSE performance, we also propose a noise-enhanced blind estimation in Chapter 5.

Simulated performance of the proposed schemes are presented in the last section of chapter 3-5 and show that the detector using the ERs estimated by our schemes yield performance almost as good as that with perfectly known ERs. Furthermore, a hybrid of the EVLA and ISI (or ISI-VLA) methods is capable of offering significant variance reduction. Both analysis and simulations prove that the ISI estimator exhibits a stochastic resonance effect, i.e., its MSEE performance is improved by injecting noise into the received samples and there exits an optimal injected noise power that achieves the maximum improvement. Finally, concluding remarks and future work are provided in Chapter 6.

Chapter 2

Adaptive blind data detection in cooperative communication network

This chapter begins with descriptions of a generic system model, assumptions and related parameter definitions. The expressions of the ML data detector and blind ER estimator are then given. The second and third subsections review some side information aided blind ER estimators for single- and multiple-relay networks. We will frequently refer to these materials in subsequent discussions.

2.1 System model, ML detection and blind ER estimator

We follow the conventional assumption of using a two-phase time division duplex cooperative communication scheme in which the SN in Fig. 1.2 transmits a sequence of independent and identical distributed (i.i.d.) ± 1-valued data $\{x[n]\}$ and all L RNs listen, decode and re-encode the received message in the first phase. The synchronous samples received by the DN and the kth RN in this phase are

$$y_{sd}[n] = h_{sd}[n]\sqrt{P_s}x[n] + w_{sd}[n], \tag{2.1.a}$$

$$y_{sr_k}[n] = h_{sr_k}[n]\sqrt{P_s}x[n] + w_{sr_k}[n], \tag{2.1.b}$$

where P_s is the signal power and the additive noise components, $w_{sd}[n], w_{sr_k}[n]$, are independent zero-mean complex white Gaussian random variables with variances σ_d^2 and

σ_r^2, respectively. We assume that the complex link gains, $h_{ij}[n]$, for the link from node i to node j, where $(i,j) \in \{(s,r_k),\ (s,d),\ (r_k,d); k = 1,\cdots,L\}$, and the corresponding noise terms, $w_{ij}[n]$, are mutually independent. The RNs send the re-encoded message to the DN in the second phase. Since RNs may detect erroneously, the re-transmitted signals are not necessarily equal to $x[n]$. If we denote by $\hat{x}_{r_k}[n]$ the signal sent by the kth relay and $y_{r_kd}[n]$ the corresponding received sample at the DN in this phase, then

$$y_{r_kd}[n] \quad = \quad h_{r_kd}[n]\sqrt{P_{r_k}}\hat{x}_{r_k}[n] + w_{r_kd} \tag{2.2}$$

where P_{r_k} is the transmitted signal power of the kth RN and $w_{r_kd}[n]$ has the same distribution as $w_{sd}[n]$. For frequency-flat fast Rayleigh fading links, $|h_{ij}|^2$ are independent exponentially distributed random variables with variance σ_{ij}^2.

Define the memoryless nonlinearity

$$f_T(z;\varepsilon) = \ln\left[\frac{\varepsilon + (1-\varepsilon)e^z}{(1-\varepsilon) + \varepsilon e^z}\right], \quad 0 < \varepsilon < 1/2 \tag{2.3}$$

and, for $k = 1,\cdots,L$, the weighting functions

$$q_0(y[n]) = \Re\left\{4h_{sd}^*[n]\sqrt{P_s}y[n]/\sigma_d^2\right\}, \tag{2.4.a}$$

$$q_k(y[n]) = \Re\left\{4h_{r_kd}^*[n]\sqrt{P_{r_k}}y[n]/\sigma_d^2\right\}, \tag{2.4.b}$$

where $\Re\{z\}$ denotes the real part of z. Then the ML detector for BPSK signals is given by [24]

$$\hat{x}[n] = \mathrm{sgn}\left[q_0(y_{sd}[n]) + \sum_{i=1}^{L} f_T\left(q_k(y_{rd}[n]); e_{sr_k}\right)\right] \tag{2.5}$$

where $\mathrm{sgn}[z]$ denotes the sign of the real number z and e_{sr_k} is the ER of the link between the source and the kth RN. (2.3) and (2.4.a)-(2.4.b) indicate that besides the instantaneous received complex amplitude-to-noise-power ratio, $\frac{\sqrt{P_{r_k}}h_{r_kd}[n]}{\sigma_d^2}$ and $\frac{\sqrt{P_s}h_{sd}[n]}{\sigma_d^2}$, the hidden SR link's ER, e_{sr_k}, should also be known by the DN for ML detection. As the instantaneous complex links gains $h_{r_kd}[n]$ and $h_{sd}[n]$ are difficult to estimate in a high dynamic wireless environment, noncoherent signals are sometimes preferred for they

require no such estimations. Nevertheless, [25] and [26] show that ML noncoherent detections of BFSK and DPSK signals by a DN still need the LT-CSI such as ERs for both far-end (SR) and near-end (SD and RD) links or σ_d^2.

For notational brevity, we henceforth omit the subscript k associated with the kth relay, r_k, unless there is danger of ambiguity. The DN of a single-relay BPSK-based CCN has the samples $\{y_{sd}[n], y_{rd}[n]\}$ of (2.1.a) and (2.2) as the sufficient statistics for estimating the BERs of its component links. As an i.i.d. source is assumed, we can easily verify that the probability density function (pdf) of $y_{sd}[n]$ is independent of e_{sr} and so is that of $y_{rd}[n]$. With N coherently received sample pairs, $\{(q_0(y_{sd}[i]), q_1(y_{rd}[i]))\}_{i=1}^N \triangleq \{(q_0^{(i)}, q_1^{(i)})\}$, the joint conditional pdf $f(y_{sd}, y_{rd}|\mathrm{I}_{csi})$ of the matched filter outputs, y_{sd} and y_{rd} given CSI, $\{h_{sd}, h_{rd}, \sigma_d^2, e_{sr}\}= \mathrm{I}_{csi}$, and unit transmit powers, $P_s = P_r = 1$ is a mixture density and the ML blind e_{sr} estimator is given by (see Appendix A.1)

$$\begin{aligned}\hat{e}_{sr} &= \arg \max_{0\le e_{sr}<0.5} \log f(\{y_{sd}[i]\}_{i=1}^N, \{y_{rd}[i]\}_{i=1}^N|\mathrm{I}_{csi}) \\ &= \arg \max_{0\le e_{sr}<0.5} \sum_{i=1}^N \log\left[\cosh\left(\frac{q_0^{(i)}+q_1^{(i)}}{2}\right) - 2\sinh\left(\frac{q_0^{(i)}}{2}\right)\sinh\left(\frac{q_1^{(i)}}{2}\right)e_{sr}\right] \end{aligned} \tag{2.6}$$

The reliability of the ML estimator depends on the sample size N, the true e_{sr} and two other component links' statistics which, in turn, determine those of $q_0^{(i)}$ and $q_1^{(i)}$. For practical ERs, we usually need large N for the ML estimator to converge. The difficulty in implementing this estimator comes at least from three other concerns: (i) the computing complexity of solving the associated nonconvex optimization problem, (ii) there exists no recursive formula for updating the objective function whenever a new received signals pair becomes available, and (iii) the large required storage space. These implementation considerations convince us to turn to estimators based on the binary sample sequence $\{\widehat{y}_{rd}[n], \widehat{y}_{sd}[n]\}$ produced by

$$\widehat{y}_{rd}[n] = \mathrm{sgn}\left[q_1(y_{rd}[n])\right], \quad \widehat{y}_{sd}[n] = \mathrm{sgn}\left[q_0(y_{sd}[n])\right], \tag{2.7}$$

Besides their simplicity, an important advantage of such estimators is that they can be

easily extended to noncoherent binary modulations while the form of the ML estimator is highly modulation-dependent.

As a prelude to the study of simultaneous blind estimation of all component links' ERs, we start with the simpler case of SR link ER estimation, assuming the ST-CSI needed and the ERs of either all or some of the remaining component links are available.

2.2 Side-Information-Aided Blind Single ER Estimation

Since a cascaded link is composed of two (i.e. SR and RD) binary symmetric links (BSLs) with ERs, e_{sr} and e_{rd}, the end-to-end ER e_{srd} is given by $e_{srd} = e_{sr}(1-e_{rd})+(1-e_{sr})e_{rd} = e_{sr} + e_{rd} - 2e_{sr}e_{rd}$. A single-relay CCN can thus be regarded as the composition of two BSLs connecting the source and the destination. We assume stationary component links with time-invariant ERs and refer to the probability $p = \mathrm{P}_r\left(\widehat{y}_{sd} = \widehat{y}_{rd}\right)$ as the success matching probability (SMP). Using the identity $p = e_{sd}e_{srd} + (1 - e_{sd})(1 - e_{srd})$ and the i.i.d. source assumption, we immediately have following identity which relates various ERs to the SMP between a direct SD link and a cascaded SRD link

$$e_{sr} = \frac{1 - e_{sd} - e_{rd} + 2e_{sd}e_{rd} - p}{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}} \stackrel{def}{=} e_{sr}(p) \tag{2.8}$$

Since the links are assumed to be stationary, $W[i] \stackrel{def}{=} I\left(\widehat{y}_{sd}[i] = \widehat{y}_{rd}[i]\right)$, where $I(\mathbf{E}) = 1$ if the statement $\mathbf{E}$ is true; otherwise it is zero, is Bernoulli distributed with success probability p. Furthermore, the SMP can be estimated by

$$\widehat{p}^{(N)} = \sum_{i=1}^{N} \frac{I\left(\widehat{y}_{sd}[i] = \widehat{y}_{rd}[i]\right)}{N}, \tag{2.9}$$

where the superscript (N) indicates that N sample pairs are used to obtain the estimator. This average count based estimator is the sample mean of the Bernoulli process $\{W[i]\}$ and is an uniform minimum variance unbiased estimator if i.i.d. samples are received [35].

Using the sample mean estimator (A.2) as $\hat{p}$, the method of moments and (2.8) suggests the estimator

$$\widehat{e}_{sr} = \frac{1 - e_{sd} - e_{rd} + 2e_{sd}e_{rd} - \widehat{p}}{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}} = e_{sr}(\hat{p}) \tag{2.10}$$

if both e_{rd} and e_{sd} are known. The derivations of (2.8) and (2.10) are given in Appendix A.2.

As $0 \leq e_{sr} \leq 0.5$, our estimator $\widehat{e}_{sr}$ may have to be modified by the soft-limiter

$$\mathcal{J}(\widehat{e}_{sr}) = \min\left[\max(\widehat{e}_{sr}, 0), 0.5\right] \tag{2.11}$$

In addition, we can easily derive a recursive relation for $\widehat{p}^{(i)}$ to sequentially estimate p and therefore e_{sr}.

The ER estimator (2.10) has many desired properties which we summarize in the following two lemmas.

Lemma 2.1. The estimator $\hat{e}_{sr}$ defined by (2.10) is (i) unbiased and attains the Cramer-Rao lower bound (CRLB), (ii) an uniformly minimum variance unbiased (UMVU) and ML estimator with variance

$$\mathrm{Var}(\hat{e}_{sr}) = \frac{p(1-p)}{N(1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd})^2}. \tag{2.12}$$

where N is sample size.

Proof. See Appendix A.3. □

Lemma 2.2. For any $\varepsilon > 0$, we have,

$$\Pr(|\widehat{e}_{sr} - e_{sr}| \geq \varepsilon) \leq 2\exp\left[-\min\left(\frac{N^2\varepsilon^2 C_1^2}{4p}, \frac{N\varepsilon C_1}{2}\right)\right] \tag{2.13}$$

where $C_1 = 1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}$, N is sample size, and the soft-limiting effect (2.11) is neglected.

Proof. See Appendix A.4. □

The properties given in *Lemma 2.1* are resulted from the fact that $\hat{e}_{sr}$ is a linear function of $\hat{p}$ and the invariance property of an ML estimator. *Lemma 2.2*, which is derived from using Chernoff's inequality, implies that the estimator $\widehat{e}_{sr}$ converges to e_{sr} in probability.

2.3 Multiple-Relay-Aided Blind Multiple ER Estimation

When there are L RNs, we have $\binom{L+1}{2}$ *combinatorial diversities* from pairwise hard-decision matchings. For any (k, l) RN pair, $k \neq l$, the random variable, $W_{kl} = I\left(\widehat{y}_{r_k d} = \widehat{y}_{r_l d}\right)$, is Bernoulli distributed with success (matching) probability p_{kl} =$\mathrm{P}_r[\widehat{y}_{r_k d} = \widehat{y}_{r_l d}]$ which satisfies the identity

$$p_{kl} = Q_k Q_l + (1 - Q_k)(1 - Q_l) \tag{2.14}$$

with Q_k being the cascaded link ER given by

$$Q_k = e_{sr_k} + e_{r_k d} - 2e_{sr_k} e_{r_k d} \stackrel{def}{=} e_{sr_k d} \tag{2.15}$$

The above equations and (2.8) imply that p_{sr_k} and p_{kl} are related to the parameter sets $\{e_{sd},\ e_{sr_k},\ e_{r_k d}\}$ and $\{e_{sr_k},\ e_{r_k d},\ e_{sr_l},\ e_{r_l d}\}$, respectively. Following the approach used for the case $L = 1$, we replace p_{sr_k} and p_{kl} in (2.8) and (2.14) by the *average sample count* (sample mean) estimators

$$\widehat{p}_{sr_k} = \sum_{j=1}^{N} \frac{I\left(\widehat{y}_{sd}[j] = \widehat{y}_{r_k d}[j]\right)}{N},\ k = 1, \cdots, L \tag{2.16.a}$$

$$\widehat{p}_{kl} = \sum_{i=1}^{N} \frac{I\left(\widehat{y}_{r_k d}[i] = \widehat{y}_{r_l d}[i]\right)}{N},\ 1 \leq k < l \leq L \tag{2.16.b}$$

to obtain $\binom{L+1}{2}$ equations, all of the form similar to (2.14), involving the unknown ERs, $\{Q_i\}$ and e_{sd}.

When the RNs are dedicated stationary nodes and $\{e_{r_k d}\}$ can be reliably estimated, there are only $L+1$ unknown parameters, $\{e_{sd}, e_{sr_k}, k = 1, \cdots, L\}$, which can be solved if

there are at least $L+1$ independent equations. Since $\binom{L+1}{2} \geq L+1$ whenever $L \geq 2$, the unknown link parameters can be estimated as long as more than two RNs are available.

For general multiple ER estimation in an L-relay CCN, $L > 2$, we can divide the problem into a sequence of subproblems, each deals with a smallest two-relay problem. For the detail, please refer to Table 2.1. The three-link (two relays plus a direct SD link) CCN is referred to as a basic network in which the link ER is governed by a set of nonlinear equations called a *basic (nonlinear) system*.

$$\begin{bmatrix} 1 - Q_1 - e_{sd} + 2e_{sd}Q_1 \\ 1 - Q_1 - Q_2 + 2Q_1Q_2 \\ 1 - e_{sd} - Q_2 + 2e_{sd}Q_2 \end{bmatrix} = \begin{bmatrix} p_{sr_1} \\ p_{12} \\ p_{sr_2} \end{bmatrix} \approx \begin{bmatrix} \widehat{p}_{sr_1} \\ \widehat{p}_{12} \\ \widehat{p}_{sr_2} \end{bmatrix} \tag{2.17}$$

where $\widehat{p}_{sr_l}$, $l = 1, 2$, and $\widehat{p}_{12}$ are obtained via (2.16.a) and (2.16.b). A similar nonlinear system arose in [32] where the estimations of the error rates e_{sd} and Q_l's were attempted. Unlike our case, there is no cascaded links and hence no need to estimate the ERs of the SR and RD links. It can be shown that the solution to the above *basic system* gives the *basic estimators* [36] (The derivation is given in Appendix A.5)

$$\widehat{Q}_i = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{(2\widehat{p}_{ij} - 1)(2\widehat{p}_{ik} - 1)}{2\widehat{p}_{jk} - 1}}, \quad i, j, k \in \{0, 1, 2\} \tag{2.18}$$

where $\widehat{Q}_0 = \widehat{e}_{sd}$, $\widehat{p}_{01} = \widehat{p}_{sr_1}$, and $\widehat{p}_{02} = \widehat{p}_{sr_2}$.

The above equation indicate that the presence of multiple RD links enables us to estimate e_{sd} and removes the need for e_{sd} side information, i.e., the relay diversity can be traded for the degree of LT-CSI. To estimate the ERs of the multiple hidden (far-end) SR links, we invoke the relation (2.15), assuming the ERs of all RD links are known, to obtain

$$\widehat{e}_{sr_k} = \frac{\widehat{Q}_k - e_{r_k d}}{1 - 2e_{r_k d}}, \; k = 1, 2. \tag{2.19}$$

The flowchart of the proposed estimator is summarized in Table 2.1.

Note that an L-relay CCN induces $\binom{L+1}{2}$ basic systems (diversities) where each relay is involved in more than one system so that multiple estimates for a given Q_i may be

Table 2.1: A blind ER estimation algorithm for a multiple-relay CCN with knowledge of $e_{r_k d}$.

Input:	Received sample $\mathbf{y}$ and $e_{r_k d}$.
1.	Compute the SMP estimates $\widehat{p}_{sr_k}$, and $\widehat{p}_{(2\ell-1)2\ell}, \ell = 1, 2, \cdots, \lfloor L/2 \rfloor$ by (2.16.a) and (2.16.b), add $\widehat{p}_{1L}$ if L is odd.
2.	For $k = 2\ell, \ell = 1, 2, \cdots$, compute (i) $(\widehat{e}_{sd}, \widehat{Q}_{k-1}, \widehat{Q}_k)$ by (2.18) using $\widehat{p}_{sr_{k-1}}$, $\widehat{p}_{sr_k}$, and $\widehat{p}_{(k-1)k}$
3.	Compute the ERs, $\widehat{e}_{sr_k}$ by (2.19) using $e_{r_k d}$ and $\widehat{Q}_k$ An improved $\widehat{e}_{sd}$ is obtained by taking average of all the $\widehat{e}_{sd}$ computed in 2.
Output:	$\widehat{e}_{sd}$ and $\widehat{e}_{sr_k}$

obtained. Dixit [32] had proposed a complex method to take advantage of this fact and obtained improved ER estimates. On the other hand, [33] shows that the basic estimators given by (2.18) asymptotically achieve the accuracy achieved by the ML pilot-aided estimator based on the two sequences of hard-decision pairs, $\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}_{i=1}^N$, $l = 1, 2$, for finite N, a better estimate is obtained by maximizing the log-likelihood functions, $\Gamma(\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}) \stackrel{def}{=} \log f(\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}_{i=1}^N)$, $l = 1, 2$, defined as

$$\begin{aligned}
&\Gamma(\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}) \\
&= \prod_{i=1}^{N} \left(e_{sd}^{1-I(\widehat{y}_{sd}[i]=x[i])} (1-e_{sd})^{I(\widehat{y}_{sd}[i]=x[i])} \prod_{l=1}^{2} Q_l^{1-I(\widehat{y}_{r_l d}[i]=x[i])} (1-Q_l)^{I(\widehat{y}_{r_l d}[i]=x[i])} \right)
\end{aligned}$$

The derivation of the above function is similar to that given in [33, Section III] with additional consideration of cascaded link ER Q_l. In [33] an EM based approach was proposed to obtain blind (unknown $x[i]$) estimates of Q_i which outperforms Dixit's method. However, our numerical experiments conclude that, for both approaches, *the performance improvement over the basic estimators is rather limited and do not worth the additional high complexity*; see Section VI and Fig. 4.4.

Before presenting our main results in the following sections, we would like to emphasize that most estimators to be developed are based on some variation or extension of the basic system (2.17) and their expressions, e.g., (3.5.a)-(3.6), (3.9.a)-(3.9.d), and

(3.11.a)-(3.12.b), are derivable from variations or extensions of the basic estimators, (2.18) and (2.19).

Chapter 3

Blind Multiple ERs Estimation using Virtual Links

We first examine the ER ambiguity issue associated with the estimation of a far-end component link's ER and then present a novel solution to resolve this ambiguity. The extension to other binary modulations–BFSK and DPSK–is discussed at the end of this chapter.

3.1 ER ambiguity in a cascaded link

As can be seen from (2.17), when there are sufficient relays, the resulting equation set leads to formulae for the estimates of e_{sd} and Q_k but not those for e_{sr_k} and e_{r_kd}. This is due to the fact that the ER of an SRD link, as (2.15) has shown, is a symmetric function of the ERs of the associated component SR and RD links, i.e., there are infinite many (e_{sr_k}, e_{r_kd}) pairs that result in the same Q_k. In fact, the legitimate candidates for the latter two ERs consist of the lower-left part of the hyperbola defined by (2.15), $(1-2Q_k)/4 = (e_{sr_k} - \frac{1}{2})(e_{r_kd} - \frac{1}{2})$, that lies within the square $\mathcal{S} \stackrel{def}{=} \{(e_{sr_k}, e_{r_kd}) | 0 < e_{sr_k} < \frac{1}{2}, 0 < e_{r_kd} < \frac{1}{2}\}$. The ambiguity in (2.15) is resolved in the scenario discussed in the last section by specifying e_{r_kd} so that $\widehat{e}_{sr_k}$ is obtained via (2.19) (see Fig. 3.1). Geometrically, this is equivalent to finding the intersection of the hyperbola and the line $e_{r_kd} = e$ within the square $\mathcal{S}$, where e is the true ER of the RD link.

When the LT-CSI, e_{r_kd}, is not available, we need to find a curve which represents

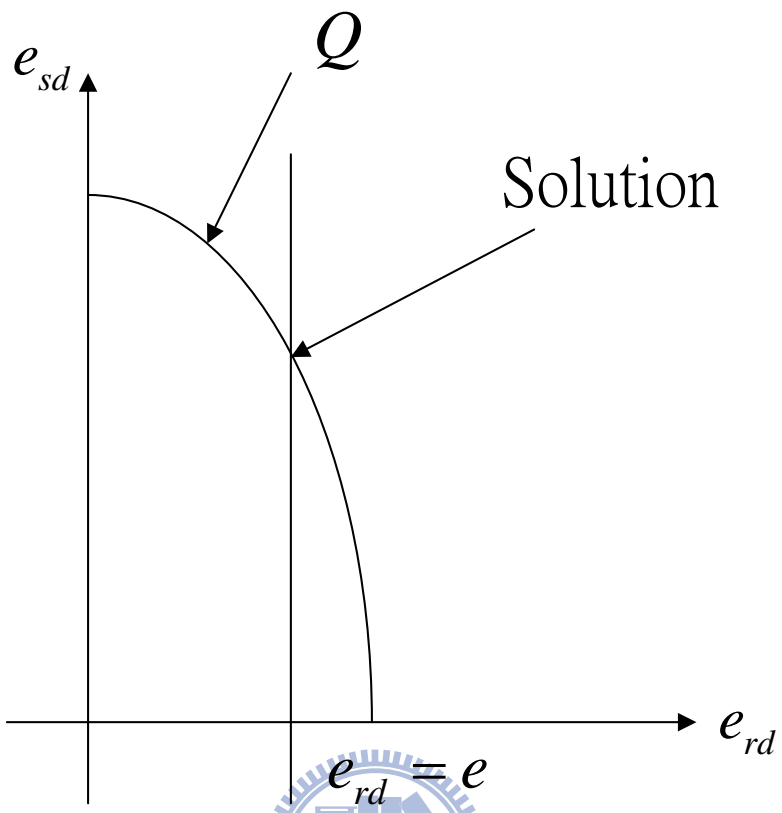


Figure 3.1: Ambiguity problem can be solved if the knowledge $e_{rd} = e$ is available.

another set of legitimate ER pairs and which has only one intersection point with (2.15) in $\mathcal{S}$. Since the hyperbola is symmetric with respect to the line $e_{r_k d} = e_{sr_k}$ and we have access to the outputs of the RD and SD links only, finding a curve which has a unique intersection with (2.15) is possible if an alternate RD link is provided. This can be seen by noting that a RD link with a different average bit SNR $\overline{\gamma}$ yields a different equivalent cascaded link with ER Q'_k and therefore a curve of the form $(1 - 2Q'_k)/4 = (e_{sr_k} - \frac{1}{2})(\alpha e_{r_k d} - \frac{1}{2})$, where α is such that $0 < \alpha e_{r_k d} \triangleq e_{r'_k d} < \frac{1}{2}$.

3.2 Virtual link methods

To have an alternate physical link (PL), one can purposely vary the power of the bit stream so that the transmitted sequence is equivalent to one formed by multiplexing two

data sources with different powers. If the locations of these two parts in the multiplexed data stream are known, the DN then perform separate comparison and counting based on (2.16.a) and (2.16.b). Although such a two-level amplitude modulation makes possible solving the e_{sr_k}, e_{r_kd} ambiguity, allocating unequal powers to different parts of the transmitted data stream is often undesirable. This dilemma can be avoided by creating a virtual link (VL) without modifying the existing link.

A VL can be created by rotating the received I-Q vector counter-clockwise by an angle θ between 0° and 90°. This is equivalent to introducing an artificial phase offset to the received samples which are then used as outputs from another link. Since the noise is circular symmetric, the rotation results in an equivalent signal power degradation $\cos^2\theta$ without altering the noise statistic. Such a virtual SNR loss cannot be accomplished by simply multiplying the BPSK matched filter output by a positive constant less than one.

An alternate method is to add an extra zero-mean white Gaussian noise component to the received in-phase samples. Both schemes give a VL with a smaller $\overline{\gamma}$. The second scheme–the addition of a perturbation term–incurs no hardware increase but requires the estimation of noise power σ_d^2, which is needed in subsequent ML detection anyway. As the phase-rotation scheme leads to an SNR degradation of magnitude $\cos^2\theta$, the second scheme has to generate i.i.d. zero-mean Gaussian random samples with variance $\sigma_v^2 = \sigma_d^2(1/\cos^2\theta - 1)$ to achieve the same SNR loss. Although both approaches achieve the same effect for BPSK signals, the phase-rotating approach cannot produce a VL for noncoherent systems while the method of inserting extra noise suits both coherent and noncoherent applications. Hence, except for the coherent system discussed in this section, we will adopt the noise-injection approach in the following sections.

We use the superscript (v) to indicate that a parameter is associated with a VL, i.e., the kth RD link's synchronous output samples and their rotated (VL) versions are denoted by $y_{r_kd}[n]$, $y_{r_kd}^{(v)}[n]$ and the corresponding ERs by e_{r_kd} and $e_{r_kd}^{(v)}$. Since for a BPSK

system operating in a flat Rayleigh fading environment, we have [37]

$$P_b^{psk}(\overline{\gamma}) = \frac{1}{2}\left(1 - \sqrt{\frac{\overline{\gamma}}{1+\overline{\gamma}}}\right), \tag{3.1}$$

which is equivalent to

$$\overline{\gamma} = \frac{(1-2P_b^{psk})^2}{1-(1-2P_b^{psk})^2}. \tag{3.2}$$

The two ERs are then related by

$$\frac{(1-2e_{r_kd})^2}{1-(1-2e_{r_kd})^2} = \frac{1}{\cos^2\theta}\frac{(1-2e_{r_kd}^{(v)})^2}{1-(1-2e_{r_kd}^{(v)})^2}. \tag{3.3}$$

Following a procedure similar to that for solving (2.17), we can easily show that the nonlinear system which consists of (2.15), (3.3) and the new cascaded link's ER equation

$$Q_k^{(v)} = e_{sr_k} + e_{r_kd}^{(v)} - 2e_{sr_k}e_{r_kd}^{(v)} \tag{3.4}$$

has the closed-form solution

$$e_{sr_k} = \frac{1-\sqrt{1-4t}}{2}, \quad e_{r_kd} = \frac{Q_k - e_{sr_k}}{1-2e_{sr_k}}, \tag{3.5.a}$$

$$e_{r_kd}^{(v)} = \frac{Q_k^{(v)} - e_{sr_k}}{1-2e_{sr_k}}. \tag{3.5.b}$$

where

$$t = \frac{(1-2Q_k^{(v)})^2 Q_k(1-Q_k)}{(1-2Q_k^{(v)})^2 - \cos^2\theta(1-2Q_k)^2} - \frac{\cos^2\theta(1-2Q_k)^2 Q_k^{(v)}(1-Q_k^{(v)})}{(1-2Q_k^{(v)})^2 - \cos^2\theta(1-2Q_k)^2}. \tag{3.6}$$

Based on this solution, we can obtain a complete blind algorithm to estimate the ERs of all component links by using the estimates for Q_k and $Q_k^{(v)}$ which are computed via (2.18) using another, say lth ($l \neq k$) relay link; ER side information is no longer needed. In short, to estimate the triplet $(e_{sd}, e_{sr_k}, e_{r_kd})$ associated with an SD and an SRD links without the help of CSI, one needs another independent relay. The auxiliary relay requirement can be waived if one creates a virtual SD link to obtain additional combinational diversities. In general, the rotation angle for producing a virtual SD link can be different from that for a virtual RD link. However, we lose no generality by

assuming both rotation angles are the same, say θ. Denote by $\widehat{p}_{(vs)r}$, $\widehat{p}_{s(vr)}$ and $\widehat{p}_{(vs)(vr)}$ the estimates for the SMPs, $\mathrm{P}_r(\widehat{y}_{sd}^{(v)} = \widehat{y}_{rd})$, $\mathrm{P}_r(\widehat{y}_{sd} = \widehat{y}_{rd}^{(v)})$, and $\mathrm{P}_r(\widehat{y}_{sd}^{(v)} = \widehat{y}_{rd}^{(v)})$, respectively, and by $Q = e_{srd}, Q^{(v)} = e_{s(vr)d}$, the ERs for the SRD and the SR-plus-virtual relay links. We obtain four nonlinear relations for a single-relay CCN:

$$\widehat{p}_{sr} = e_{sd}Q + (1 - e_{sd})(1 - Q) \tag{3.7.a}$$

$$\widehat{p}_{(vs)r} = e_{sd}^{(v)}Q + (1 - e_{sd}^{(v)})(1 - Q) \tag{3.7.b}$$

$$\widehat{p}_{s(vr)} = e_{sd}Q^{(v)} + (1 - e_{sd})(1 - Q^{(v)}) \tag{3.7.c}$$

$$\widehat{p}_{(vs)(vr)} = e_{sd}^{(v)}Q^{(v)} + (1 - e_{sd}^{(v)})(1 - Q^{(v)}) \tag{3.7.d}$$

With the additional PL-VL relation

$$\frac{(1 - 2e_{sd})^2}{1 - (1 - 2e_{sd})^2} = \frac{1}{\cos^2\theta}\frac{(1 - 2e_{sd}^{(v)})^2}{1 - (1 - 2e_{sd}^{(v)})^2} \tag{3.8}$$

the nonlinear system (3.7.a)–(3.8) yields the closed-form estimators

$$\widehat{e}_{sd} = \frac{1}{2}\left[\frac{1 - \widehat{p}_{sr} - \widehat{Q}}{1 - 2\widehat{Q}} + \frac{1 - \widehat{p}_{s(rv)} - \widehat{Q}^{(v)}}{1 - 2\widehat{Q}^{(v)}}\right] \tag{3.9.a}$$

$$\widehat{Q} = \frac{1 - \sqrt{1 - 4t_1}}{2}, \quad \widehat{Q}^{(v)} = \frac{1 - \sqrt{1 - 4t_2}}{2} \tag{3.9.b}$$

$$t_1 = \frac{\cos^2\theta(2\widehat{p}_{sr} - 1)^2(\widehat{p}_{(vs)r} - 1)\widehat{p}_{(vs)r}}{(2\widehat{p}_{(vs)r} - 1)^2 - \cos^2\theta(2\widehat{p}_{sr} - 1)^2} - \frac{(2\widehat{p}_{(vs)r} - 1)^2(\widehat{p}_{sr} - 1)\widehat{p}_{sr}}{(2\widehat{p}_{(vs)r} - 1)^2 - \cos^2\theta(2\widehat{p}_{sr} - 1)^2} \tag{3.9.c}$$

$$t_2 = \frac{\cos^2\theta(2\widehat{p}_{s(vr)} - 1)^2(\widehat{p}_{(vs)(vr)} - 1)\widehat{p}_{(vs)(vr)}}{(2\widehat{p}_{(vs)(vr)} - 1)^2 - \cos^2\theta(2\widehat{p}_{s(vr)} - 1)^2} - \frac{(2\widehat{p}_{(vs)(vr)} - 1)^2(\widehat{p}_{s(vr)} - 1)\widehat{p}_{s(vr)}}{(2\widehat{p}_{(vs)(vr)} - 1)^2 - \cos^2\theta(2\widehat{p}_{s(vr)} - 1)^2}. \tag{3.9.d}$$

Estimators, $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$, can be derived from solving the nonlinear system which includes (2.15), (3.3) and an equation similar to (3.4). An analytic solution of this nonlinear system is obtained by substituting (3.9.b) into (3.6) and then (3.5.a). As has been mentioned in Section I, we refer to ER estimation algorithms using the approach described in this section as virtual link aided (VLA) estimators. The corresponding estimation procedure is included in Table 3.1.

Note that the SMP formulae (2.14) and (3.7.a)-(3.7.d) are not valid for the SMP between a PL and its virtual version since their outputs are correlated. Actually, this

Table 3.1: A blind ER estimation algorithm for BPSK modulation.

Input:	Received samples, $\mathbf{y}$, noise variance, σ_d^2, and scaling factor values, $a_{sd}^{(v)}$, $a_{rd}^{(v)}$.
1:	Create virtual SD and RD links by injecting complex Gaussian noise samples with scaling factors, $a_{sd}^{(v)}$ and $a_{r_i d}^{(v)}$.
2:	Compute SMPs for all physical, virtual SD-SRD link pairs.
3:	Compute $\widehat{Q}$, $\widehat{Q}^{(v)}$, and $\widehat{e}_{sd}$ through (3.9.b) and (3.9.a) with $a_{sd}^{(v)} = a_{rd}^{(v)} = \frac{1}{\cos^2\theta}$.
4:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ via (3.5.a)-(3.6).
Output:	$\widehat{e}_{sd}$, $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$.

SMP is the sum of two conditional SMPs defined by (4.28) and (4.29) which are derived in Appendix D. Obviously, a system involves these two nonlinear expressions does not easily render a closed-form solution. On the other hand, a VL can provide a new SMP relation similar to (2.14) with each different PL or its virtual version and a single-relay CCN can offer two uncorrelated VLs to render a basic system that consists of three independent SMP equations, we thus conclude that, by using both virtual RD and SD links, one can estimate all ERs of a single-relay CCN without side information.

3.3 Blind ER Estimation for BFSK and DPSK Signals

Although we have limited our discussion to BPSK signals so far, such a restriction does not lose any generality as far as the VL concept is concerned. The proposed blind estimation method in the last section can easily be extended to noncoherent binary modulations because our estimation scheme can be applied for any binary symmetric channel. Besides using (noncooperative) noncoherent detectors, the DN adds a complex Gaussian perturbation term to each of the received noncoherent sample to generate the corresponding VL with the desired equivalent average SNR.

As in the BPSK case, we first perform the noncoherent detection to obtain the hard

decisions. Since, for the noncoherent case, the definition and estimation of SMPs are the same as those of the BPSK-based system, we have four nonlinear equations similar to (3.7.a)–(3.7.d) which relate the SMPs to the corresponding ERs of the connecting SD and cascaded SRD links. The relation between the ER of a cascaded link and its two component links remains the same we thus obtain two equations similar to (2.15) and (3.4). However, as different modulation type is involved, the equation governing the relation between e_{sd}'s for the physical and the virtual links is different from (3.3), so is that between the two e_{rd}'s. The new relation can be expressed in the generic form

$$F_e(z) = a^{(v)} F_e(z^{(v)}), \tag{3.10}$$

where $z = e_{sd}$ or e_{rd} and, as before, the superscript (v) on the right-hand side denotes the corresponding set of parameters for the VL. (3.10) is similar to (3.3) but the actual expression for $F_e(z)$ depends on the modulation used and $a^{(v)}$ is a scaling parameter related to the variance of the injected noise (normalized with respect to σ_d^2).

Solving the nonlinear system consisting of the SMP equations and $F_e(e_{sd}) = a^{(v)} F_e(e_{sd}^{(v)})$, we obtain

$$\widehat{Q} = \frac{(1-2\widehat{p}_{sr})\left[1-\widehat{p}_{(vs)r}\right] - a^{(v)}\left[1-2\widehat{p}_{(vs)r}\right](1-2\widehat{p}_{sr})}{(1-2\widehat{p}_{sr}) - a^{(v)}\left[1-2\widehat{p}_{(vs)r}\right]} \tag{3.11.a}$$

$$\widehat{Q}^{(v)} = \frac{\left[1-2\widehat{p}_{s(vr)}\right]\left[1-\widehat{p}_{(vs)(vr)}\right]}{\left[1-2\widehat{p}_{s(vr)}\right] - a^{(v)}\left[1-2\widehat{p}_{(vs)(vr)}\right]} - \frac{a^{(v)}\left[1-2\widehat{p}_{(vs)(vr)}\right]\left[1-2\widehat{p}_{s(vr)}\right]}{\left[1-2\widehat{p}_{s(vr)}\right] - a^{(v)}\left[1-2\widehat{p}_{(vs)(vr)}\right]} \tag{3.11.b}$$

$$\widehat{e}_{sd} = \frac{1}{2}\left[\frac{1-\widehat{p}_{sr}-\widehat{Q}}{1-2\widehat{Q}} + \frac{1-\widehat{p}_{s(vr)}-\widehat{Q}^{(v)}}{1-2\widehat{Q}^{(v)}}\right] \tag{3.11.c}$$

Similarly, we have

$$\widehat{e}_{sr} = \frac{\widehat{Q}^{(v)} - 2\widehat{Q}\widehat{Q}^{(v)} - a^{(v)}\widehat{Q} + 2a^{(v)}\widehat{Q}\widehat{Q}^{(v)}}{1-2\widehat{Q} - a^{(v)} + 2a^{(v)}\widehat{Q}^{(v)}} \tag{3.12.a}$$

$$\widehat{e}_{rd} = \frac{\widehat{Q} - \widehat{e}_{sr}}{1-2\widehat{e}_{sr}}. \tag{3.12.b}$$

The explicit forms of $F_e(z)$ for different modulations and the corresponding relations used for computing the ER estimators are listed in Table 3.2. The estimation procedure is summarized in Table 3.3.

Table 3.2: The required CSI and the solutions of nonlinear systems under various modulations.

	$F_e(x)$	Q	$Q^{(v)}$	e_{sd}	e_{sr}	e_{rd}
BPSK	$\frac{(1-2x)^2}{1-(1-2x)^2}$	(3.9.b)	(3.9.b)	(3.9.a)	(3.5.a)	(3.5.a)
BFSK	$\frac{1-2x}{x}$	(3.11.a)	(3.11.b)	(3.11.c)	(3.12.a)	(3.12.b)
DPSK	$\frac{1-2x}{2x}$	(3.11.a)	(3.11.b)	(3.11.c)	(3.12.a)	(3.12.b)

Table 3.3: A blind ER estimation algorithm for BFSK/DPSK modulation.

Input:	Received samples, $\mathbf{y}$, noise variance, σ_d^2, and scaling factor values, $a_{sd}^{(v)}$, $a_{rd}^{(v)}$.
1:	Create virtual SD and RD links by injecting complex Gaussian noise samples with scaling factors, $a_{sd}^{(v)}$ and $a_{r_id}^{(v)}$.
2:	Compute SMPs for all physical, virtual SD-SRD link pairs.
3:	Compute $\widehat{Q}$, $\widehat{Q}^{(v)}$, and $\widehat{e}_{sd}$ by using (3.11.a)-(3.11.c).
4:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ based on (3.12.a) and (3.12.b).
Output:	$\widehat{e}_{sd}$, $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$.

3.4 Simulation Results

For convenience of reference we refer to the ML detector using the ER estimators presented in Section II as the physical-link-only (PLO) detector and that using a VLA estimator as the VLA detector. The ML detector with perfect CSI is called the ideal detector. Let $d_{sr_k}, d_{r_kd}, d_{sd}$ be the distances of the kth SR, RD links and the SD link and θ_{sr_k} be the angle between the SD and kth RD links; see Fig. 1.2. Without loss of generality, we use the normalization, $d_{sd} = 10$ so that

$$d_{sr_k}^2 = d_{r_kd}^2 + d_{sd}^2 - 2d_{r_kd}d_{sd}\cos\theta_{sr_k} = 100 + d_{r_kd}^2 - 20d_{r_kd}\cos\theta_{sr_k}, \tag{3.13}$$

We assume the path loss model, $\sigma_{ij}^2 \propto d_{ij}^{-\alpha}$ with the normalization $\sigma_{sd}^2 = 1$ and $\alpha > 0$. Denote by σ_{ij}^2 the variance of the Rayleigh faded link gain and d_{ij} the distance between node i and node j, $(i,j) \in \{(s,r_k),(r_k,d),\ k = 1,\cdots,L\}$. All the simulated performance curves are obtained by sequentially applying the proposed methods, i.e., the estimated ERs are updated sequentially as each new sample becomes available and the updated

estimates are then used for detecting each received bit. As in [25], we define the SH average SNR as the average received SNR for the direct SD link without relaying, $\bar{\gamma}_{sd}$. Simulation for a given $\bar{\gamma}_{sd}$ terminates whenever the number of error events in the detector output exceeds 500. We assume that noise powers at DN and RN are the same, $\sigma_d^2 = \sigma_r^2$, and use the normalization $P = P_s + \sum_{i=1}^{L} P_{r_i} = 1$ such that $\bar{\gamma}_{sd} = 1/\sigma_d^2$. To reduce the complexity of the ML detector, [24] suggested a piecewise linear function to approximate the nonlinearity (2.3). As it causes negligible performance degradation with respect to that of the ML detector so long as $e_{sr_k} < \frac{1}{2}$, we use the same approximation in our simulation efforts. Fig. 3.2 illustrates the block diagram of the maximum likelihood (ML) detector, where $f_T(t)$ is approximated by a piecewise linear function: $f_T(t; e_{sr_k}) \approx \min(\max(t, -T), T)$ and $T = \ln\left(\frac{1-e_{sr_k}}{e_{sr_k}}\right)$.

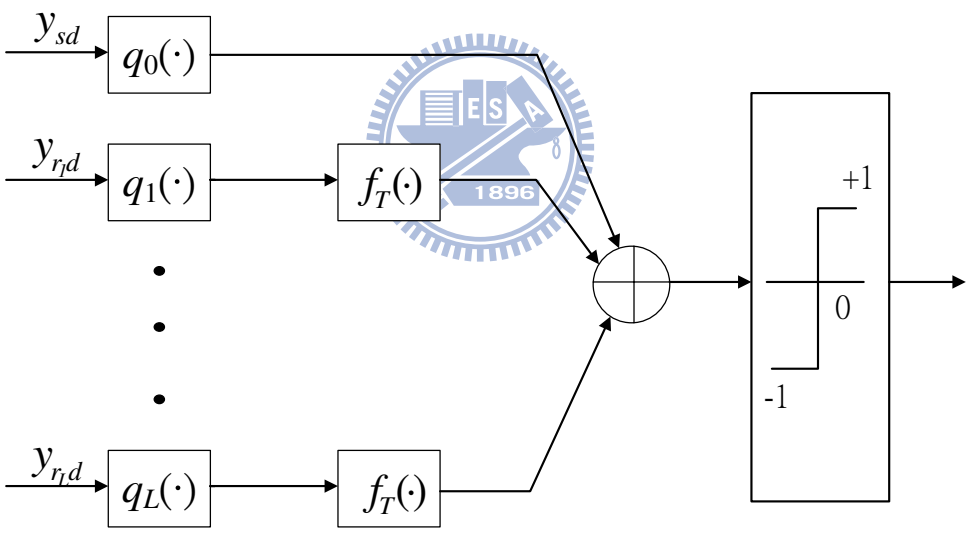


Figure 3.2: Block diagram of the maximum likelihood (ML) detector.

The performance of the PLO and VLA detectors for the simplest case, $L = 1$ with BPSK modulation, is illustrated in Fig. 3.3. For the PLO detector, only e_{sr} is unknown while the VLA detector assumes ERs of other component links are also unavailable and uses a rotation angle $\theta = 45°$, which is equivalent to injecting noise with $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$. The performance of both detectors are found to approach that of the ideal ML detector.

We also investigate the effect of correlated fading on the performance of the VLA detector for DPSK signals and the result is shown in the same figure. Modified Jake's model [38] with normalized Doppler frequency $J = f_d T_s = 0.001$, where f_d and T_s being the Doppler frequency and the sampling period, respectively, is used to generate the component link gains, $\{h_{sd}[n]\}$, $\{h_{sr}[n]\}$, and $\{h_{rd}[n]\}$, as a function of sampling epochs. For the DPSK system, we use the noise-injected VLA detector with scaling factors $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$; see (3.10). Obviously, the performance of the VLA detector is almost the same as that of the ML detector within the range of interest, indicating that the i.i.d. assumption gives accurate ER estimates for moderately correlated fading environments.

Fig. 3.3 also show the performance for the cases of two and four RNs. In the two-relay case, we assume that the PLO detector knows $e_{r_k d}$ perfectly. Again, both PLO and VLA detectors yield performance almost identical to that of the ML detector. For the four-relay case, we decompose the problem into four single-relay CCN subproblems, each involves only one SRD and the SD links. It can be seen that at the low SH-SNR region ($0 \sim 2$ dB), the performance of the VLA detector is slightly worse than that of the optimal detector. This is due to fact that the sample size used is not large enough to offer a very reliable BER estimate. Nevertheless its performance is still superior to that of the MRC detector.

In the communication systems, the transmission are packet-based. To investigate the effect of finite length packet, we first estimate the ERs given two fixed finite samples and simulate the bit error rate performance given the estimation results. If the ERs are small, the ER estimates are quite bad given insufficient samples. Hence, it can be expected that the performance will be degenerated, especially at high SNR region, as shown in Fig. 3.4. Moreover, increasing the sample size can also improve the bit error rate performance due to the better ER estimates.

Finally, we consider the effect of ER estimate error in bit error rate performance. Fig. 3.6 shows that the distribution of y_{sd} has a peak near 0 while y_{rd} is more smooth.

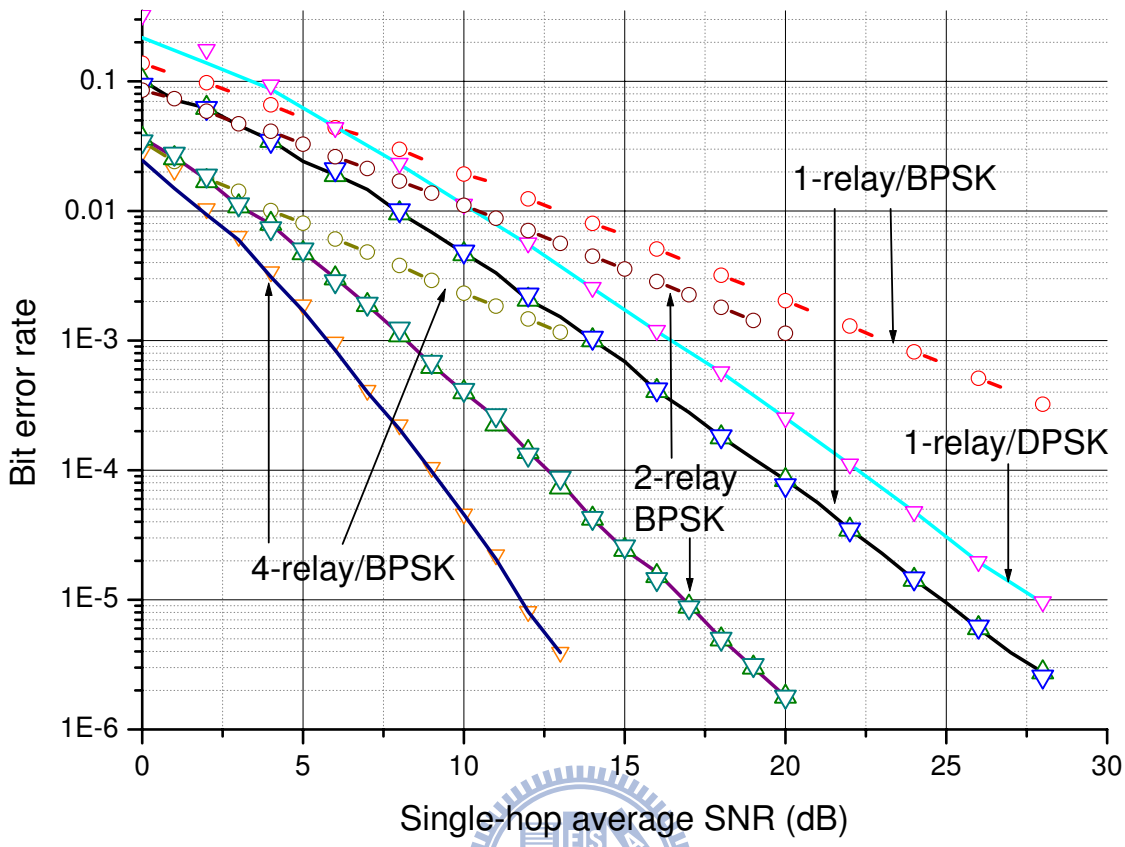


Figure 3.3: Bit error rate performance of the ML (solid curves), MRC (○), PLO (△) and VLA (∇) detectors. The following system parameter values are used. (i) single-relay system: $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.8$, $\theta_{sr} = 0°$, and $a^v_{sd} = a^v_{rd} = 2$ (single-relay), (ii) 2-relay system: $d_{r_1d} = d_{sr_2} = 7/10$, $\theta_{sr_1} = \theta_{sr_2} = 0°$, $P_s = 0.5P, P_{r_1} = P_{r_2} = 0.25P$, and $\theta = 45°$, (iii) 4-relay system: $d_{sd} = 10, d_{sr_1} = 5, \theta_{sr_1} = 45°, d_{sr_2} = 6, \theta_{sr_2} = 30°, d_{sr_3} = 4, \theta_{sr_3} = 60°, d_{sr_4} = 5, \theta_{sr_1} = 0°, P_p = 0.5P, P_{r_i} = 0.125P$, for $i = 1, 2, 3, 4$ and $\theta = 30°$.

If y_{sr} is negative and near 0, then high value of y_{rd} will be underestimated when $\widehat{e}_{sr}$ is overestimated (see Fig. 3.5). In this case, the probability that $y_{sr} + f_T(y_{rd}) < 0$ increases, yielding more errors. On the other hand, if y_{sr} is positive and near 0, then high value of y_{rd} will be overestimated when $\widehat{e}_{sr}$ is underestimated as the case of error-free. Hence, the error event also occurs more frequently. Nevertheless, if the ER estimate error is not large, then there is not a great difference between the nonlinear function with ER estimate and that with true ER, as shown in Fig. 3.5. Consequently, the performance is not degenerated significantly.

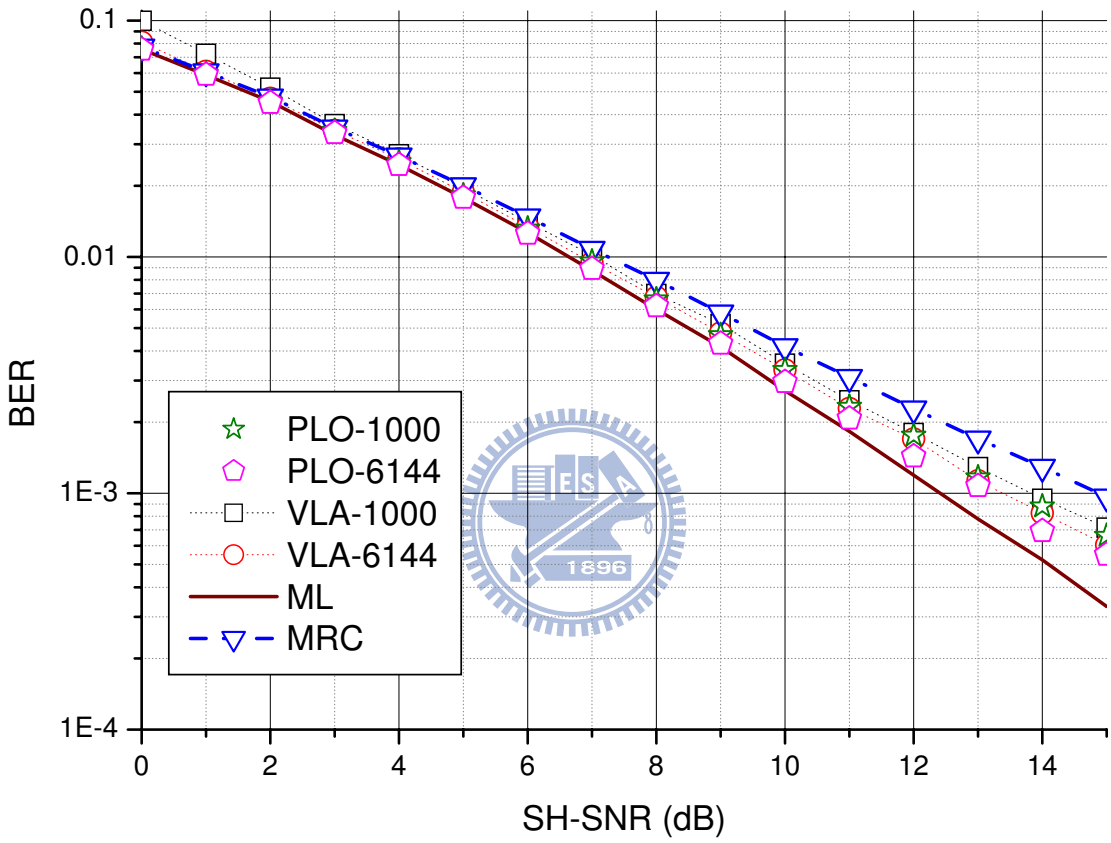


Figure 3.4: Bit error rate performance of the ML, MRC, PLO, and VLA detectors. The following system parameter values are used. Single-relay system: $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.$, $\theta_{sr} = 45°$, and $a^v_{sd} = a^v_{rd} = 2$.

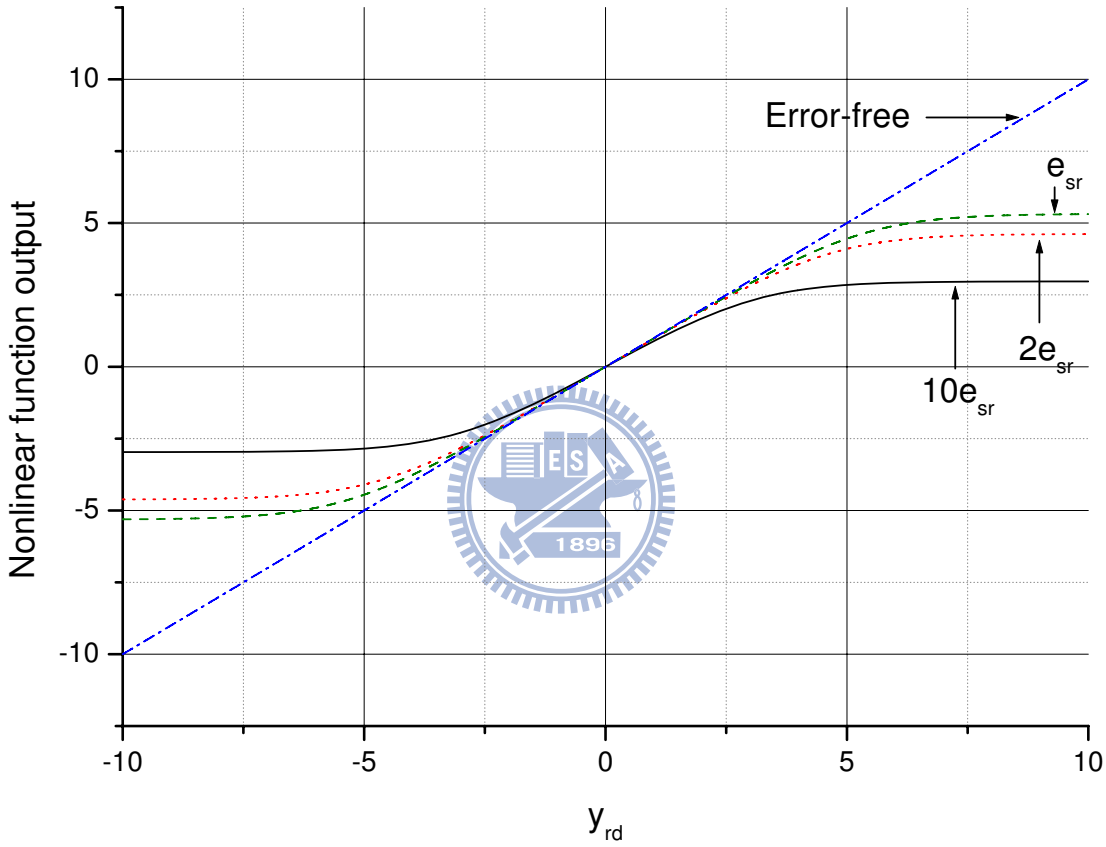


Figure 3.5: The nonlinear function (2.3) under various ERs. The system parameter values are $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.$, $\theta_{sr} = 45°$ and SNR= 8 dB ($e_{sr} = 0.0049$).

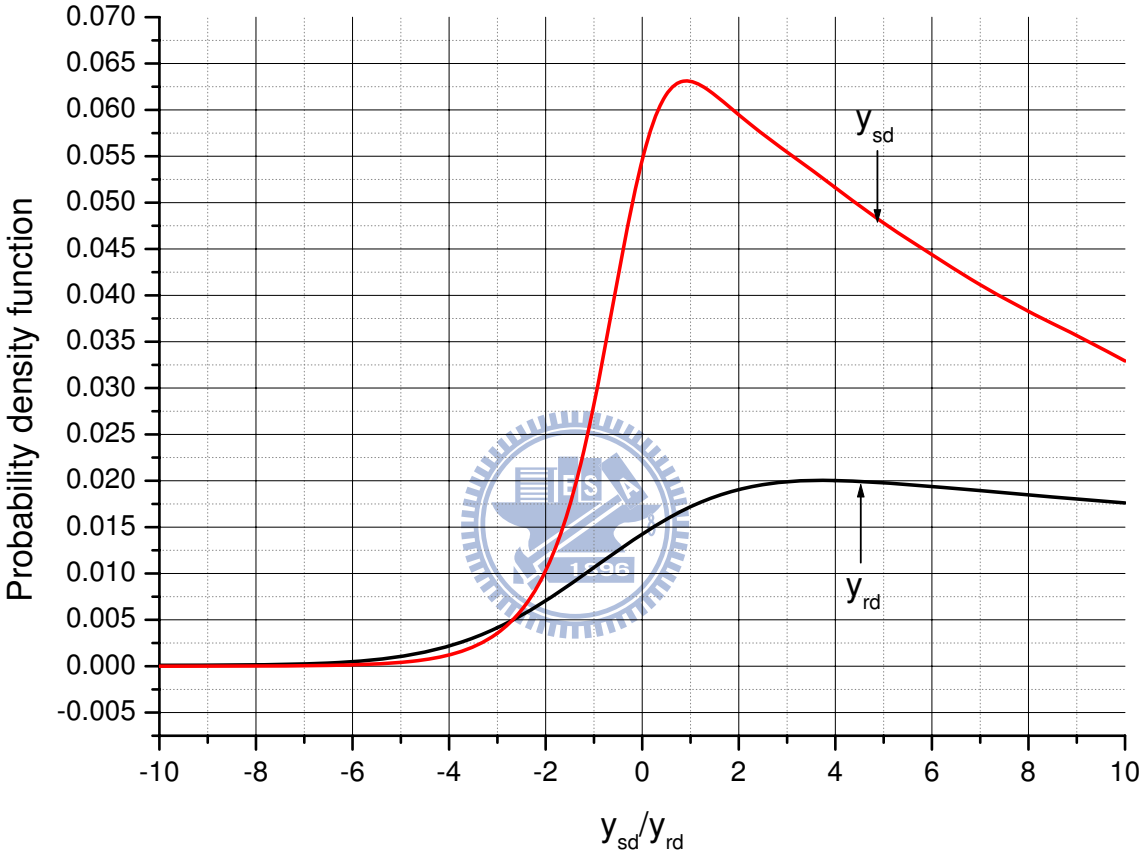


Figure 3.6: The probability density functions of y_{sd} and y_{rd} given $x = 1$ in single-relay system. The system parameter values are $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.$, $\theta_{sr} = 45°$ and SNR= 8 dB.

Chapter 4

Noise-Enhanced ER Estimations

4.1 Convergence consideration and a simple variance reduction method

It is easy to see that, like the estimator for the SMP p defined in Chapter 2, $\widehat{p}_{sr}$, $\widehat{p}_{(vs)r}$, $\widehat{p}_{s(vr)}$, and $\widehat{p}_{(vs)(vr)}$ converge in probability. As the proposed estimators are continuous functions of these estimates, the continuous mapping theorem [39] implies that the estimators $\{\widehat{e}_{sr}, \widehat{e}_{rd}, \widehat{e}_{sd}\}$ converge in probability as well and their variances depend on those of the SMP estimators. The latter are all derived from the same compare-and-count process which is similar to that used in simulation-based ER estimations [40]. The main difference is that, for the latter, the desired detector output is known perfectly and one has complete information and control of the operating average SNR and the link output statistic. In contrast, our scheme can only rely on blind counting without a pilot sequence and the link statistic is either unavailable or only partially known. Both estimation methods, however, have the same order of convergence rate and require a large number of samples to obtain a reliable estimate if the true ER is small; see *Lemma 2.1* and [40].

A straightforward approach to improve the convergence performance is to use multiple VLs, i.e., we add $n_{vl} - 1$ virtual RD and/or SD links with the same noise power. Each VL renders a set of new estimates and the final estimates are obtained by taking average of the n_{vl} estimates. This method is called the *enhanced-VLA* (EVLA) estima-

tor which yields a reduced variance for a given sample size, or equivalently, achieves the same variance as that of the original ($n_{vl} = 1$) estimator with a smaller sample size.

4.2 A brief introduction to importance sampling

To further improve the convergence/variance performance, the above-mentioned analogy between our method and the simulation-based estimator suggests that we apply a variance reduction method used in the latter approach called importance sampling (IS). For self-contained, we briefly introduce the IS technique.

As we known, variance is a key measure to evaluate the efficiency of a estimator. Low variance implies the low number of samples required for a fixed precision. Hence, several techniques had been proposed to reduce the variance significantly [41] and importance sampling is one of them [40], [42].

To understand importance sampling technique, we consider the following problem:

$$p = \int H(x) f(x) dx \tag{4.1}$$

where H is a system performance function and f is a probability density function. A concrete example is the estimation of error rate for BPSK modulation in AWGN channel. In this case, $f(x)$ is the distribution of received signal and $H(x)$ is an indicator function of an error event. To estimate the value p, we can generate several samples x_i according to the distribution $f(x)$, evaluate the system performance function, and calculate the average weighted sum:

$$\widehat{p} = \frac{1}{N} \sum_{i=1}^{N} H(x_i) \tag{4.2}$$

For importance sampling technique, we should choice another probability density function $g(x)$ such that $g(x) \neq 0$ whenever $H(x)f(x) \neq 0$. Then, the estimation problem (4.1) becomes

$$p = \int H(x) \frac{f(x)}{g(x)} g(x) dx \tag{4.3}$$

and we can estimate p by

$$\widehat{p} = \frac{1}{N}\sum_{i=1}^{N} H(x_i)\frac{f(x_i)}{g(x_i)} \tag{4.4}$$

which is called the importance sampling estimator. Notice that the samples x_i are drawn from the distribution $g(x)$ instead of $f(x)$ in this case. The probability density function $g(x)$ is usually called the importance sampling density, proposal density, or instrumental density [43]. In addition, the ratio of densities,

$$W(x) = \frac{f(x)}{g(x)} \tag{4.5}$$

is called the likelihood ratio.

The importance sampling estimator is unbiased because

$$E\left[\widehat{p}\right] = \frac{1}{N}\sum_{i=1}^{N}\int_x H(x_i)\frac{f(x_i)}{g(x_i)}g(x_i)dx_i = \frac{1}{N}\sum_{i=1}^{N}\int_x H(x_i)f(x_i)dx_i = \frac{N}{N}p = p$$

Moreover, the variance of $\widehat{p}$ is

$$Var\left[\widehat{p}\right] = \frac{1}{N}\sum_{i=1}^{N} Var\left[H(x_i)\frac{f(x_i)}{g(x_i)}\right]$$

To find the optimal importance sampling density $g(x)$, we can minimize the variance of $\widehat{p}$, which equivalent to minimize the variance $Var\left[H(x_i)\frac{f(x_i)}{g(x_i)}\right]$. Notice that

$$\begin{aligned}\left(\int |H(x)|f(x)dx\right)^2 &= \left(\int \frac{|H(x)|f(x)}{[g(x)]^{1/2}}[g(x)]^{1/2}dx\right)^2 \\ &\leq \int \frac{|H(x)|^2 f(x)^2}{g(x)}dx \int g(x)dx = \int \frac{|H(x)|^2 f(x)^2}{g(x)}dx\end{aligned}$$

Hence, the variance $Var\left[H(x_i)\frac{f(x_i)}{g(x_i)}\right]$ is minimized if $g_{opt}(x_i) \propto |H(x)|f(x)$. Since $g(x)$ is a probability density function, we have

$$g_{opt}(x_i) = \frac{|H(x)|f(x)}{\int |H(x)|f(x)dx} \tag{4.6}$$

Moreover, if $H(x) \geq 0$, we have $g_{opt} = H(x)f(x)/p$ and

$$Var\left[H(x)\frac{f(x)}{g_{opt}(x)}\right] = E\left[H^2(x)\frac{f(x)^2}{g_{opt}(x)^2}\right] - p^2 = E\left[p^2\right] - p^2 = 0 \tag{4.7}$$

As can be seen, if we can find the optimal importance sampling density, the variance becomes zero if $H(x) \geq 0$. However, it also implies that we should known the value p, which is the value we want to estimate and could not be known in advance. In general, how to select a good importance sampling density depends on the problem we encounter and a bad choice of importance sampling density may even yield large variance.

To sum up, the basic idea of IS method is to modifies the distribution of received sample so that it follows a proposal probability distribution that makes the important (error) event occurs much more often than the original unmodified case does.

In [42], the IS technique is applied to the estimation of error probabilities over non-linear channel. A variance reduction factor for a given sample size is defined as

$$\gamma \approx \frac{MSEE_{IS}}{MSEE_o} \tag{4.8}$$

where $MSEE_o$ and $MSEE_{IS}$ are the MSEE of the direct Monte-Carlo and importance sampling estimations, respectively. The smaller the variance reduction factor is, the better the importance sampling estimation is. Because $MSEE$ is inverse proportional to the sample size N, the variance reduction factor also indicates the sample size reduction inversely. For example, if $\gamma = 0.01$ and the direct Monte-Carlo estimation needs 10^6 samples for some $MSEE$, then the importance sampling estimator only needs 10^4 samples with the same performace.

Moreover, a importance sampling density $g(x) = \sqrt{\frac{1-\alpha}{(2\pi)^\alpha}}[f(x)]^{1-\alpha}$ is proposed in [42] and it is shown that the optimal value of α is

$$\alpha_{opt} = \frac{-3 + \sqrt{9 + 4T^2(1 + T^2)}}{2T^2} \tag{4.9}$$

for memoryless Gaussian channel with $H(x) = I(x > T)$ and $T > 0$ is an arbitrary threshold. Similar to the case, we will propose an importance sampling inspired enhanced estimator with a parameter to change the distribution of the received signals and find the optimal value of this parameter.

4.3 An importance sampling inspired noise-enhanced estimator

The difficulty in applying the IS theory to our scenario, besides the fundamental differences just mentioned in section 4.1, is due to the fact that the estimators, as was shown in (3.11.a)–(3.12.b) and other similar equations presented before, are derived from SMPs and, perhaps, other ERs. The complete control of their statistics through dependent variables whose probability distributions are unknown is impossible. For instance, in the case of a BPSK based CCN, an SMP depends on the inner product of the SD and RD link outputs whose probability distributions depend on, among other parameters, the true ER of the SR link, which needs to be estimated in the first place. In other words, the optimal (variance-minimizing) importance distribution is a function of the parameters whose values we either do not know or want to estimate.

The following observations, however, indicate that a *suboptimal* importance distribution is obtainable. Firstly, the ultimate parameters of interest are the link ERs not the pairwise SMPs and the IS theory says that convergence is faster if the ER to be estimated by simulation is properly increased, which may be realized by simply adjusting the corresponding link output's variance. Secondly, some ER estimator formulae are functions of other ERs and SMPs, hence if the estimates of other ERs can be improved while those for SMPs remain unchanged, e.g., the ER estimator of e_{sr} through (2.8), we can obtain an improved estimator. Finally, it is reasonable to assume that link outputs' statistics are partially known, e.g., their noise variances. But even if we are able to partially control the distributions of related parameters, there still exist the problem of weighting the resulting counts, which is needed in a conventional IS-based procedure and can only be done if both the original and modified link output distributions are known.

Our solution which overcomes all these difficulties can be expressed by the following toy example. Similar to the simulation case, we consider a point-to-point communica-

tion system with BPSK modulation in Rayleigh fading channel and assume that the transmission signal is perfectly known at receiver. That is, we encounter a data-aided error rate estimation problem.

To estimate the error rate, a simple but widely used estimator is compare-and-count estimator:

$$\widehat{e} = \frac{\sum_{i=1}^{N} I(\widehat{y}_i = x_i)}{N} \tag{4.10}$$

where x_i and $\widehat{y}_i$ are the ith transmitted signal and its decision at receiver. It has been shown that the normalized MSEE of this estimator is

$$MSEE_o = \frac{1-e}{Ne} \tag{4.11}$$

where e is the system error rate. This formula indicates that it requires more samples to estimate a smaller error rate. In other words, given a fixed normalized MSEE, the required number of samples decreases with the increase of error rate. It motives us to transform the estimation problem of error rate e into another one with larger error rate, say $e^{(w)}$, which is similar to the concept of IS technique. To increase the system error rate can be done by injecting some noise into the received signals and the distribution of received signals changes. Recall that when we use IS technique, we need to calculate the likelihood ratio:

$$W(y) = \frac{f(y)}{g(y)} \tag{4.12}$$

In this toy example, $f(y)$ and $g(y)$ are

$$f(y) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(y_i - h_i\sqrt{P}x_i)}{N_0}\right) \tag{4.13}$$

$$g(y) = \frac{1}{\sqrt{\pi a^w N_0}} \exp\left(-\frac{(y_i - h_i\sqrt{P}x_i)}{a^{(w)}N_0}\right) \tag{4.14}$$

where P is the transmission power and we inject Gaussian noise such that the noise variance is $a^{(w)}N_0$, $a^{(w)} \geq 1$. It is not reasonable to assume that P is also known

because the error rate can be calculated by the formula (3.1) [37]. In terms of P and N_0, (3.1) can be rewritten as

$$e = \frac{1}{2}\left(1 - \sqrt{\frac{P}{P + N_0}}\right) \tag{4.15}$$

if the variance of channel gain h is 1. Hence, it is difficult to evaluate the likelihood ratio.

To overcome this difficulty, we observe that we have the error rate formula (3.1) and inverse formula (3.2). We can first estimate the error rate $e^{(w)}$ and find the equivalent $SNR^{(w)}$ by (3.2). Since $a^{(w)}$ is known, we can transform $SNR^{(w)}$ to $SNR = a^{(w)}SNR^{(w)}$. Finally, use the formula (3.1) to obtain the estimate of e. The flowchart and noise-enhanced ER estimation algorithm is given in Fig. 4.1 and Table 4.1, respectively. Notice that we can combine the transformations by the following formula (Appendix B.1):

$$e = \frac{1}{2}\left(1 - \sqrt{\frac{a^{(w)}(1 - 2e^{(w)})^2}{1 - (1 - 2e^{((w))})^2 + a^{(w)}(1 - 2e^{(w)})^2}}\right) \tag{4.16}$$

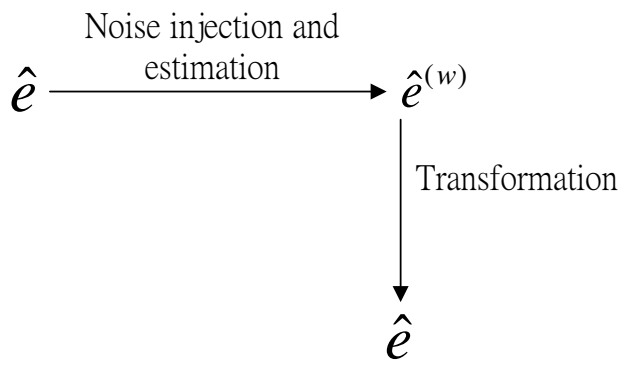


Figure 4.1: The flowchart of the proposed noise-enhanced estimation for data-aided BPSK system in Rayleigh fading channel

According to the toy example, the proposed noise-enhanced estimator for CCN proceeds as follows. We first add zero-mean complex Gaussian samples with variances,

Table 4.1: A noise-enhanced ER estimation algorithm for data-aided BPSK system in Rayleigh fading channel.

Input:	Received samples, $\mathbf{y}$, noise variance, σ_d^2, and $a^{(w)}$.
1:	Add noise-enhanced complex zero-mean Gaussian samples with variances $a^{(w)} - 1$ to $\mathbf{y}$.
2:	Obtain $\widehat{e}^{(w)}$ by (4.10).
3:	Find out $\widehat{e}$ by .
7:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ via (3.5.a)-(3.6).
8:	**elseif** modulation type is BFSK or DPSK **then**
9:	Compute $\widehat{Q}$, $\widehat{Q}^{(v)}$, and $\widehat{e}_{sd}$ by using (3.11.a)-(3.11.c).
10:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ based on (3.12.a) and (3.12.b).
11:	**end if**
12:	**end for**
13:	Convert the estimates of the noise-injected ERs back to those of the original (uncontaminated) ERs via (4.17) or (4.18).
14	Take averages of all estimates, if available, to obtain the final estimates $\widehat{e}_{sd}$ and $\widehat{e}_{rd}$.
15	Compute $\widehat{e}_{sr}$ via (2.10) using the estimates obtained in 14 and $\widehat{p}$ derived from the uncontaminated received samples.
Output:	$\widehat{e}_{sd}$, $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$.

N_{sd}, N_{rd}, to the received SD and RD link output samples y_{sd} and y_{rd}, respectively. This results in link outputs with larger variances. By solving the nonlinear system associated with the estimated SMPs of the noise-injected links, we obtain the estimates $\left\{\widetilde{e}_{sd}^{(w)}, \widetilde{e}_{sr}^{(w)}, \widetilde{e}_{rd}^{(w)}\right\}$, where the superscript (w) is used to signify the fact that the estimates are computed by inserting artificial noises. As the noise-injection effectively reduces the average SNR, the scaling relation (3.10) with $a^{(v)} = a^{(w)} = 1 + N_{sd}/\sigma_d^2$ or $1 + N_{rd}/\sigma_d^2$ enables us to weight and convert the estimates, $\left\{\widetilde{e}_{sd}^{(w)}, \widetilde{e}_{rd}^{(w)}\right\}$, back to the estimates $\{\widehat{e}_{sd}^{(w)}, \widehat{e}_{rd}^{(w)}\}$ of the true ERs $\{e_{sd}, e_{rd}\}$. For instance, in a noncoherent BFSK or DPSK based CCN, the relation, $\frac{1-2e}{e} = a^{(w)}\frac{1-2e^{(w)}}{e^{(w)}}$, for $\max\{e, e^{(w)}\} < 1/2$, suggests that DN use the conversion rule

$$\widehat{e}^{(w)} = \widetilde{e}^{(w)}/(a^{(w)} + 2\widetilde{e}^{(w)} - 2a^{(w)}\widetilde{e}^{(w)}), \tag{4.17}$$

where the subscripts, "sd" and "rd" associated with the estimators $\widetilde{e}^{(w)}$ and $\widetilde{e}^{(w)}$ are

omitted to simplify the expression. Similarly, the conversion rule for a BPSK based network is

$$\widehat{e}^{(w)} = \frac{1}{2}\left[1 - \sqrt{\frac{a^{(w)}\left(1-2\widetilde{e}^{(w)}\right)^2}{1-\left(1-2\widetilde{e}^{(w)}\right)^2 + a^{(w)}\left(1-2\widetilde{e}^{(w)}\right)^2}}\right]. \tag{4.18}$$

The above two conversion rules bypass the need for complete statistics by directly using the ER conversion based only on, $a^{(w)}$, the ratio between the noise-injected and original SNRs (instead of individual SNRs). They also imply that $\widehat{e}^{(w)} < \widetilde{e}^{(w)}$, which has been expected as we have purposely made $e^{(w)}$ larger by injecting noise. If VLs are needed, we have to inject an additional noise term into the noise-injected PLs to create VLs. Hence, the scaling factor is $a^{(v)}$, $a^{(w)}$, or $a^{(v)}a^{(w)}$, depending on whether the link is a VL, noise-injected PL or a noise-injected VL. We call the class of estimators based on the above concept as the *importance sampling inspired* VLA (ISI-VLA) estimator. In the following sections, we show, via both analysis and simulations, that the ISI-VLA estimator does offer significant performance enhancement. The novel estimation procedure in this and previous section are summarized in Table. 4.2.

4.4 Properties and performance analysis of the noise-enhanced estimator

For the above approach, noise-injection is performed to improve the ER estimators not the SMP p observed at the DN. In fact, it results in a smaller SMP $p^{(w)}$ and if we want to estimate the original p through $p^{(w)}$, we obtain a worse SMP estimate, i.e.,

Lemma 4.1. Let p and $p^{(w)}$ be the true SMP's of the original and noise-injected links, $\widehat{p}^{(w)}$ and $\widehat{p}$ be the estimates of p with and without the aid of the noise-injected link. Then

$$\mathrm{Var}\left[\widehat{p}\right] \leq \mathrm{Var}\left[\widehat{p}^{(w)}\right]. \tag{4.19}$$

Proof. See Appendix B.2. □

Table 4.2: A unified blind noise-enhanced ER estimation algorithm.

Input:	Received samples, $\mathbf{y}$, noise variance, σ_d^2, scaling factor values, $a_{sd}^{(v)}$, $a_{rd}^{(v)}$, $a_{sd}^{(w)}$, $a_{rd}^{(w)}$ and the number of VLs, n_{vl}.
1:	**for** i=1 to n_{vl} **do**
2:	Add noise-enhanced complex zero-mean Gaussian samples with variances $a_{sd}^{(w)} - 1$ and/or $a_{r_i d}^{(w)} - 1$ to the received SD- and RD-link output samples.
3:	Create virtual SD and RD links by injecting complex Gaussian noise samples with scaling factors, $a_{sd}^{(v)}$ and $a_{r_i d}^{(v)}$.
4:	Compute SMPs for all physical, virtual and/or noise-enhanced SD-SRD link pairs.
5:	**if** modulation type is BPSK **then**
6:	Compute $\widehat{Q}$, $\widehat{Q}^{(v)}$, and $\widehat{e}_{sd}$ through (3.9.b) and (3.9.a) with $a_{sd}^{(v)} = a_{rd}^{(v)} = \frac{1}{\cos^2\theta}$.
7:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ via (3.5.a)-(3.6).
8:	**elseif** modulation type is BFSK or DPSK **then**
9:	Compute $\widehat{Q}$, $\widehat{Q}^{(v)}$, and $\widehat{e}_{sd}$ by using (3.11.a)-(3.11.c).
10:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ based on (3.12.a) and (3.12.b).
11:	**end if**
12:	**end for**
13:	Convert the estimates of the noise-injected ERs back to those of the original (uncontaminated) ERs via (4.17) or (4.18).
14	Take averages of all estimates, if available, to obtain the final estimates $\widehat{e}_{sd}$ and $\widehat{e}_{rd}$.
15	Compute $\widehat{e}_{sr}$ via (2.10) using the estimates obtained in 14 and $\widehat{p}$ derived from the uncontaminated received samples.
Output:	$\widehat{e}_{sd}$, $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$.

As we can only inject noise into samples received by the DN, y_{sd} and (or) y_{rd}, e_{sr} remains intact and $\widehat{e}_{sr}^{(w)} = \widetilde{e}_{sr}^{(w)}$ if this estimator is obtained by substituting $\widehat{e}_{rd}^{(w)}, \widehat{e}_{sd}^{(w)}$ and $\widehat{p}^{(w)}$ into (2.8). The above *lemma* suggests that we should replace $\widehat{p}^{(w)}$ by $\widehat{p}$ in the substitution procedure for estimating e_{sr}. As mentioned in the last section, a better estimate for e_{sr} can thus be obtained by using the noise-enhanced estimates, $\widehat{e}_{rd}^{(w)}, \widehat{e}_{sd}^{(w)}$, and the original $\widehat{p}$; see (2.8).

The range of the appropriate values for the scaling factor $a^{(w)}$ is certainly dependent on the true ERs e and the noise injected ERs $e^{(w)}$. As will be shown in *Theorem 4.4* and

numerically in next section that the MSEE performance is improved by injecting proper noise power into the received samples and there is an optimal injected noise power that achieves the maximum MSEE improvement. This phenomena is called the stochastic resonance effect which has been observed in some nonlinear signal processing systems.

In a BPSK-based single-relay CCN with perfect SD link ($e_{sd} = 0$), when both the average transmitted relay power P_r and the magnitude of the slow-faded RD link gain $|h_{rd}|$ are known, we show in Appendix B.3 that the optimal scaling factor is approximately equal to the RD link output SNR

$$a_{opt}^{(w)} \approx \frac{P_r |h_{rd}|^2}{\sigma_d^2}. \tag{4.20}$$

In addition, we also discuss the performance in the data-aided point-to-point BPSK communication and the results are given in the following Theorem

Theorem 4.2. For a data-aided point-to-point BPSK communication, the optimal scaling factor $a^{(w)}$ is equal to

$$a_{opt}^{(w)} = \text{SNR}, \tag{4.21}$$

The minimum achievable MSEE reduction ratio, γ_{min}, is given by

$$\gamma_{min} = \frac{4\text{SNR}}{(1 + \text{SNR})^2} \tag{4.22}$$

Moreover, noise injection using the optimal scaling factor is beneficial if SNR is larger than 1 (0 dB).

Proof. See Appendix B.4 □

For blind estimation in CCN, we need the following *lemma* to derive a closed-form expression of the optimal scaling factor for the more practical case addressed in *Theorem 4.4*.

Lemma 4.3. [Noncoherent modulation] For a network that consists of three independent (SD or cascaded) flat Rayleigh fading links with ERs e_i. If the ISI scheme is applied

with a common noise-injected ER $e_i^{(w)} = \epsilon$ using the scaling factors $a_i^{(w)}, i = 1, 2, 3$, the variance of the noise enhanced estimator $\widehat{e}_i$ using the conversion rule (4.17) is given by

$$\mathrm{Var}\left[\widehat{e}_i\right] \approx \frac{\left(a_i^{(w)}\right)^2}{\left(a_i^{(w)} + 2\epsilon - 2a_i^{(w)}\epsilon\right)^4} \frac{\epsilon - 2\epsilon^2 + 2\epsilon^3 - \epsilon^4}{(2\epsilon - 1)^2 N}. \tag{4.23}$$

Proof. See Appendix B.5. □

In subsequent discourse, we denote by $\widehat{y}_i, e_i$, SNR_i, the hard-decision output, ER, average SNR of the ith link (direct or cascaded) and by $a_i^{(v)}$ and $a_i^{(w)}$, the associated scaling factor used. To characterize the stochastic resonance effect and the noise enhanced performance we define the MSEE reduction ratio, $\gamma \stackrel{def}{=} \mathrm{MSEE}_{\mathrm{ISI}}/\mathrm{MSEE}_{\mathrm{O}}$, where $\mathrm{MSEE}_{\mathrm{ISI}}$ and $\mathrm{MSEE}_{\mathrm{O}}$ are the MSEE's of the ISI-VLA and VLA estimators with the same sample size. Using the above *lemma* we obtain

Theorem 4.4. For a network with three independent flat Rayleigh fading links, the optimal scaling factor under the common noise-injected ER constraint, $e_i^{(w)} = \epsilon, i = 1, 2, 3$, is approximately equal to

$$a_{i,opt}^{(w)} \approx t_1 \mathrm{SNR}_i, \tag{4.24}$$

where $t_1 = 0.3085$ (DPSK) or 0.15428 (BFSK). The minimum achievable MSEE reduction ratio, γ_{min}, for $\mathrm{SNR}_i \gg 1$ is given by

$$\gamma_{min} \approx \begin{cases} 9.8277 \frac{\mathrm{SNR}_i^2}{(1+\mathrm{SNR}_i)^3}, & \text{DPSK} \\ 19.655 \frac{\mathrm{SNR}_i^2}{(2+\mathrm{SNR}_i)^3}, & \text{BFSK} \end{cases} \tag{4.25}$$

Moreover, noise injection using the optimal scaling factor is beneficial if SNR_i is larger than 3.241 (DPSK) or 6.483 (BFSK).

Proof. See Appendix B.6. □

Following a procedure similar to that used in proving *Lemma 4.3* and *Theorem 4.4* and using the relation governing the ER ϵ of a noise-injected BPSK link and the associated scaling factor, $\epsilon = \frac{1}{2}\left(1 - \sqrt{\frac{\mathrm{SNR}_i}{a_i^{(w)} + \mathrm{SNR}_i}}\right)$, we can prove that

Theorem 4.5. For a 3-link BPSK based network in a flat Rayleigh fading environment, the optimal scaling factors that ensure a common noise-injected ER is $a_i = t_i \mathrm{SNR}_i$ and the MSEE reduction ratio γ for link i is

$$\gamma|_{a_i=t_1\mathrm{SNR}_i} = \frac{24.68\mathrm{SNR}_i^2}{3 + 10\mathrm{SNR}_i + 11\mathrm{SNR}_i^2 + 4\mathrm{SNR}_i^3}, \tag{4.26}$$

where $t_1 = \frac{-1+\sqrt{7}}{3}$. Noise injection using the optimal scaling factor is beneficial if $\mathrm{SNR}_i > 1.823$.

To evaluate the MSEE performance of VLA and ISI-VLA estimators in a CCN, as shown in B.5, we need to compute the covariance and matrix $\mathbf{C}$ of the pairwise matching indicators $I(\hat{y}_k[t] = \hat{y}_j[t])$ and the associated Jacobian matrix $\mathbf{J}$. The entries of these two matrices are functions of the (not necessarily pairwise) SMPs whose expressions are given below[1].

Lemma 4.6. For a two-link BPSK-based network, the SMP, $p_{12(v1)}$, that direct PLs 1, 2 and $a_1^{(v)}$-scaled VL 1 (denoted by $v1$) all yield the same hard decision is given by

$$p_{12(v1)} = e_2 p_{em}(e_1, a_1^{(v)}) + (1 - e_2) p_{cm}(e_1, a_1^{(v)}) \tag{4.27}$$

where the *conditional erroneous matching probability* $p_{em}(e_1, a_1^{(v)}) \stackrel{def}{=} \Pr(\widehat{y}_1 = \widehat{y}_1^{(v)} = -s|s)$ and the *conditional correct matching probability*, $p_{cm}(e_1, a_1^{(v)}) \stackrel{def}{=} \Pr(\widehat{y}_1 = \widehat{y}_1^{(v)} = s|s)$, $s = \pm 1$ being the normalized transmitted BPSK signal, are

$$p_{em}(e_1, a_1^{(v)}) = \frac{e_1}{2} + \frac{1}{2\pi}\left[\tan^{-1}\left(\frac{1}{\sqrt{a_1^{(v)} - 1}}\right) - \frac{\tan^{-1}\left(\frac{(1-2e_1^{(v)})^{-1}}{\sqrt{a_1^{(v)}-1}}\right)}{(1 - 2e_1^{(v)})^{-1}}\right] \tag{4.28}$$

$$p_{cm}(e_1, a_1^{(v)}) = 1 - e_1 - e_1^{(v)} + p_{em}(e_1, a_1^{(v)}) \tag{4.29}$$

[1]Matching probabilities and variance analysis for noise-enhanced estimators are similar. Depending on where the noise-injected links are located, the resulting expressions are obtained by replacing e_i and (or) $e_i^{(v)}$ by $e_i^{(w)}$ and (or) its VL version; the scaling factors are also modified when necessary. This apply to *Lemmas 4.6, 4.8* as well as *Theorems 4.7, 4.9*.

If PL 1 is a cascaded link, the SMP becomes

$$\begin{aligned} p_{12(v1)} =& e_2 \left[p_{em}(e_1, a_1^{(v)})(1-e_{sr}) + p_{cm}(e_1, a_1^{(v)}) e_{sr} \right] \\ &+ (1-e_2) \left[p_{em}(e_1, a_1^{(v)}) e_{sr} + p_{cm}(e_1, a_1^{(v)})(1-e_{sr}) \right] \end{aligned} \tag{4.30}$$

where e_{sr} is ER of the hidden component link of PL 1. The SMP that direct PL 1, cascaded PL 2, $a_i^{(v)}$-scaled VLs 1 and 2 all yield the same hard decision is given by

$$\begin{aligned} p_{12(v1)(v2)} =& p_{em}(e_1, a_1^{(v)}) \left[p_{em}(e_2, a_2^{(v)})(1-e_{sr}) + p_{cm}(e_2, a_2^{(v)}) e_{sr} \right] \\ &+ p_{cm}(e_1, a_1^{(v)}) \left[p_{em}(e_2, a_2^{(v)}) e_{sr} + p_{cm}(e_2, a_2^{(v)})(1-e_{sr}) \right] \end{aligned} \tag{4.31}$$

Finally, we have the two joint pairwise SMPs

$$\begin{aligned} &\Pr\left(\widehat{y}_1 = \widehat{y}_2,\ \widehat{y}_1^{(v)} = \widehat{y}_2^{(v)}\right) = p_{12(v1)(v2)} \\ &+ \left[e_1 - p_{em}(e_1, a_1^{(v)}) \right] \left\{ (1-e_{sr}) \left[e_2 - p_{em}(e_2, a_2^{(v)}) \right] + e_{sr} \left[e_2^{(v)} - p_{em}(e_2, a_2^{(v)}) \right] \right\} \\ &+ \left[e_1^{(v)} - p_{em}(e_1, a_1^{(v)}) \right] \left\{ (1-e_{sr}) \left[e_2^{(v)} - p_{em}(e_2, a_2^{(v)}) \right] + e_{sr} \left[e_2 - p_{em}(e_2, a_2^{(v)}) \right] \right\} \end{aligned} \tag{4.32}$$

and

$$\begin{aligned} &\Pr\left(\widehat{y}_1 = \widehat{y}_2^{(v)},\ \widehat{y}_1^{(v)} = \widehat{y}_2\right) = p_{12(v1)(v2)} \\ &+ \left[e_1 - p_{em}(e_1, a_1^{(v)}) \right] \left\{ (1-e_{sr}) \left[e_2^{(v)} - p_{em}(e_2, a_2^{(v)}) \right] + e_{sr} \left[e_2 - p_{em}(e_2, a_2^{(v)}) \right] \right\} \\ &+ \left[e_1^{(v)} - p_{em}(e_1, a_1^{(v)}) \right] \left\{ (1-e_{sr}) \left[e_2 - p_{em}(e_2, a_2^{(v)}) \right] + e_{sr} \left[e_2^{(v)} - p_{em}(e_2, a_2^{(v)}) \right] \right\} \end{aligned} \tag{4.33}$$

Proof. See Appendix B.7. □

With the above formulae and the pairwise SMP given by (B.2), we use a procedure similar to that presented in B.5 to evaluate the covariance matrix of the ER estimators and obtain

Theorem 4.7. For a two-link BPSK-based network using a virtual R_1D link, the variances of the VLA estimators, $\widehat{e}_1, \widehat{e}_2$ are given by

$$\mathrm{Var}[\widehat{e}_1|e_1, e_2, a_1^{(v)}] = \frac{b_2^2 p_{12}(1-p_{12}) - 2b_2b_3 + p_{(v1)2}(1-p_{(v1)2})}{(1-2e_2)^2(b_2-b_1)^2N} \tag{4.34}$$

$$\mathrm{Var}[\widehat{e}_2|e_1, e_2, a_1^{(v)}] = \frac{b_1^2 p_{12}(1-p_{12}) - 2b_1b_3 + p_{(v1)2}(1-p_{(v1)2})}{(1-2e_1)^2(b_2-b_1)^2N} \tag{4.35}$$

where $p_{12}, p_{(v1)2}$ are the SMPs for the link pairs $(1,2)$ and $(v1,2)$, $b_1 = \frac{a_1^{(v)}}{[a_1^{(v)}-(a_1^{(v)}-1)(1-2e_1)^2]^{3/2}}$, $b_2 = \frac{1}{[a_1^{(v)}-(a_1^{(v)}-1)(1-2e_1)^2]^{1/2}}$, $b_3 = p_{12(v1)} - p_{12}p_{2(v1)}$. Moreover, if noise of power $(a_i^{(w)}-1)\sigma_d^2$ is injected, then the variance of the noise-enhanced ISI-VLA estimator $\widehat{e}_i^{(w)}$ is given by

$$\frac{a_i^{(w)}}{\left[1+(a_i^{(w)}-1)(1-2e_i^{(w)})^2\right]^3}\mathrm{Var}\left[\widehat{e}_i^{(w)}\,\middle|\,e_1^{(w)}, e_2^{(w)}, a_1^{(v)}\right] \tag{4.36}$$

If a virtual R_2D link is used instead, then (4.34)-(4.36) should be modified by replacing $a_1^{(v)}$, e_1, $p_{(v1)2}$, and $p_{12(v1)}$ with $a_2^{(v)}$, e_2, $p_{1(v2)}$ and $p_{12(v2)}$, respectively.

Note that the notations used in (4.34) and (4.35) imply that the variance of $\widehat{e}_i$ is a function of e_1, e_2, and $a_1^{(v)}$ only. All other parameters, e.g. b_i's, depend on these three parameters. For the case addressed in *Theorem 4.7*, the optimal scaling factors can be obtained by finding the extreme points of (4.36)–a highly nonlinear function of $a_1^{(w)}, a_2^{(w)}$.

The performance analysis of an ISI-VLA estimator for the hidden SR link is more involved. We need the following preliminary result.

Lemma 4.8. For a single-relay CCN with single virtual SD and RD link, the (i,j)th entry of the covariance matrix $\mathbf{C}$ of the indicator vector, $\left[I\left(\widehat{y}_1=\widehat{y}_2\right)\ I\left(\widehat{y}_1^{(v)}=\widehat{y}_2\right)\ I\left(\widehat{y}_1=\widehat{y}_2^{(v)}\right)\ I\left(\widehat{y}_1^{(v)}=\widehat{y}_2^{(v)}\right)\right]^T$, is given by

$$\mathbf{C}_{ij} = \begin{cases} p_{kl}(1-p_{kl}) & \text{if } k=l,\ k'=l' \\ \Pr(\widehat{y}_k=\widehat{y}_l, \widehat{y}_{k'}=\widehat{y}_{l'}) - p_{kl}p_{k'l'} & \text{otherwise} \end{cases} \tag{4.37}$$

for $i,j = 1,\cdots,4$, with the mapping $i \to (k,l)$ defined by

$$k = \begin{cases} 1 & \text{if } i \text{ is odd} \\ (v1) & \text{otherwise} \end{cases},\ l = \begin{cases} 2 & i \le 2 \\ (v2) & i \ge 3 \end{cases} \tag{4.38}$$

and a similar mapping from j to (k', l'). The corresponding inverse Jacobian $\mathbf{J}^{-1}$ is given by

$$\begin{pmatrix} 2e_2-1 & (2e_1-1)(1-2e_2) & (2e_1-1)(1-2e_{sr}) & 0 \\ (2e_2-1)h'(e_1,a_1^{(v)}) & (2e_1^{(v)}-1)(1-2e_2) & (2e_1^{(v)}-1)(1-2e_{sr}) & 0 \\ 0 & (2e_3-1)(1-2e_2^{(v)}) & (2e_3-1)(1-2e_{sr})h'(e_2,a_2^{(v)}) & 2e_2^{(v)}-1 \\ 0 & (2e_3^{(v)}-1)(1-2e_2^{(v)}) & (2e_3^{(v)}-1)(1-2e_{sr})h'(e_2,a_2^{(v)}) & (2e_2^{(v)}-1)h'(e_1,a_1^{(v)}) \end{pmatrix} \tag{4.39}$$

where $h'(x,a) = \frac{a^{(v)}}{[a^{(v)}+(1-a^{(v)})(1-2x)^2]^{3/2}}$.

We immediately have

Theorem 4.9. For a single-relay BPSK-based CCN with a $a_{sd}^{(v)}$-scaled virtual SD link and a $a_{rd}^{(v)}$-scaled virtual RD link, as described by (3.7.a)–(3.8), the variances for the VLA estimators $\widehat{e}_{sr}, \widehat{e}_{rd}$ and $\widehat{e}_{sd}$ are given by

$$\mathrm{Var}[\widehat{e}_{sr}] = \frac{\widetilde{\mathbf{C}}_{22}}{N}, \mathrm{Var}[\widehat{e}_{rd}] = \frac{\widetilde{\mathbf{C}}_{33}}{N}, \tag{4.40}$$

$$\mathrm{Var}[\widehat{e}_{sd}] = \frac{\widetilde{\mathbf{C}}_{11}+\widetilde{\mathbf{C}}_{14}+\widetilde{\mathbf{C}}_{41}+\widetilde{\mathbf{C}}_{44}}{4N} \tag{4.41}$$

where $\widetilde{\mathbf{C}} = \mathbf{J}\mathbf{C}\mathbf{J}^T$, and $\widetilde{\mathbf{C}}_{ij}$ denotes the element in the ith row and jth column of $\mathbf{C}$.

Furthermore, the variance of the ISI-VLA estimators, $\widehat{e}_{sr}^{(w)}$, $\widehat{e}_{rd}^{(w)}$, and $\widehat{e}_{sd}^{(w)}$, are

$$\mathrm{Var}[\widehat{e}_{sr}^{(w)}] = \frac{\widetilde{\mathbf{C}}_{22}^{(w)}}{N}, \tag{4.42}$$

$$\mathrm{Var}[\widehat{e}_{rd}^{(w)}] = \frac{a_{rd}^{(w)}}{\left[1+(a_{rd}^{(w)}-1)(1-2e_{rd}^{(w)})^2\right]^3}\frac{\widetilde{\mathbf{C}}_{33}}{N} \tag{4.43}$$

$$\mathrm{Var}[\widehat{e}_{sd}^{(w)}] = \frac{a_{sd}^{(w)}}{\left[1+(a_{sd}^{(w)}-1)(1-2e_{sd}^{(w)})^2\right]^3}\frac{\widetilde{\mathbf{C}}_{11}^{(w)}+\widetilde{\mathbf{C}}_{14}^{(w)}+\widetilde{\mathbf{C}}_{41}^{(w)}+\widetilde{\mathbf{C}}_{44}^{(w)}}{4N} \tag{4.44}$$

where $\widetilde{\mathbf{C}}^{(w)} = [\widetilde{\mathbf{C}}_{ij}^{(w)}] = \mathbf{J}^{(w)}\mathbf{C}^{(w)}\left(\mathbf{J}^{(w)}\right)^T$, and $\mathbf{J}^{(w)}$, $\mathbf{C}^{(w)}$ are computed after noise-injection into all but the SR link.

We summarize below a few remarks regarding the above properties, their extensions and the proposed noise-enhanced estimator in general.

R1 The noise samples play the dual role of i) generating VLs to eliminate the needs for CSI and extra RNs and resolve the symmetric ambiguity and ii) altering the statistical property of the received samples.

R2 As the identity, (2.14), which relates an SMP to the associated ERs, involves two independent links, the three-link network has the special property of offering $\binom{3}{2}$=3 link-pairs such that each link participates in two link-pairs. Such an "uniform participation" is important to guarantee uniform performance, i.e., the MSEE performance for each link is the same if the true ERs are identical. In general for a network with four or more links the number of link-pairs is larger than the number of independent links and the performance of an ER estimator for a particular link depends on the number of link-pairs it has participated.

R3 Although *Theorems 4.4, 4.5* consider a three-link network only, extensions to networks with more independent component links are straightforward but closed-form expressions for the corresponding optimal scaling factor and noise benefit interval can only be determined numerically. Nevertheless, for the special cases considered by both theorems, the minimum achievable MSEE reduction ratio tends to $O\left(\frac{1}{\mathrm{SNR}_i}\right)$ at high SNRs.

R4 *Theorems 4.7, 4.9* give the MSEE expressions for BPSK-based VLA and ISI-VLA estimators but we are not able to derive closed-form expressions for the noncoherent modulation based networks. The optimal injected-noise power levels for noncoherent networks with correlated links seem to be mathematically intractable. However, our analysis indicates that a key factor in the MSEE expression is the square of the first derivative of the conversion function (rule) with respect to the scaling factor which is of order $(a_i^{(w)})^{-2}$ for small ERs; see, e.g., (4.23). The increase of $a_i^{(w)}$ reduces this factor's value but it also impact on other parameters that might increase the MSEE. For examples, in (4.23) $a_i^{(w)}$ is fixed by the identical

$e_i^{(w)} = \epsilon$ constraint and is not a independent parameter while in (4.36) $a_i^{(w)}$ affects every parameters on the second rational term. The optimal $a_i^{(w)}$ strikes a best balance between these conflicting effects. Numerical experiments reported in the next section show that, similar to the special cases addressed in *Theorems 1, 2*, there is a proper range of injected noise power levels for enhancing the performance with added noise and an optimal scaling factor (added noise power level) does exist.

R5 Similar to the EVLA scheme, we can add $n_{vl} - 1$ virtual RD and/or SD links to obtain the same number of estimates for $\{\widehat{e}_{sd}\}$ and/or $\{\widehat{e}_{rd}\}$, each with the same reduced variance, and then take average on the resulting n_{vl} estimators. This sample-mean approach guarantees improved performance but the improvement ratio is bounded by $1/n_{vl}$ due to the correlations amongst VLs. The resulting multiple VLs algorithm is called the enhanced ISI-VLA (EISI-VLA) estimator.

4.5 Numerical Results

To verify our MSEE analysis, we first consider the data-aided point-to-point BPSK communication in Rayleigh fading channel (toy example case), illustrated in Fig. 4.2. As can be seen, our analysis is consistent with simulation results and there is indeed a optimal injected noise power. Based on Theorem 4.2, the optimal value of $a^{(w)}$ is $1000 = 30$ (dB), which is the channel quality. In practice, the noise variance should be estimated with some estimation error. Fortunately, this simulation also indicates that the performance is not sensitive to the optimal value of $a^{(w)}$. This nonsensitive property shows the possibility to implement the noise-enhanced estimator in the real world.

Then, we consider a 3-link wireless sensor network in Fig. 4.3 which shows that, for all three binary modulations considered, the analytic predictions are very close to those obtained by simulations even when the sample size is small, and both give identical results if the sample size is large. Similar performance trend for the ISI-VLA scheme in a BPSK-based single-relay CCN is found in the same figure. The normalized MSEE

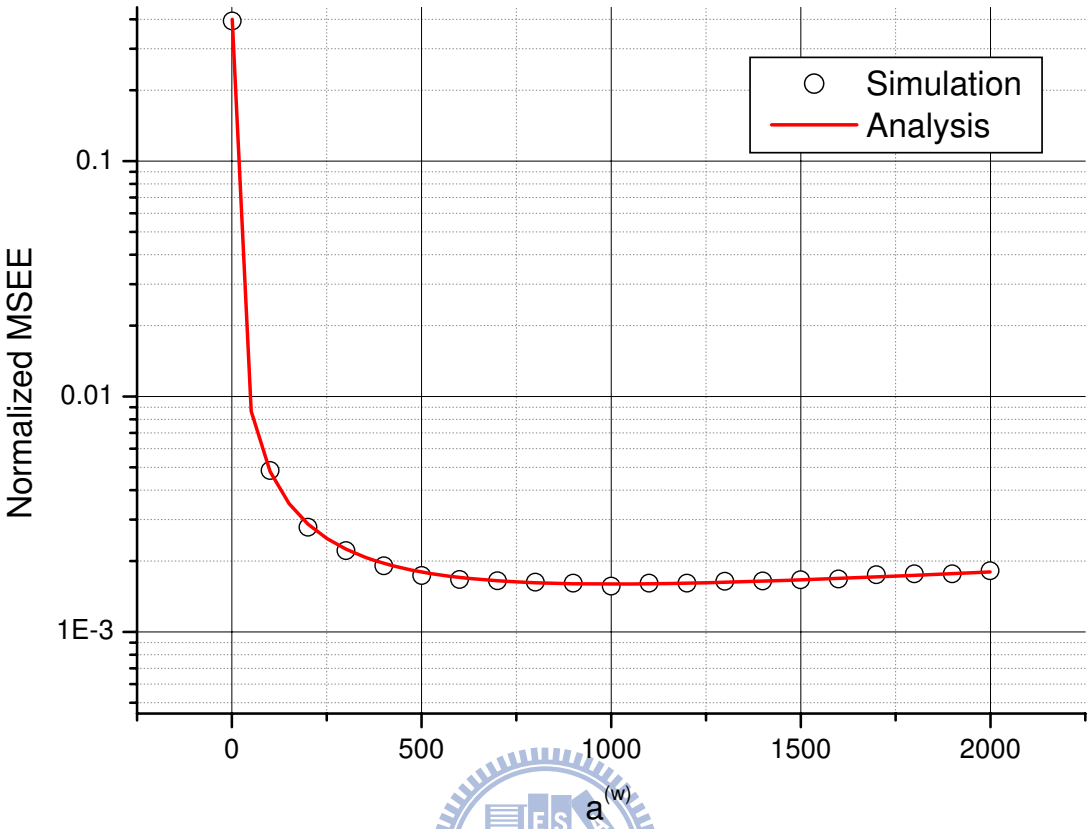


Figure 4.2: Normalized MSEE performance of the noise-enhanced estimator for data-aided point-to-point BPSK communication in Rayleigh fading channel. The channel quality is 30 dB (error rate $e = 2.5 \times 10^{-4}$). The sample size is 10000.

performance, $E[(\widehat{e} - e)^2]/e^2$, where e is the true ER, of the VLA, VLA-EM, and EISI-VLA estimation schemes for a BFSK-based single-relay CCN network is shown in Fig. 4.4. The VLA-EM scheme refers to a modified version of the EM based estimator of [33], which did not consider the hidden SR link. The modifications are needed to apply a VL for resolving the ambiguity and replace the normalization factor such that the equation for updating the ER estimate for the cascaded link becomes

$$Q_k^{(i+1)} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{I(\widehat{y}_j=\widehat{y}_k)}\left(1-Q_j^{(i)}\right)^{1-I(\widehat{y}_j=\widehat{y}_k)}}{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{I(\widehat{y}_j=\widehat{y}_k)}\left(1-Q_j^{(i)}\right)^{1-I(\widehat{y}_j=\widehat{y}_k)}} + \frac{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{1-I(\widehat{y}_j=\widehat{y}_k)}\left(1-Q_j^{(i)}\right)^{I(\widehat{y}_j=\widehat{y}_k)}}{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{I(\widehat{y}_j=\widehat{y}_k)}\left(1-Q_j^{(i)}\right)^{1-I(\widehat{y}_j=\widehat{y}_k)}}\right)^{-1} \tag{4.45}$$

where Q_i's are defined in chapter 2 with the superscripts denote the associated iteration number. The ISI method injects additional noise to estimate the ERs of the resulting links and then converting them back to $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ via the analytic formulas given in Table 3.2. The performance curves clearly demonstrate that the advantage of the VLA-EM scheme against the VLA estimator is negligible while the EISI-VLA scheme far outperforms the other two schemes.

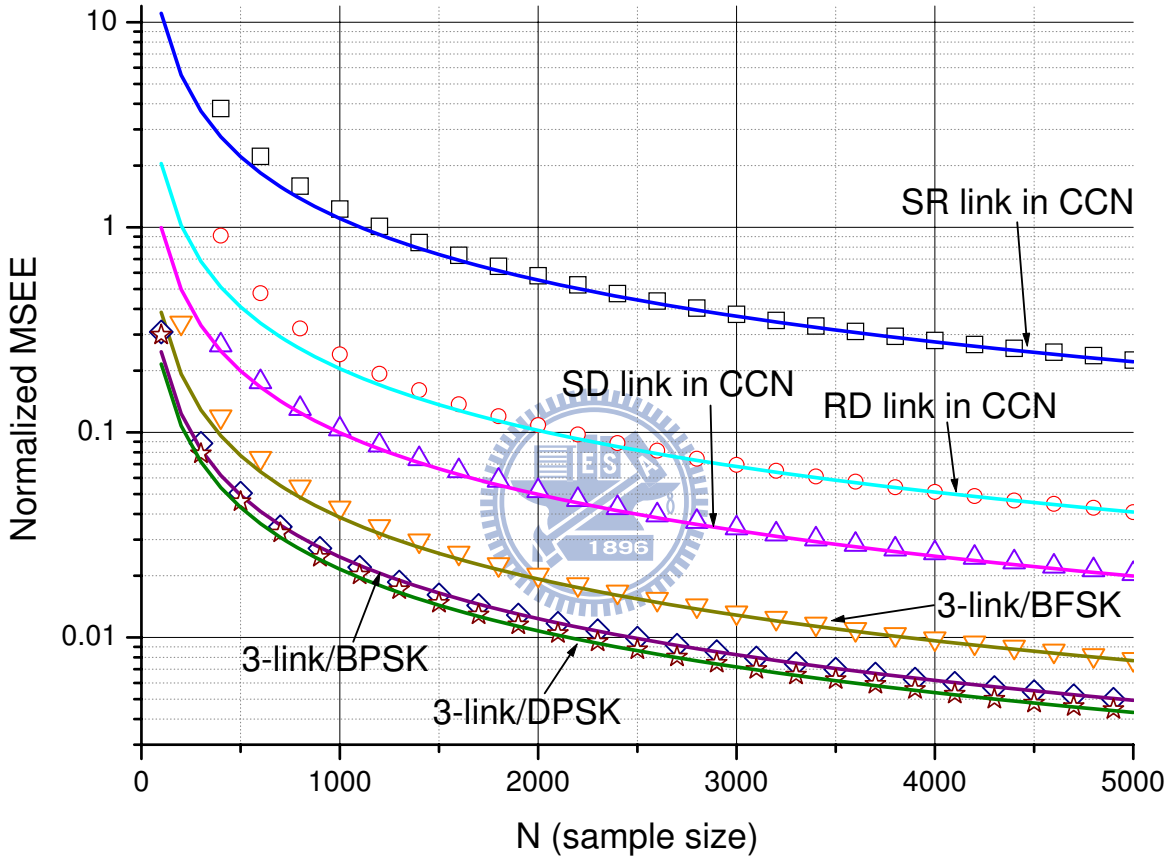


Figure 4.3: Normalized MSEE performance of the ISI-VLA scheme for (a) various binary modulated 3-link networks ($e_1 = 0.003$, $e_2 = 0.002, e_3 = 0.001$; the injected noise power is such that SH SNR=2 for link 1 and $e_1^{(w)} = e_2^{(w)} = e_3^{(w)}$) and (2) BPSK-based single-relay CCN ($e_{sr} = 0.02922$, $e_{rd} = 0.001988$, $e_{sd} = 0.04356$, $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$, $a_{sd}^{(w)} = 1$ and $a_{rd}^{(w)} = 30$). For 3-link networks, only the performance of $\widehat{e}_1$ is shown. The analytic predictions (solid curves) for these two scenarios are based on (4.23) and (4.42)–(4.44), respectively.

Fig. 4.5 plots the MSEE reduction ratio as a function of the scaling factor $a_{sd}^{(w)}$ while the other scaling factor $a_{rd}^{(w)}$ is chosen such that $e_{rd}^{(w)} = e_{sd}^{(w)}$. These curves reveal that

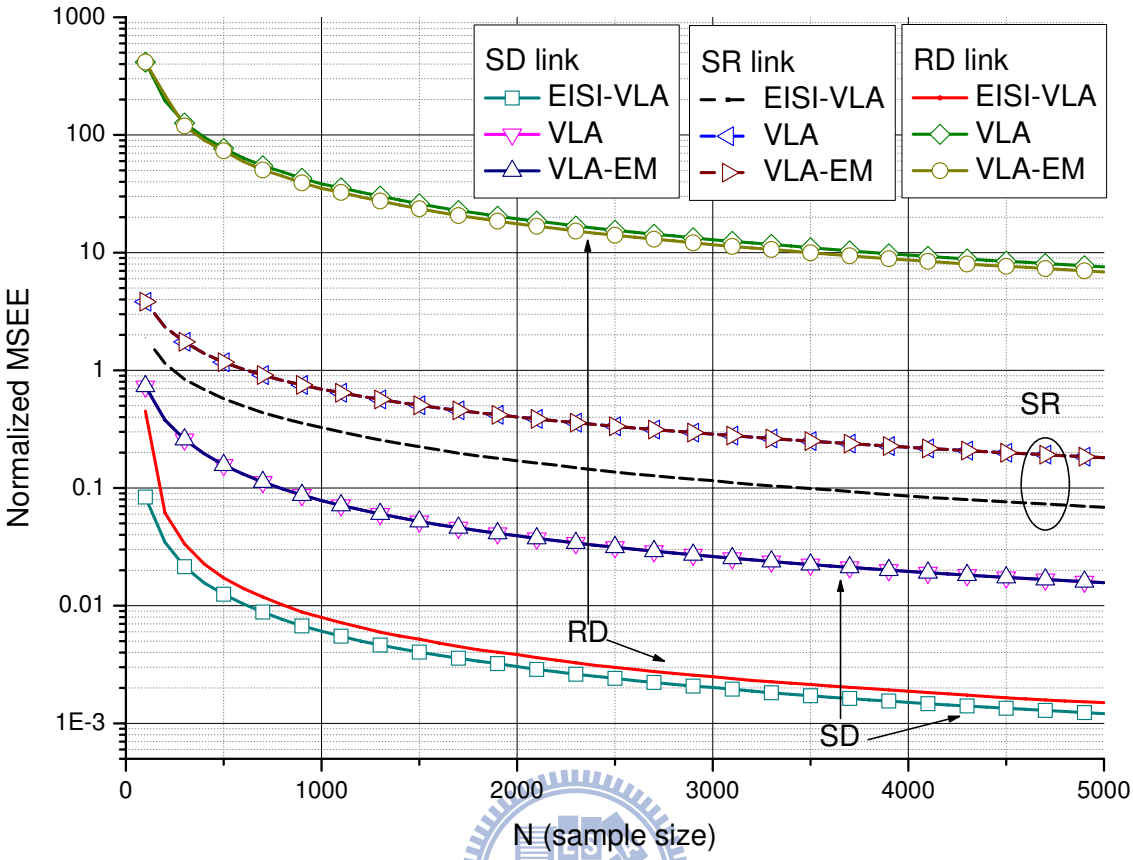


Figure 4.4: Normalized MSEE performance of VLA, VLA-EM, and EISI-VLA schemes in a BFSK-based single-relay CCN with $e_{sr} = 0.0127$, $e_{rd} = 5.0711 \times 10^{-5}$ and $e_{sd} = 0.0298$. Other parameter values used are: $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$, $e_{sd}^{(w)} = e_{rd}^{(w)} = 0.05$ and $n_{vl} = 30$.

the MSEE performance is improved by injecting proper noise power into the received samples and there is an optimal injected noise power that achieves the maximum MSEE improvement. This phenomena is called the stochastic resonance effect which has been observed in some nonlinear systems; see [6] and reference therein. We also notice that the improvement is more impressive when the true ER becomes smaller, which is consistent with what the IS theory has predicted. The noise benefit interval (NBI), defined as the range of the scaling factor values within which the MSEE reduction ratio is less than 1, is a function of the true e_{sd} and e_{rd}. As mentioned before, we are not able to derive closed-form expressions for the optimal scaling factors used in a noncoherent network. Nevertheless, extensive simulations suggest that it is a good strategy to make

$e_{sd}^{(w)} = e_{rd}^{(w)} \approx 0.05$ if both e_{sd} and e_{rd} are much smaller than 0.05. As was explained in Chapter 4, because of the availability of improved estimates for e_{sd} and e_{rd}, the performance of $\widehat{e}_{sr}$ is also improved although we do not and could not inject noise into samples received at RNs.

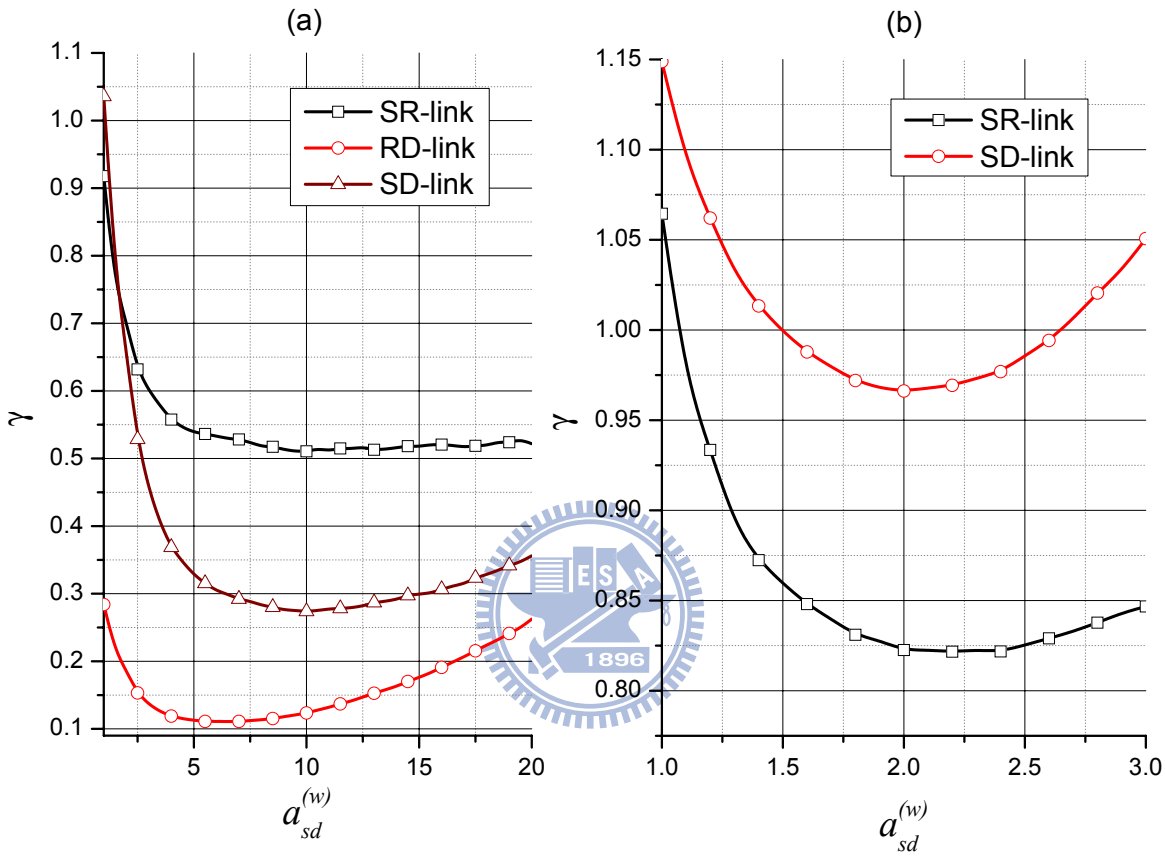


Figure 4.5: MSEE reduction ratio (γ) performance of the ISI-VLA estimator with BFSK modulation and $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$. Part (a) is obtained by assuming $d_{sr} = 5$, SH-SNR=25 dB with the path loss exponent = 2 (which leads to $e_{sr} = 0.0016, e_{rd} = 0.0016, e_{sd} = 0.0062$). Part (b) assumes that $d_{sr} = 8$, SH-SNR=18 dB with path loss exponent = 4 so that $e_{sr} = 0.0127, e_{rd} = 5.0711 \times 10^{-5}, e_{sd} = 0.0298$. The MSEE reduction ratio of the RD link is not shown in part (b) as it is relatively small ($\sim O(10^{-3})$).

Although proper noise-injection does improve the convergence rate performance, in some cases such as those shown in Fig. 4.5, the improvement is not quite as significant as one wishes. The MSEE reduction ratio can be further improved by the enhanced ISI-VLA estimator as is shown in Fig. 4.6 where the simulation conditions are identical

to those assumed in Fig. 4.5(b). As expected, the performance is improved with the increase of n_{vl} and the improvement is much more impressive when the true ER is small: the required sample size reduction is more than 10 times for the SD link and is greater than 8000 times for the RD link when $n_{vl} = 30$. Another benefit of using multiple VLs is that the NBI becomes larger as n_{vl} increases.

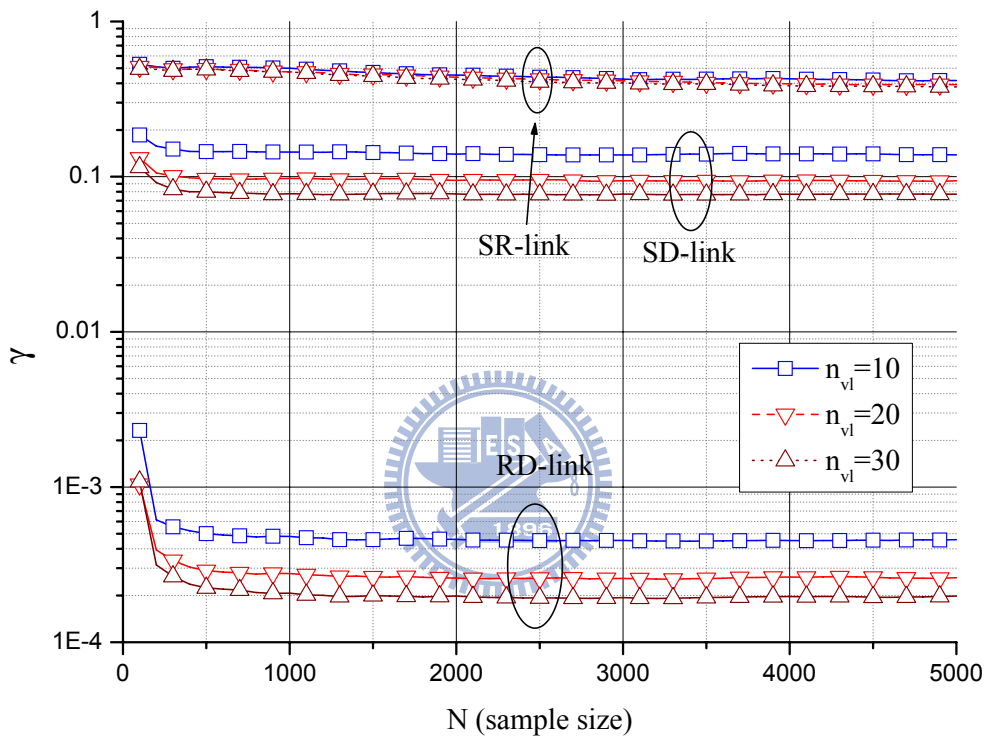


Figure 4.6: MSEE reduction ratio behavior of the EISI-VLA estimator for BFSK based CCN with different n_{vl}. Other system parameter values are the same as those of Fig. 4.5(b).

Chapter 5

Data fusion and blind multiple error rate estimation in a non-binary modulation based wireless sensor network

For optimal M-hypothesis detection in a parallel system, the performance of the sensor nodes must be available at the fusion center (FC). Such information can be obtained by the LJW blind estimator [34] based on a multinomial distribution model with parameters related to each links' ERs. To get the estimates, we need to solve a nonlinear optimization problem. As shown in section 5.2, this algorithm is not feasible for large M due to the prohibitively high computational complexity. To approximate the optimal detector, we propose a suboptimal detector based on bit-level representation and a corresponding blind estimator to estimate the error rate of sensor nodes in section 5.3. The complexity of our estimator is much lower than that of LJW as we are able to obtain a closed-form salutation instead of employing an iterative algorithm for solving a nonlinear optimization. To further improve the convergence rate, we propose a noise-enhanced estimator in section 5.5. Simulation results show that the proposed suboptimal detector using the proposed blind estimator render negligible performance loss with respect to that of the optimal detector. A stochastic resonance phenomenon is observed in the estimator's mean square estimation error performance.

5.1 System model and optimal detector

The parallel sensing system illustrated in Fig. 5.1 consists of one source, L sensors, and one FC. In a wireless sensor network, source transmits a signal from M candidate signals $\{s_0, s_1, \cdots, s_{M-1}\}$ with a prior probability $\mathrm{P}(H_i)$ =P(s_i is transmitted). Each sensor detects and forwards its decision to the FC which then determines which hypothesis (H_i) is true based on the signals forwarded by the sensors. In this dissertation, we assume that either the sensing (source-sensor) channels or the reporting (sensors-FC) channels is error-free. Such an assumption losses no generality as each combined source-sensor-FC link can be modeled as an equivalent composite channel [34]. The optimal fusion rule depends on the parameters of the equivalent channels which can be estimated by the method described below.

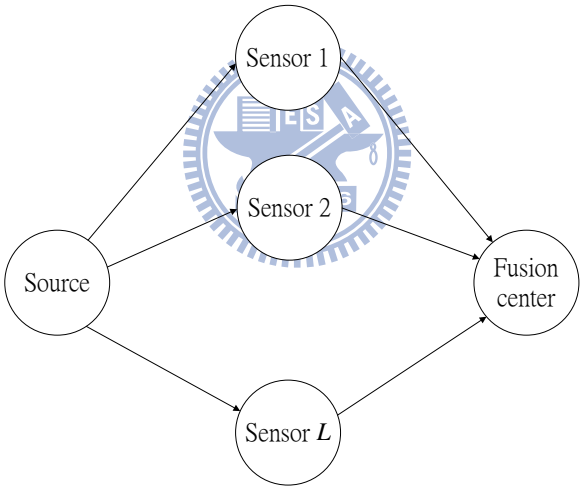


Figure 5.1: Parallel distributed detection system

Denote by $\widehat{y}_j \in \{d_i, i = 0, \cdots, M-1\}$ the jth sensor's hard decision and define the event probability e_j^{ik}

$$e_j^{ik} = \mathrm{P}(\widehat{y}_j = d_k | H_i),\ j = 1, \cdots, L,\ i, k = 1, \cdots, M, i \neq k$$

Assuming $\mathrm{P}(H_i) = 1/M,\ i = 0, \cdots, M-1$ and the sensor-FC links are noiseless, we

have the optimal data fusion rule [34]

$$\widehat{y} = \arg \max_{i=0,\cdots,M-1} \prod_{j=1}^{L} \prod_{k=1}^{M} \left(e_j^{ik}\right)^{I(\widehat{y}_j = d_k)} \tag{5.1}$$

where $I(\widehat{y}_j = d_k)$ is the binary valued function indicating if the statement $\widehat{y}_j = d_k$ is true.

5.2 LJW blind ER estimator

In practice, the error rates e_j^{ik} are usually unknown at FC and need to be estimated. Pilot-assisted estimates can be obtained by sensors and forwarded to FC, which, however, is not feasible for many sources and/or battery-limited sensors.

For convenience of reference, we briefly outline the LJW estimator[1] in the followings.

Let $l = \sum_{j=1}^{L} u_j M^{j-1}$, where $u_j = k$ if $\widehat{y}_j = d_k$, then the probability $\mathrm{P}(\mathbf{u}_l) = \mathrm{P}(u_1, u_2, \cdots, u_L) = p_l$ can be expressed as

$$p_l = \sum_{i=0}^{M-1} \mathrm{P}(H_i) \prod_{j=1}^{L} \mathrm{P}(u_j|H_i) = \sum_{i=0}^{M-1} \mathrm{P}(H_i) \prod_{j=1}^{L} e_j^{iu_j} \tag{5.2}$$

As there are $C = M^L$ sensor decision combinations of $\mathbf{u}_l$, if we denote the number of the jth decision combination by x_i, then the occurrence numbers of $\mathbf{u}_l$, $(x_1, \cdots, x_C)$, are multinomial distributed with parameters $\{p_l\}$

$$\mathrm{P}(x_1, \cdots, x_C) = \frac{N}{x_1! \cdots x_C!} p_1^{x_1} \cdots p_C^{x_C}$$

where N is the number of reported sensing samples. Given the occurrence number, the LJW blind estimator is

$$\widehat{\theta} = \arg \max_{\theta} p_1^{x_1} \cdots, p_C^{x_C} \tag{5.3}$$

where $\theta = \{e_j^{ik}\}$, for $j = 1, \cdots, L,\ i, k = 0, \cdots, M-1$.

[1]Notice that the assumption of the prior probabilities in [34] is different from ours. In [34], the prior probabilities $\mathrm{P}(H_i)$ are unknown and can be estimated by LJW blind ER estimator.

For an M-hypothesis, L-sensor problem, we have $L(M-1)M$ unknown parameters e_j^{ik}. There are 720 unknown parameters to be estimated for the case $M = 16$ and $L = 3$. An iterative method is often needed to solve the corresponding nonlinear optimization problem (5.3). Such a large scale optimization requires extremely high computational complexity and large memory size. Thus the LJW blind estimator is not very practical unless M and L are very small.

5.3 Optimal/suboptimal fusion rules for nonbinary signals

For $M = 2$, the blind estimator proposed in previous chapters provides a low-complexity suboptimal alternate to (5.3). This estimator is based on the pairwise comparisons of two independent link outputs which are categorized into matched (same hard decisions) or unmatched (different hard decisions) outputs. The extension to the nonbinary case becomes much more complicated as either a matched or unmatched result is caused by multiple joint events in each selected link pair's M-ary decisions.

We notice that the LJW blind estimator does not take the sensors' transmission method into account and, with all its generality, requires high computational complexity for large M and L. In contrast, we assume that the sensors (or the source) use a nonbinary modulation scheme and show that the structure of the signal constellation can be used to simplify a large-scale problem greatly or to decompose a large-scale problem into several small-scale problems, which reduce the computational complexity significantly. We consider two modulation schemes: orthogonal and M-QAM modulations; blind estimators for networks employing other nonbinary modulations such as M-PAM and M-PSK can be similarly treated and derived.

Recall that the optimal detector is

$$\widehat{y} = \arg \max_{i=0,\cdots,M-1} \prod_{j=1}^{L} \prod_{k=1}^{M} \left(e_j^{ik}\right)^{I(\widehat{y}_j = d_k)} \tag{5.4}$$

and the number of unknown parameters is $L(M-1)M$ due to the unequal of e_j^{ik} in general. If all of e_j^{ik} are equal for $i \neq k$, the optimal detector can reduce to be

$$\widehat{y} = \arg \max_{i=0,\cdots,M-1} \prod_{j=1}^{L} (1-e_j)^{I(\widehat{y}_j=d_i)} \prod_{k=1,\neq i}^{M} \left(\frac{e_j}{M-1}\right)^{I(\widehat{y}_j=d_k)} \tag{5.5}$$

where e_j is the symbol error rate of the jth link. To perform the optimal detection, we only need to estimation L unknown parameters.

For M-QAM modulation scheme, we first express the symbol-level decision for the jth sensor-FC link $\widehat{y}_j$ in terms of its bit-level decisions, i.e., $\widehat{y}_j = (\hat{y}_j^1, \hat{y}_j^2, \cdots, \hat{y}_j^k)^T$, where $M = 2^k$ and $\hat{y}_j^i$ represents the ith bit for the hard decision $\widehat{y}_j$. Converting an M-ary communication link into k parallel binary symmetric links with with crossover probabilities e_j^k, the bit error probability of $\widehat{y}_j^k$ (Fig. 5.2). Then, we have

$$\frac{\mathrm{P}(\widehat{y}_j^k|b_k=0)}{\mathrm{P}(\widehat{y}_j^k|b_k=1)} = \frac{\prod_{j=1}^{L}(e_j^k)^{\widehat{y}_j^k}(1-e_j^k)^{1-\widehat{y}_j^k}}{\prod_{j=1}^{L}(e_j^k)^{1-\widehat{y}_j^k}(1-e_j^k)^{\widehat{y}_j^k}} = \prod_{j=1}^{L}\left(\frac{1}{e_j^k}-1\right)^{1-2\widehat{y}_j^k} \stackrel{def}{=} \Lambda(\widehat{y}_j^k) \tag{5.6}$$

where b_k is the k bit of the source output symbol.

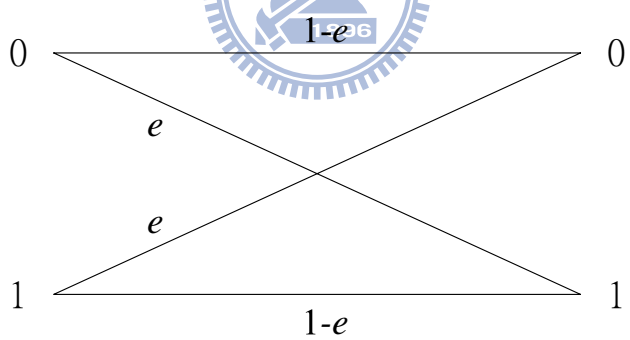


Figure 5.2: A binary symmetric channel with parameter e

Since

$$\ln \Lambda(\widehat{y}_j^k) = \sum_{j=1}^{L}(1-2\widehat{y}_j^k)\ln\left(\frac{1}{e_j^k}-1\right) \tag{5.7}$$

the optimal bit-level fusion rule, assuming $b_k = 0$ or 1 with equal probability, is

$$\widehat{y}^k = \begin{cases} 0 & \text{if } \sum_{j=1}^{L}\left(1-2\widehat{y}_j^k\right)\ln\left(\frac{1}{e_j^k}-1\right) > 0 \\ 1 & \text{otherwise} \end{cases} \tag{5.8}$$

where only the bit error rate e_j^k is involved.

A symbol-level fusion decision can be easily obtained once all bit-level decisions are known. This fusion rule reduces the number of parameters to be estimated from $LM(M-1)$ in (5.1) to $L\log_2 M$.

The above derivation assumes equivalent binary symmetric links, i.e., bit error rates for information bit 0 and 1 are identical. Clearly, QPSK modulation satisfies this assumption if the source shows no preference while M-QAM ($M > 4$) modulation may not. In the latter case, we need to consider a binary *asymmetric* channel (Fig. 5.3). The log likelihood ratio for such a channel with error rates e^0 and e^1 is given by

$$\begin{aligned}
&\ln\left(\frac{\mathrm{P}(\widehat{y}_j{}^k|b_k=0)}{\mathrm{P}(\widehat{y}_j{}^k|b_k=1)}\right)\\
&=\ln\left(\frac{\prod_{j=1}^{L}(e_j^{k,0})^{\widehat{y}_j^k}(1-e_j^{k,0})^{1-\widehat{y}_j^k}}{\prod_{j=1}^{L}(e_j^{k,1})^{1-\widehat{y}_j^k}(1-e_j^{k,1})^{\widehat{y}_j^k}}\right)\\
&=\sum_{j=1}^{L}\widehat{y}_j^k\ln\left(\frac{e_j^{k,0}}{1-e_j^{k,1}}\right)+(1-\widehat{y}_j^k)\ln\left(\frac{1-e_j^{k,0}}{e_j^{k,1}}\right)
\end{aligned}\tag{5.9}$$

where $e_j^{k,b}$ denotes the error rate of the kth bit in the binary representation of $\widehat{y}_j$ given bit b has been transmitted. The corresponding bit-level fusion rule becomes

$$\widehat{y}^k=\sum_{j=1}^{L}\widehat{y}_j^k\ln\left(\frac{e_j^{k,0}}{1-e_j^{k,1}}\right)+(1-\widehat{y}_j^k)\ln\left(\frac{1-e_j^{k,1}}{e_j^{k,1}}\right)\underset{1}{\overset{0}{\gtrless}}0\tag{5.10}$$

For the above asymmetric case, there are $2L\log_2 M$ parameters, $\{e_j^{k,0}, e_j^{k,1}\}$, to be estimated, which is still much less than the number of unknown parameters in (5.1).

5.4 Blind symbol/bit ERs estimator

At the beginning, we focus on the orthogonal modulation and want to estimate the symbol ERs. For convenience, we consider a three links wireless sensor networks. Later, we will show that a wireless sensor network with more links can be decomposed into several wireless sensor networks involving three links only.

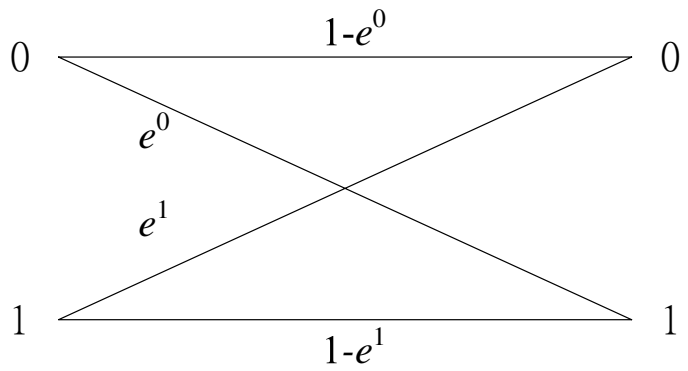


Figure 5.3: A binary nonsymmetric channel with parameter e^0 and e^1

To derive a estimator, we first define a SMP as $\mathrm{P}(\widehat{y}_{j_1} = \widehat{y}_{j_2})$ (the probability that the hard decisions from the ith and jth links are identical). Then, by the law of total probability, we have

$$\begin{aligned}
&\mathrm{P}(\widehat{y}_{j_1} = \widehat{y}_{j_2}) = \sum_{k=1}^{M} \mathrm{P}(\widehat{y}_{j_1} = \widehat{y}_{j_2}|H_k)\mathrm{P}(H_k) \\
=&\sum_{k=1}^{M} \frac{1}{M}\mathrm{P}(\widehat{y}_{j_1} = \widehat{y}_{j_2} \text{ are correct}|H_k) + \sum_{k=1}^{M} \frac{1}{M}\mathrm{P}(\widehat{y}_{j_1} = \widehat{y}_{j_2} \text{ are incorrect}|H_k) \\
=&\sum_{k=1}^{M} \frac{1}{M}(1-e_{j_1})(1-e_{j_2}) + \sum_{k=1}^{M} \frac{e_{j_1}e_{j_2}}{M(M-1)} \\
=&1 - e_{j_1} - e_{j_2} + \frac{M}{M-1}e_{j_1}e_{j_2}
\end{aligned} \tag{5.11}$$

where we use the facts that $\mathrm{P}(H_k) = \frac{1}{M}$ and $\mathrm{P}(\widehat{y}_{j_1} = \widehat{y}_{j_2} \text{ are incorrect}|H_k) = \frac{e_{j_1}e_{j_2}}{M-1}$.

For a three-links wireless sensor network, one has three SMPs, i.e. $\mathrm{P}(\widehat{y}_1 = \widehat{y}_2)$, $\mathrm{P}(\widehat{y}_2 = \widehat{y}_3)$, and $\mathrm{P}(\widehat{y}_1 = \widehat{y}_3)$. These three SMPs and (5.11) yield the following nonlinear system

$$\begin{bmatrix} 1 - e_2 - e_3 + \frac{M}{M-1}e_2e_3 \\ 1 - e_1 - e_3 + \frac{M}{M-1}e_1e_3 \\ 1 - e_1 - e_2 + \frac{M}{M-1}e_1e_2 \end{bmatrix} = \begin{bmatrix} \mathrm{P}(\widehat{y}_2 = \widehat{y}_3) \\ \mathrm{P}(\widehat{y}_1 = \widehat{y}_3) \\ \mathrm{P}(\widehat{y}_1 = \widehat{y}_2) \end{bmatrix} \triangleq \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \tag{5.12}$$

and the solutions are

$$e_i = \frac{-b_i - \sqrt{b_i^2 - 4a_ic_i}}{2a_i} \tag{5.13}$$

where, for $i, j, k \in \{1, 2, 3\}$ and $i \neq j \neq k$,

$$a_i = M^2 p_i - M,\ b_i = -2(M-1)(Mp_i - 1),$$
$$c_i = (M-1)(p_i(M-1) + (p_j + p_k - 1) - Mp_j p_k)$$

Hence, the procedure of the proposed ER estimator consists of two steps and is similar to that shown in Chapter 2. In the first step, we estimate $\{p_i\}_{i=1}^3$ by

$$\widehat{\mathrm{P}}(\widehat{y}_{j_1} = \widehat{y}_{j_2}) = \sum_{l=1}^{N} \frac{I\left(\widehat{y}_{j_1}[l] = \widehat{y}_{j_2}[l]\right)}{N} \tag{5.14}$$

where $\widehat{y}_{j_1}[l]$ ($\widehat{y}_{j_1}[l]$) is the lth hard decision from the j_1th (j_2th) sensor node. Then, by the method of moments, we can estimate ERs by solving (5.12) and we have

$$\widehat{e}_i = \frac{-b_i - \sqrt{b_i^2 - 4a_i c_i}}{2a_i} \tag{5.15}$$

where, for $i, j, k \in \{1, 2, 3\}$ and $i \neq j \neq k$,

$$a_i = M^2 \widehat{p}_i - M,\ b_i = -2(M-1)(M\widehat{p}_i - 1),$$
$$c_i = (M-1)(\widehat{p}_i(M-1) + (\widehat{p}_j + \widehat{p}_k - 1) - M\widehat{p}_j\widehat{p}_k)$$

For QAM modulation scheme, we need to estimate the *bit* ERs $\{e_j^k\}$, not *symbol* ERs. Hence, we need another step to transform the symbol-level decision into bit-level one, yielding a three-step estimator. Specifically, we first de-map the symbol hard decision $\widehat{y}_i$ into bit decisions $\widehat{y}_i^k$, which is often required in a typical digital link. In the second step, we estimate the bit-level SMPs $\mathrm{P}(\widehat{y}_{j_1}^k = \widehat{y}_{j_2}^k)$ for $j_1 \neq j_2$. We then have the the following a basic nonlinear system in a three-link sensor network ((5.12) with $M = 2$)

$$\begin{bmatrix} 1 - e_1^k - e_2^k + 2e_1^k e_2^k \\ 1 - e_2^k - e_3^k + 2e_2^k e_3^k \\ 1 - e_1^k - e_3^k + 2e_1^k e_3^k \end{bmatrix} = \begin{bmatrix} \mathrm{P}(\widehat{y}_1^k = \widehat{y}_2^k) \\ \mathrm{P}(\widehat{y}_2^k = \widehat{y}_3^k) \\ \mathrm{P}(\widehat{y}_1^k = \widehat{y}_3^k) \end{bmatrix} \approx \begin{bmatrix} \widehat{\mathrm{P}}(\widehat{y}_1^k = \widehat{y}_2^k) \\ \widehat{\mathrm{P}}(\widehat{y}_2^k = \widehat{y}_3^k) \\ \widehat{\mathrm{P}}(\widehat{y}_1^k = \widehat{y}_3^k) \end{bmatrix} \tag{5.16}$$

Using the SMP estimates

$$\widehat{\mathrm{P}}(\widehat{y}_{j_1}^k = \widehat{y}_{j_2}^k) = \sum_{l=1}^{N} \frac{I\left(\widehat{y}_{j_1}^k[l] = \widehat{y}_{j_2}^k[l]\right)}{N} \tag{5.17}$$

we obtain the following ER estimate via the method of moments (which can be obtained by (5.15) with $M = 2$)

$$\widehat{e}_j^k = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{(2\widehat{\mathrm{P}}(\widehat{y}_j^k = \widehat{y}_{j_1}^k) - 1)(2\widehat{\mathrm{P}}(\widehat{y}_j^k = \widehat{y}_{j_2}^k) - 1)}{2\widehat{\mathrm{P}}(\widehat{y}_{j_1}^k = \widehat{y}_{j_2}^k) - 1}}, \quad j, j_1, j_2 \in \{1, 2, 3\} \tag{5.18}$$

For some nonbinary modulations, e_j^k is independent of k, hence the countings on the RHS of (5.17) for different k should be averaged to obtain an improved estimator. We summarize the complete procedure in Table 5.1. For general $L > 3$, we can decompose the estimation problem into several estimation subproblems involving only three links. For instance, if $L = 5$, we can consider two estimation subproblems. The first subproblem considers the first three links while the other one involves the last three links. In this case, the ER of the third link is estimated in both subproblems and we can average these two results to get a new estimate with better performance.

Table 5.1: A blind ER estimation algorithm for high order modulation in sensor network.

Input:	Received samples, $\mathbf{y}$
1:	Detect the signals $\widehat{y}_i$ and transform them into bit-level $\widehat{y}_i^k$
2:	Compute SMPs for all link pairs by (5.17).
3:	Compute $\widehat{e}_i$, $i = 1, 2, 3$ via (5.18) using the SMPs obtained in 2
Output:	$\widehat{e}_1$, $\widehat{e}_2$ and $\widehat{e}_3$.

To evaluate the performance of the blind bit-level ER estimation, we can analyze the performance on the binary symmetric channel. Although it is not binary symmetric channel for M-QAM modulation for $M > 4$, the results under this assumption can an approximation of the performance on the binary asymmetric channel. To derive the formula, we use the Delta method and inverse function theorem, as shown in Appendix B.5 and [33]. Based on the Delta method, the MSEE of the blind bit-level ER estimation on the binary symmetric channel is

$$\mathrm{MSEE}[\widehat{e}_j^k] \approx \frac{\left[\mathbf{JCJ^T}\right]_{j,j}}{2N} \tag{5.19}$$

where $\mathbf{C}$ is the covariance matrix, $\mathbf{J}$ is the Jacobian matrix and $[\mathbf{X}]_{i,j}$ denotes the element of the matrix $\mathbf{X}$ in the ith row and jth column. Unlike the binary case in Chapter 4, there is a factor 2 in the denominator. This comes from the factor that there are in-phase and quadrature-phase in M-QAM modulation. Since BERs of the kth bit in these two phases are identical, we have two samples to estimation the BER given a received sample.

Assuming independent links, the covariance matrix of pairwise matching indicator vector $(I(\hat{y}_1 = \hat{y}_2), I(\hat{y}_1 = \hat{y}_3), I(\hat{y}_2 = \hat{y}_3))$ can be shown to be ((B.11) or [33])

$$\mathbf{C} = \begin{pmatrix} p_{12}(1-p_{12}) & p_{123} - p_{12}p_{13} & p_{123} - p_{12}p_{23} \\ p_{123} - p_{12}p_{13} & p_{13}(1-p_{13}) & p_{123} - p_{13}p_{23} \\ p_{123} - p_{12}p_{23} & p_{123} - p_{13}p_{23} & p_{23}(1-p_{23}) \end{pmatrix} \tag{5.20}$$

where

$$p_{ij} = \left(1 - e_i^k\right)\left(1 - e_j^k\right) + e_i^k e_j^k \tag{5.21}$$

$$p_{ijl} = \left(1 - e_i^k\right)\left(1 - e_j^k\right)\left(1 - e_l^k\right) + e_i^k e_j^k e_l^k \tag{5.22}$$

where e_i^k can be evaluated based on (5.28) in the next section. We can derive the Jacobian matrix based on (5.18) by the definition of the Jacobian matrix and expressed it as a function of e_i^k, $i = 1, 2, 3$. This procedure is quite complex. Instead, we first derive the inverse of the Jacobian matrix based on (5.16) and the inverse function theorem. The associated inverse Jacobian matrix is

$$\mathbf{J}^{-1} = \begin{pmatrix} \left(2e_2^k - 1\right) & \left(2e_1^k - 1\right) & 0 \\ \left(2e_3^k - 1\right) & 0 & \left(2e_1^k - 1\right) \\ 0 & \left(2e_3^k - 1\right) & \left(2e_2^k - 1\right) \end{pmatrix} \tag{5.23}$$

With (5.20)-(5.23), we can compute the MSEE on the binary symmetric channel and get the lower bound performance of the blind bit-level ER estimation on the binary asymmetric channel.

For orthogonal modulation scheme, the derivations of the MSEE formula are similar.

For covariance matrix, the formulas of p_{ij} and p_{ijl} are modified as

$$p_{ij} = \left(1 - e_i^k\right)\left(1 - e_j^k\right) + \frac{e_i^k e_j^k}{M-1} \tag{5.24}$$

$$p_{ijl} = \left(1 - e_i^k\right)\left(1 - e_j^k\right)\left(1 - e_l^k\right) + \frac{e_i^k e_j^k e_l^k}{(M-1)^2} \tag{5.25}$$

The other difference is the Jacobin matrix formula since the nonlinear systems (5.12)-(5.22) and (5.16) are not the same. In this case, the inverse Jacobian matrix is

$$\mathbf{J}^{-1} = \begin{pmatrix} \left(\frac{M}{M-1}e_2 - 1\right) & \left(\frac{M}{M-1}e_1 - 1\right) & 0 \\ \left(\frac{M}{M-1}e_3 - 1\right) & 0 & \left(\frac{M}{M-1}e_1 - 1\right) \\ 0 & \left(\frac{M}{M-1}e_3 - 1\right) & \left(\frac{M}{M-1}e_2 - 1\right) \end{pmatrix} \tag{5.26}$$

With (5.20) and (5.24)-(5.26), the MSEE for orthogonal modulation scheme is

$$\text{MSEE}[\hat{e}_j^k] \approx \frac{\left[\mathbf{JCJ^T}\right]_{j,j}}{N} \tag{5.27}$$

.

5.5 Noise-enhanced ER estimations

To further enhance the performance of the blind symbol/bit-level ER estimator, we observe that the performance of the estimator is worse with the increase of the link quality, as shown in Fig. 5.7. This observation also appears in the case with BPSK modulation in Chapter 4. In that chapter, we observe that the estimator of the SMPs is a compare-and-count process (5.17) which is similar to that used in simulation-based ER estimations and IS techniques can be applied. It motives us to inject the noise into the received samples to alter their statistics and perform the estimation based on them. Because the binary symmetric model is similar to the case in Chapter 4 with binary modulation, MSEE can be reduced possibly by injecting noise. To achieve the better performance, we propose a noise-enhanced ER estimations similar to the estimation proposed in 4 with some *modifications*.

The procedure of the proposed noise-enhanced ER estimations involves four steps for M-QAM modulation scheme. First, we inject noise into the received signal before we

perform the demodulation[2]. Then, demodulate the signals and transform the detected signals $\widehat{y}_i$ into bit-level signals. The ER estimator shown in Table 5.1 is performed to obtain the bit ER estimates $\widetilde{e}_i$. Finally, transform the estimates $\widetilde{e}_i$ into the true/original ones $\widehat{e}_i$. So far, we know how to implement the first three steps. The remaining step is how to transform $\widetilde{e}_i$ into $\widehat{e}_i$.

To find a transformation from $\widetilde{e}_i$ to $\widehat{e}_i$, we need to set up the binary labelling because the transformation depends on the binary labelling. In this dissertation, we focus on the Gray mapping labelling [44] and one example (16-QAM with Gray mapping labelling) is shown in Fig. 5.4.

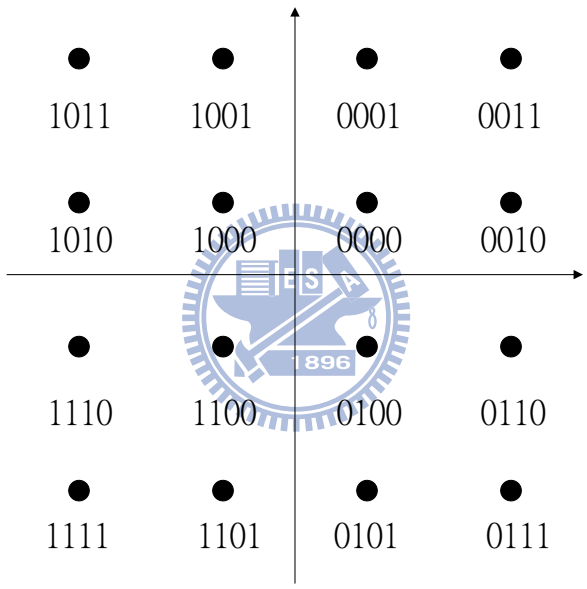


Figure 5.4: 16-QAM with Gray mapping labelling

Fig. 5.4 shows the error rate of the third bit for information bit 0, $e^{3,0}$, is different from that for information bit 1, $e^{3,1}$ because of the nonlinear decision boundary. Simulation results (Fig. 5.6), however, show that approximating the binary nonsymmetric channel with binary symmetric one induces negligible performance loss. Therefore, in the following discussion, we only consider binary symmetric channel, i.e. $e^k = e^{k,0} = e^{k,1}$. Moreover, since M-QAM consists of two independent $\sqrt{M}$-PAM, we only consider the

[2]In practice, the received signals should be quantized due to the limited reporting channel bandwidth. In-depth analysis to the effect of the quantization is beyond the scope of this work.

bits in in-phase components and relabel the index in order from left to right.

In [45], it is shown that the kth bit error probability of M-QAM in Rayleigh fading channel is expressed as

$$e^k = \frac{1}{\sqrt{M}} \sum_{i=0}^{(1-2^{-k})\sqrt{M}-1} \left\{ w(i,k,M) \left(1 - \frac{\sqrt{\frac{3(2i+1)^2 \log_2 M\eta}{2(M-1)}}}{\sqrt{\frac{3(2i+1)^2 \log_2 M\eta}{2(M-1)} + 1}} \right) \right\} \tag{5.28}$$

where η is the energy per bit to noise power spectral density ratio E_b/N_0 and

$$w(i,k,M) = (-1)^{\left\lfloor \frac{i2^{k-1}}{\sqrt{M}} \right\rfloor} \left(2^{k-1} - \left\lfloor \frac{i2^{k-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor \right)$$

Clearly, to find the inverse function is difficult and a numerical method is proposed to find the energy per bit to noise power spectral density ratio η given the error rate e^k.

Notice that several methods can be applied to find the value of η when e^k is given, i.e. Newton method method [46]. Because the problem is a one-dimensional problem, we use the bisection method [46] due to its low complexity. Hence, the upper and lower bound of η should be found first to determine the search interval.

If only the first term in (5.28) is considered, an lower bound of e^k can be obtained by neglecting the higher order terms, i.e.

$$e^k_{low} = \frac{2^{k-1}}{\sqrt{M}} \left(1 - \frac{\sqrt{\frac{3\log_2 M\eta}{2(M-1)}}}{\sqrt{\frac{3\log_2 M\eta}{2(M-1)} + 1}} \right) \tag{5.29}$$

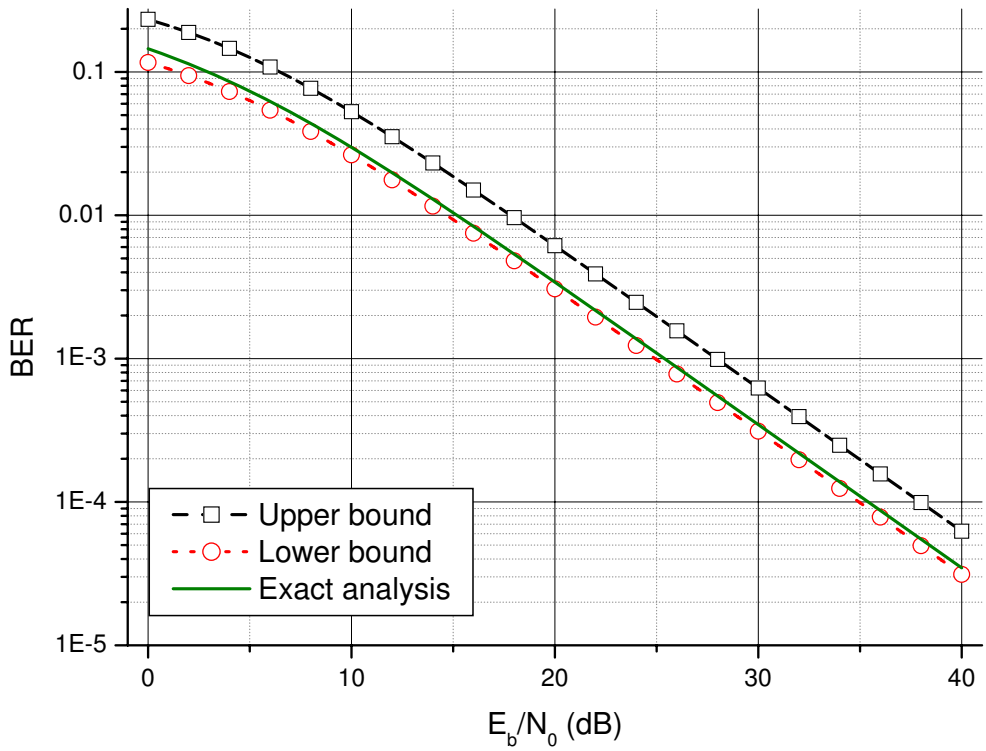
and its inverse function is

$$\eta = \left(1 - \frac{\sqrt{M} e^k_{low}}{2^{k-1}} \right)^2 \frac{2M-2}{3\left(2\sqrt{M}\frac{e^k_{low}}{2^{k-1}} - M\left(\frac{e^k_{low}}{2^{k-1}}\right)^2 \right) \log_2 M}$$

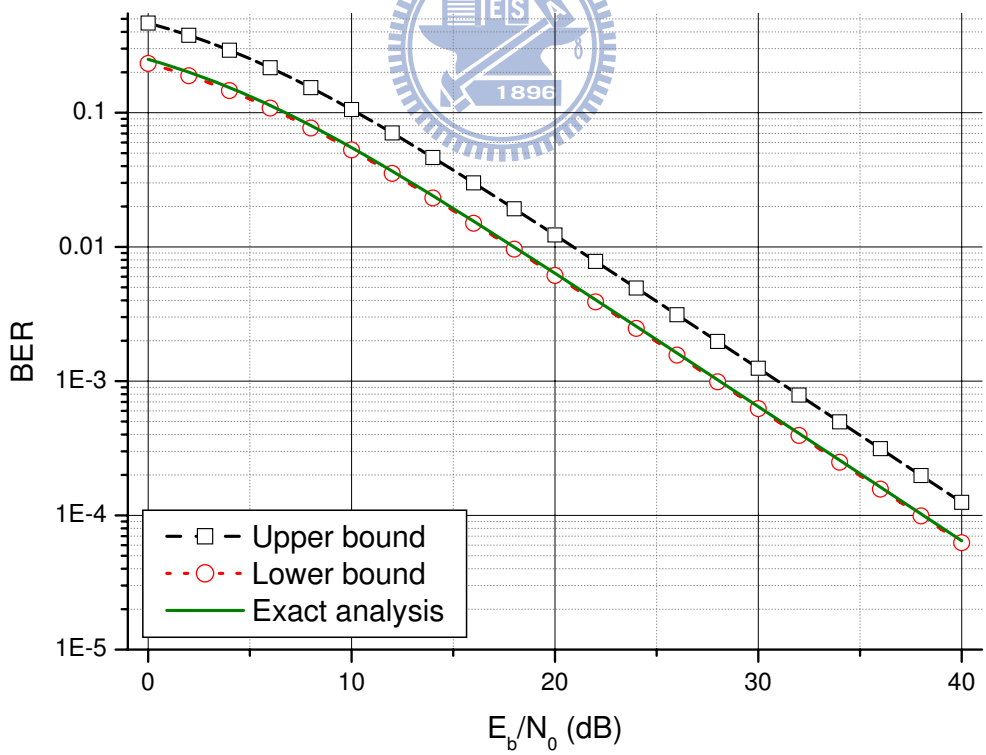
Hence, given e^k, we can find a lower bound of η:

$$\eta_{low} = \left(1 - \frac{\sqrt{M} e^k}{2^{k-1}} \right)^2 \frac{2M-2}{3\left(2\sqrt{M}\frac{e^k}{2^{k-1}} - M\left(\frac{e^k}{2^{k-1}}\right)^2 \right) \log_2 M} \tag{5.30}$$

However, the upper bound with simple form is not easy to find if it is possible. Instead of finding a upper bound, we propose a simple bound which is a upper bound



(a) The exact analysis, upper and lower bound of the first bit error rate



(b) The exact analysis, upper and lower bound of the second bit error rate

Figure 5.5: 16QAM error rate in Rayleigh fading with $\alpha = 2$.

within the range of interest, i.e.

$$e^k_{upper}(\alpha) = \alpha \frac{2^{k-1}}{\sqrt{M}} \left(1 - \frac{\sqrt{\frac{3\log_2 M\eta}{2(M-1)}}}{\sqrt{\frac{3\log_2 M\eta}{2(M-1)} + 1}} \right) \tag{5.31}$$

where $\alpha > 1$. As can be seen, we introduce a parameter α into (5.29) to get the upper bound (5.31). The parameter α depends on the range of interest and should be determined offline. For 16-QAM modulation in Rayleigh fading channel, Fig. 5.5 shows that $\alpha = 2$ is a good choice from $E_b/N_0 = 0$ to 40 (dB). In this case, we can find the upper bound of η given e^k

$$\eta_{upper} = \left(1 - \frac{\sqrt{M} e^k}{\alpha 2^{k-1}} \right)^2 \frac{2M-2}{3\left(2\sqrt{M}\frac{e^k}{\alpha 2^{k-1}} - M(\frac{e^k}{\alpha 2^{k-1}})^2\right)\log_2 M} \tag{5.32}$$

With the lower bound (5.29) and upper bound (5.31), the bisection method for finding η is described in Table 5.2.

Table 5.2: The bisection method for finding $\eta = \frac{E_b}{N_0}$.

Input:	e^k, tolerance δ, and α
1:	Compute the lower η_{low} and upper bound η_{upper} of η by (5.29) and (5.31), respectively
2:	Set η_t=$(\eta_{low} + \eta_{upper})/2$ and compute the bit error rate e^k_t given η_t by (5.28).
3:	if $e^k_t > e^k$
4:	$\eta_{low} = \eta_t$
5:	else
6:	$\eta_{upper} = \eta_t$
7:	endif
8:	If $\frac{(\eta_{upper} - \eta_{low})}{\eta_{low}} > \delta$
9:	goto 2.
10:	endif
11:	Compute $\eta = \frac{\eta_{upper} + \eta_{low}}{2}$
Output:	η.

After finding the E_b/N_0, we can rescale it and find the estimate $\widehat{e}^k$ by (5.28), again. Hence, the fourth step is complete. The flowchart of the whole estimation process is

given in Table 5.3. We end up this section by emphasizing the possibility of further performance improvement using multiple estimates, as stated in remark 5 in Chapter 4. However, unlike the binary modulation case, multiple estimates for M-QAM can occurs naturally. For M-QAM, the kth BER of a link in the in-phase part is the same as that of a link in the quadrature-phase part. Hence, we can improve the performance by average these two estimates. Moreover, in our estimator, it is implicitly assumed that we consider the three links with the same bit-level elements for convenience of representation. In fact, we can estimate ERs by considering the three links with different bit-level elements to generate multiple estimates. For example, we may consider a basic nonlinear system involving e_1^1, e_2^2, e_3^1 to obtain the estimates $\widehat{e}_1^1$, $\widehat{e}_2^2$, $\widehat{e}_3^1$.

Table 5.3: A noise-enhanced blind ER estimation algorithm for high order modulation in sensor network.

Input:	Received samples, $\mathbf{y}$, noise variance, σ_d^2, δ, α and scaling factor values, $a_1^{(w)}$, $a_2^{(w)}$ and $a_3^{(w)}$.
1:	Add noise-enhanced complex zero-mean Gaussian samples with variances $a_i^{(w)} - 1$, $i = 1, 2, 3$ to the received samples.
2:	Detect the signals $\widehat{y}_i^{(w)}$ and transform them into bit-level.
3:	Compute SMPs for all link pairs by (5.17).
4:	Compute $\widehat{e}_i^{(w)}$, $i = 1, 2, 3$ via (5.18) using the SMPs obtained in 3.
5:	Find the $\eta_i^{(w)}$ through the bisection method in Table 5.2 with $\widehat{e}_i^{(w)}$, $i = 1, 2, 3$.
6:	Compute $\eta_i = a_i^{(w)} \eta_i^w$, $i = 1, 2, 3$.
7:	Compute $\widehat{e}_i$ via (5.28), $i = 1, 2, 3$.
Output:	$\widehat{e}_1$, $\widehat{e}_2$ and $\widehat{e}_3$.

Similarly, the noise-enhanced estimation also consists of four steps. First, generate noise-enhanced receive signals. Then, demodulate the signals. Because we are interested in the symbol ERs estimation, we do not have to transform the detected signal into bit-level one. Then, estimate the symbol ERs $\widetilde{e}_i$. Finally, transform the estimates $\widetilde{e}_i$ into the true/original ones $\widehat{e}_i$. To find the transformation from $\widetilde{e}_i$ to $\widehat{e}_i$, we need the formula of symbol ER. In this dissertation, we focus on MFSK (M-ary frequency-shift keying)

modulation scheme.

As in the case of M-QAM modulation scheme, it is sufficient to find a way of obtaining SNR given a symbol ER for the ER transformation. In [37], it is shown that the symbol ER for MFSK modulation scheme is

$$e = \sum_{m=1}^{M-1} (-1)^{m+1} \binom{M-1}{m} \frac{1}{1+m(1+\eta)} \tag{5.33}$$

where η is the average signal-to-noise power ratio (SNR). Notice that (5.33) is a rational function of η. Hence, we can obtain a polynomial equation of e with one variable η. Solving the polynomial equation yields the value of η given e. For example, the symbol ER for 4-FSK modulation scheme is

$$\begin{aligned} e &= \sum_{m=1}^{3} (-1)^{m+1} \binom{3-1}{m} \frac{1}{1+m(1+\eta)} \\ \Rightarrow & 6e\eta^3 + (29e-11)\eta^2 + (46e-28)\eta + 24e - 18 = 0 \end{aligned}$$

We can solve the polynomial equation either by the bisection method or by Newton method [46].

5.6 Simulation results

For convenience of reference, we refer to the detector (5.1), (5.8), and (5.10) as the symbol-level, bit-level, and nonsymmetric bit-level detector, respectively. Moreover, the bit-level detector with the proposed ER estimator is denoted as ER-based bit-level detector. The simulated performance curve of ER-based bit-level detector is obtained by sequentially applying the proposed method. The simulations are terminated when 500 errors occurs.

First, Fig. 5.6 illustrates the performance curves with three links and QPSK modulation. Clearly, the proposed bit-level detector and symbol-level detector have no difference in BER, showing that the approximation induces negligible performance loss. In addition, the ER-based bit-level detector has the identical performance as bit-level detector.

This indicates that the ER-based bit-level detector can be applied in practice with low computational complexity but high performance.

To investigate the effect of higher order modulation and orthogonal one, we also consider 16-QAM modulation and 4-FSK in three links sensor network in the same figure (Fig. 5.6). Although there is a performance gap between the performance curve of nonsymmetric bit-level and of symbol-level detector for 16-QAM, this gap is smaller than 0.5 dB and is insignificant. Notice that the bit-level and nonsymmetric bit-level detector have no observable difference in performance. Hence, we only consider the bit-level detector with blind ER estimator (ER-based bit-level detector). Importantly, the proposed ER-based bit-level detector for 16-QAM and ER-based detector for 4-FSK have insignificant performance loss with low complexity; hence, they can be implemented in wireless sensor networks.

We consider a 3-link wireless sensor network employing 16-QAM modulation and 4-FSK in Fig. 5.7 and 5.8, respectively. To compare the performance under different link qualities, we adopt the normalized MSEE as metric. Normalized MSEE is defined by $\text{MSEE}_j^k/(e_j^k)^2$, where j and k denote the jth link and kth bit, respectively. As can be seen, the higher the link quality (E_b/N_0), the worse the performance, which is the motivation of the proposed noise-enhanced estimator.Although the analytic results (5.27) is almost identical to the simulation results for 4-FSK, the analytic results (5.19) for 16-QAM is a lower bound performance because of the model mismatch. Nevertheless, the gap between the analytic and simulation results is small.

Fig. 5.9 shows the stochastic resonance phenomenon where we consider a 3-link sensor network with 16-QAM modulation. The MSEE reduction ratio (γ) performance is defined by $\frac{MSEE_n}{MSEE_o}$, where $MSEE_n$ and $MSEE_o$ are the mean square estimation error of noise-enhanced and direct estimator (5.18). As shown in [42], the MSEE reduction ratio γ provides the insight about the reduction ratio of requried sample size for a given normalized MSEE. The smaller the γ is, the smaller the required sample size for noise-

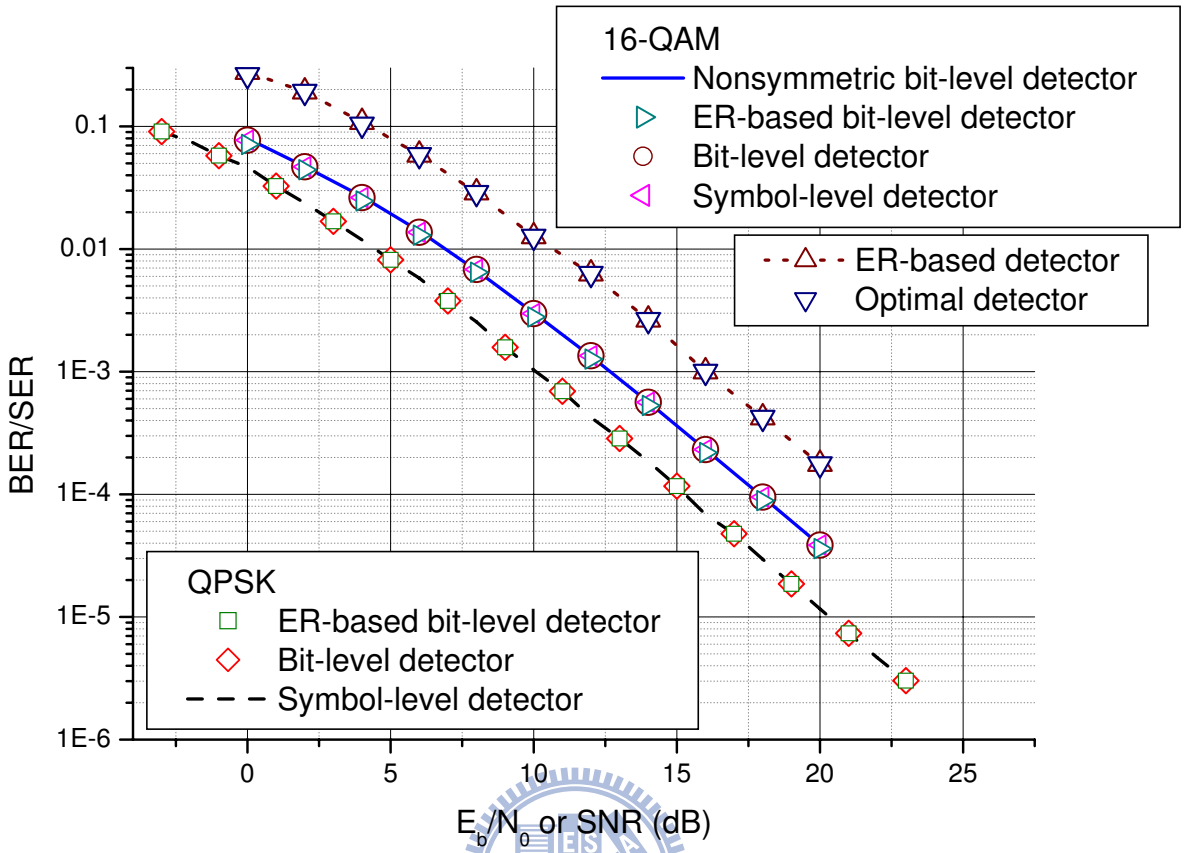


Figure 5.6: Bit (symbol) error rate performance of various detectors with QPSK/16-QAM (MFSK) modulation and Gray mapping labelling. The qualities of the three links for QPSK modulation are denoted by $\left(\frac{E_b}{N_0}, \frac{E_b}{N_0}+1, \frac{E_b}{N_0}+2\right)$ in dB. Similarly, $\left(\frac{E_b}{N_0}, \frac{E_b}{N_0}+2, \frac{E_b}{N_0}+4\right)$ and (SNR,SNR+3,SNR+6) are the link qualities for QAM and 4-FSK, respectively.

enhanced estimator is. We plots the MSEE reduction ratio as a function of the first bit's $\frac{E_b}{N_0}^{(w)}$, the $\frac{E_b}{N_0}$ after noise injection, at sample size $N = 5000$ in Fig. 5.9. Noise are also injected into the other two links such that all links have the same $\frac{E_b}{N_0}^{(w)}$. These two curves reveal that the MSEE performance is improved by injecting proper noise power into the received samples and there is an optimal injected noise power that achieves the maximum MSEE improvement. For example, the MSEE reduction of the e_3^1 is about 0.002 when $\frac{E_b}{N_0}^{(w)} = 10$ dB. This indicates that the required number of samples for a given precision can be reduced more than 500 times [40]. We also notice that the improvement is more impressive when the true ER becomes smaller, which is consistent

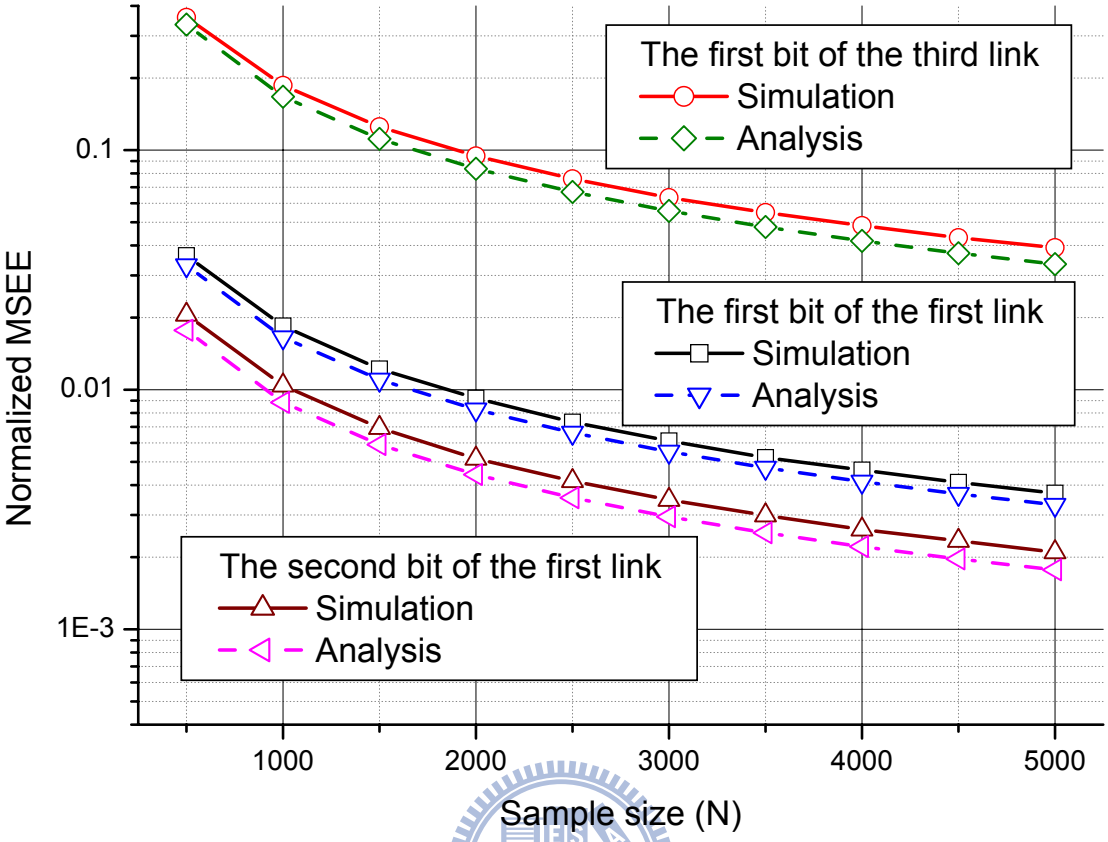


Figure 5.7: Normalized MSEE performance of blind bit-level estimator in a 16-QAM-based three link wireless sensor network. The qualities of these three links are 10, 15, and 20 (dB), respectively.

with what the importance sampling theory has predicted. Finally, we can observe that the noise benefit interval (NBI) is quite wide. For instance, the NBI is from 0 to 20 dB for e_1^1 if we define the noise benefit interval as the interval that $\gamma < 0.1$. This observation implies that the noise-enhanced estimator is robust to the noise variance estimate error.

Actually, the MSEE reduction ratio behavior depends on the noise injection strategy. To show the possibility of existence of stochastic resonance phenomenon for other strategy, we consider a strategy keeping the difference of these three links' qualities. This strategy is equivalent to add the noise with the same noise intensity to all three links. The simulation results are shown in Fig. 5.10 for 16-QAM and 5.11 for 4-FSK. As can be seen, we have the similar conclusions: the existence of stochastic resonance

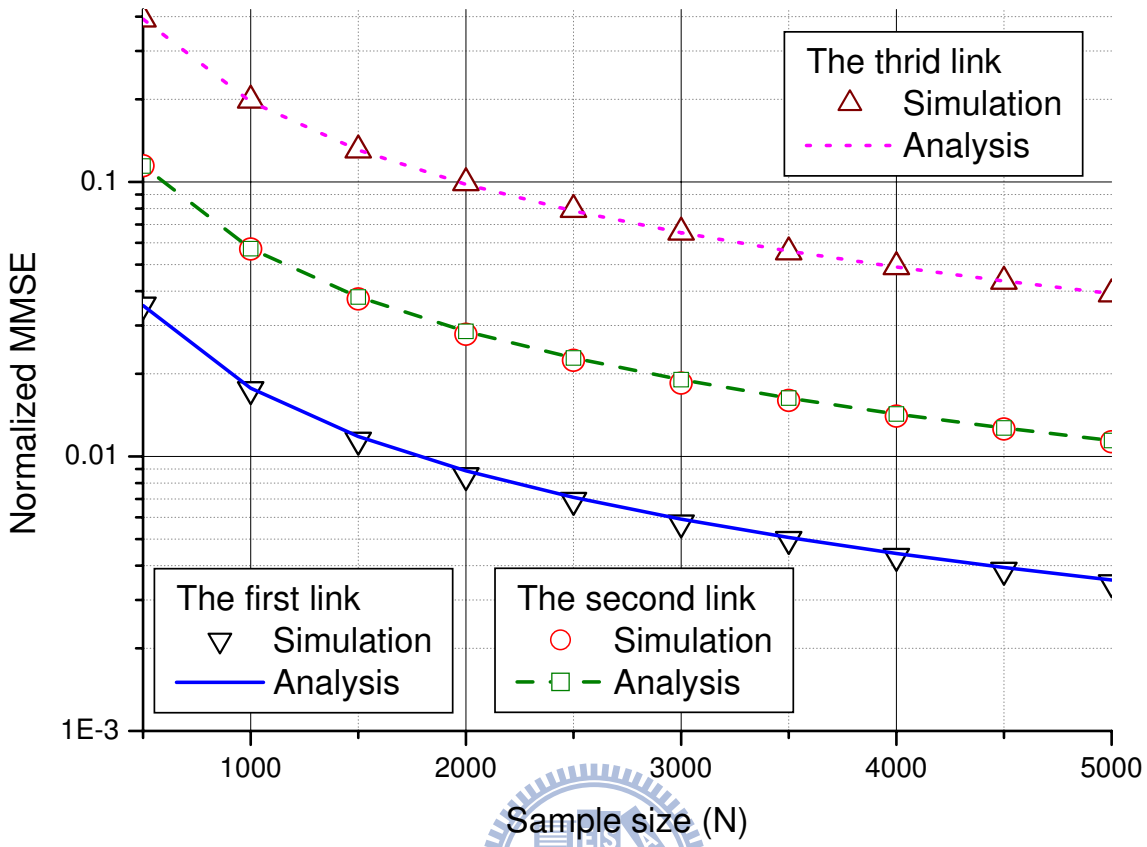


Figure 5.8: Normalized MSEE performance of blind symbol ER estimator in a 4-FSK-based three link wireless sensor network. The qualities of these three links are 15, 20, and 25 (dB), respectively.

phenomenon and robustness to the estimation error of noise variance.

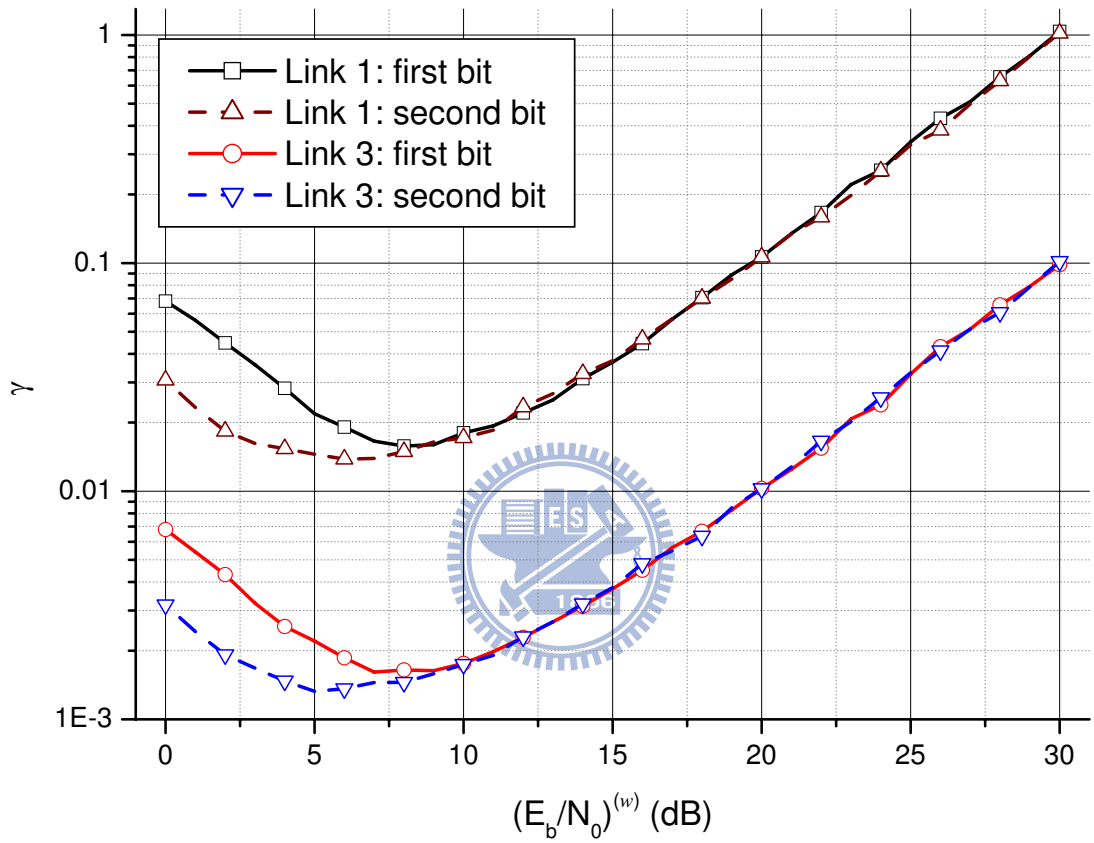


Figure 5.9: MSEE reduction ratio behavior of the noise-enhanced estimator in three links sensor network with 16-QAM modulation and Rayleigh fading channel. The qualities $\left(\frac{E_b}{N_0}\right)$ of these three links are 30, 35, and 40 (dB). Noise are injected into the three links such that all links have the same $\frac{E_b}{N_0}^{(w)}$.

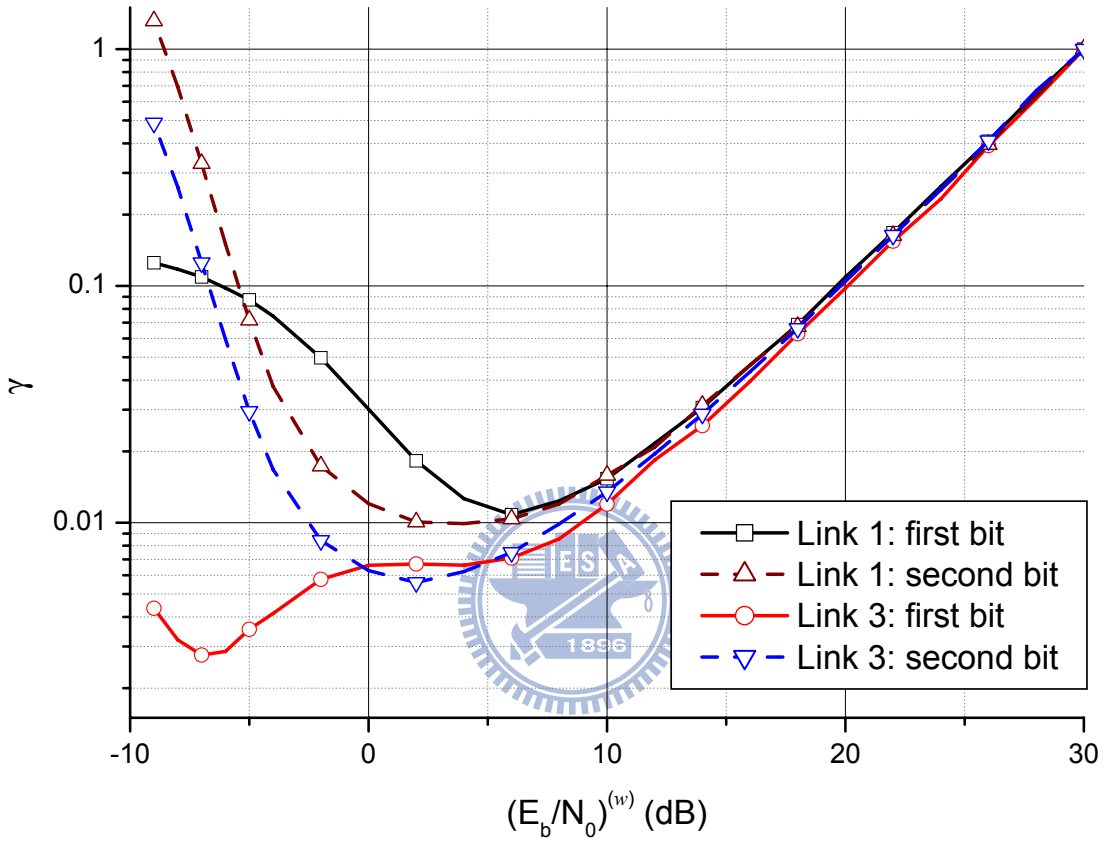


Figure 5.10: MSEE reduction ratio behavior of the noise-enhanced estimator in three links sensor network with 16-QAM modulation and Rayleigh fading channel. The qualities $\left(\frac{E_b}{N_0}\right)$ of these three links are 30, 35, and 40 (dB). We inject noise into the first link $\left(\left(\frac{E_b}{N_0}\right) = 30 \text{ (dB)}\right)$ such that the quality of the link is $\frac{E_b}{N_0}^{(w)}$. We keep the difference of the quality of these three links. That is, the other two links with noise injection have the quality $\frac{E_b}{N_0}^{(w)} + 5$ and $\frac{E_b}{N_0}^{(w)} + 10$ (dB).

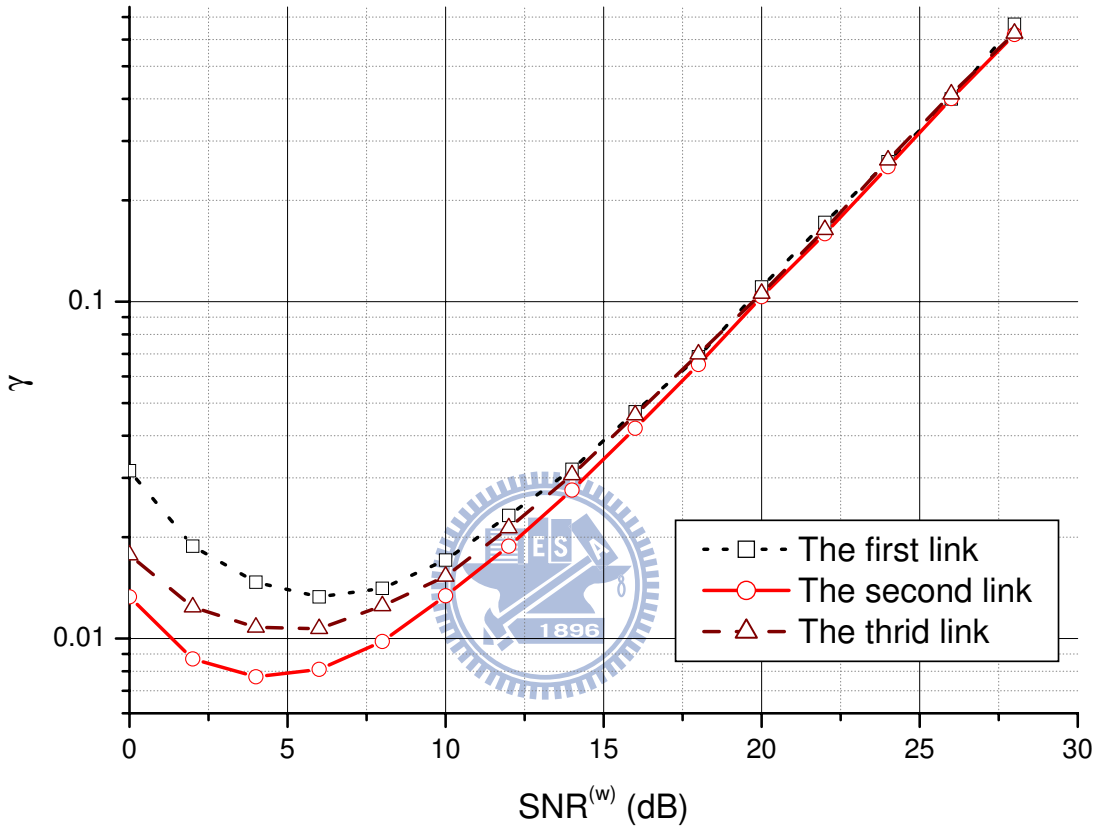


Figure 5.11: MSEE reduction ratio behavior of the noise-enhanced estimator in three links sensor network with 4-FSK modulation and Rayleigh fading channel. The qualities $\left(\frac{E_b}{N_0}\right)$ of these three links are 30, 35, and 40 (dB). We inject noise into the first link $\left(\left(\frac{E_b}{N_0}\right) = 30 \text{ (dB)}\right)$ such that the quality of the link is $\frac{E_b}{N_0}^{(w)}$. We keep the difference of the quality of these three links. That is, the other two links with noise injection have the quality $\frac{E_b}{N_0}^{(w)} + 5$ and $\frac{E_b}{N_0}^{(w)} + 10$ (dB).

Chapter 6

Conclusions and future work

Blind ER estimation is needed for data detection or fusion in wireless relay networks which include sensor networks and cooperative communication networks as subclasses. Earlier proposals suffer from slow convergence and were unable to estimate the ERs of hidden SR links. Some ambiguity issues associated with cascaded links and the lack of enough links remain unsolved before.

In this dissertation, we first propose noise-enhanced blind ER estimators for binary modulation based wireless relay networks. Noise-enhancement manifests itself in three aspects. Firstly, noise is added to the received samples to create VLs for removing the CSI requirement and resolving the ambiguity associated with an underdetermined system and that due to the symmetric nature of a cascaded link. Secondly, multiple noise-injected VLs are used to reduce the estimation variance and the number of relays needed for estimating ERs. Thirdly, inspired by the IS theory used in computer simulation based ER estimation, noise with proper power is inserted to improve the ER estimator's convergence performance. The MSEE performance of some special networks is analyzed and both analysis and simulations show that the IS inspired estimator exhibits the so-called stochastic resonance phenomenon which amounts to the effect that injecting noise with a proper power helps improving an estimator's performance and there exists an optimal injected noise power that offers the best MSEE improvement. Simulation results indicate that the performance of the ML detector using our estimators is very

close to that of the ideal ML detector which knows SR link's ER perfectly. Moreover, the Monte-Carlo based ISI approach is capable of bringing about several orders of MSEE reduction.

For networks using high order modulation, we find that the optimal symbol-level detector is not feasible because of the prohibitive computing load. However, if orthogonal signals such as MFSK is used, the optimal symbol-level detector can be greatly simplified. For general high order modulation based networks, we derive a bit-level detector which requires much smaller number of ER parameters. We propose ERs estimators for the latter two cases. These estimators require low complexity while the existing ERs estimator has to solve a large-scale optimization problem. Simulation results shows that our symbol/bit-level fusion rule using the proposed ER estimator render small performance loss (less than 0.5 dB). We also propose noise-enhanced blind ER estimators to improve the MSEE performance for the nonbinary modulation based networks. As expected, simulation results demonstrate that injecting noise with proper power does bring about significant performance improvement and an optimal injected noise power level can be found.

Our work can be extended to deal with applications in distributed source coding [47]. As the noise-enhanced estimator achieves its best performance only if the optimal injected power is known, a more efficient way to find this power level for different scenarios is needed. In Chapter 5, we have neglected the band-limiting effect and the fusion center receives complete soft outputs [48]. There are cases when only the quantized measurements are available at either the sensor nodes or the fusion center and it is desired to have a distributed estimation algorithm [49]. These are some of topics that calls for further investigations. Finally, we also believe that there are many interesting stochastic resonance phenomenon in nonlinear communication systems and networks that deserve much more research efforts to explore their applications.

Appendix A

A.1 Derivation of (2.6)

Given the complete CSI, $\{h_{sd}, h_{rd}, \sigma_d^2, e_{sr}\} = I_{csi}$ and unit transmit powers, $P_s = P_r = 1$, the conditional joint pdf of the matched filter outputs, y_{sd}, y_{rd}, can be represented as

$$\begin{aligned} f(y_{sd}, y_{rd}|I_{csi}) &= C\exp\left(-\frac{\|y_{sd}-h_{sd}\|^2}{\sigma_d^2}\right)\left[(1-e_{sr})\exp\left(-\frac{\|y_{rd}-h_{rd}\|^2}{\sigma_d^2}\right)\right. \\ &\quad \left.+e_{sr}\exp\left(-\frac{\|y_{rd}+h_{rd}\|^2}{\sigma_d^2}\right)\right]+C\exp\left(-\frac{\|y_{sd}+h_{sd}\|^2}{\sigma_d^2}\right) \\ &\quad \times\left[(1-e_{sr})\exp\left(-\frac{\|y_{rd}+h_{rd}\|^2}{\sigma_d^2}\right)+e_{sr}\exp\left(-\frac{\|y_{rd}-h_{rd}\|^2}{\sigma_d^2}\right)\right] \end{aligned}$$

where C is a normalization constant. By removing the terms independent of e_{sr}, we obtain

$$\begin{aligned} f(y_{sd}, y_{rd}|I_{csi}) &\propto \exp(q_0/2+q_1/2)+e_{sr}\exp(q_0/2)\left[-\exp(q_1/2)+\exp(-q_1/2)\right] \\ &\quad +\exp(-q_0/2-q_1/2)+e_{sr}\exp(-q_0/2)\left[-\exp(-q_1/2)+\exp(q_1/2)\right] \\ &\propto \cosh\left(\frac{q_0+q_1}{2}\right)-2e_{sr}\sinh(q_0/2)\sinh(q_1/2) \end{aligned}$$

where $q_0 \stackrel{def}{=} 4Re\{y_{sd}^* h_{sd}\}/\sigma_d^2$ and $q_1 \stackrel{def}{=} 4Re\{y_{rd}^* h_{rd}\}/\sigma_d^2$. Given N independent sample pairs, $\left\{\left(q_0^{(i)}, q_1^{(i)}\right)\right\}_{i=1}^N$, the ML estimator for e_{sr} is given by

$$\begin{aligned} \hat{e}_{sr} &= \arg\max_{0\le e_{sr}<0.5} \log f(\{y_{sd}[i]\}_{i=1}^N, \{y_{rd}[i]\}_{i=1}^N|\mathrm{I}_{csi}) \\ &= \arg\max_{0\le e_{sr}<0.5} \sum_{i=1}^N \log\left[\cosh\left(\frac{q_0^{(i)}+q_1^{(i)}}{2}\right)-2\sinh\left(\frac{q_0^{(i)}}{2}\right)\sinh\left(\frac{q_1^{(i)}}{2}\right)e_{sr}\right]. \end{aligned}$$

A.2 Derivations of (2.8) and (2.10)

To show

$$e_{sr} = \frac{1 - e_{sd} - e_{rd} + 2e_{sd}e_{rd} - p}{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}} \stackrel{def}{=} e_{sr}(p) \tag{A.1}$$

one first notices that

$$p = \Pr(\widehat{y}_{sd} = \widehat{y}_{rd}) = \Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1) + \Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 0)$$

Let S be the random variable representing the binary source output. The first term on the right hand side is

$$\Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1) = \frac{1}{2}\Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1|S = 0) + \frac{1}{2}\Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1|S = 1)$$

Given S, $\widehat{y}_{sd}$ is independent of $\widehat{y}_{rd}$, hence

$$\begin{aligned}\Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1|S = 0) &= \Pr(\widehat{y}_{sd} = 1|S = 0)\Pr(\widehat{y}_{rd} = 1|S = 0) \\ &= e_{sd}\left(e_{sr} + e_{rd} - 2e_{sr}e_{rd}\right)\end{aligned}$$

Similarly,

$$\Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1|S = 1) = (1 - e_{sd})\left(1 - e_{sr} - e_{rd} + 2e_{sr}e_{rd}\right)$$

One then has

$$\Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1) = \frac{1}{2}\left[1 - e_{sd} + 2e_{sd}e_{rd} - e_{rd} - e_{sr}(1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd})\right]$$

It is verifiable that

$$\Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 0) = \Pr(\widehat{y}_{sd} = \widehat{y}_{rd} = 1)$$

and therefore

$$\begin{aligned}p &= \Pr(\widehat{y}_{sd} = \widehat{y}_{rd}) \\ &= 1 - e_{sd} + 2e_{sd}e_{rd} - e_{rd} - e_{sr}(1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd})\end{aligned}$$

and

$$e_{sr} = \frac{1 - e_{sd} - e_{rd} + 2e_{sd}e_{rd} - p}{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}}$$

As $I(\widehat{y}_{sd}[i] = \widehat{y}_{rd}[i])$ is Bernoulli distributed, the sample mean estimator

$$\widehat{p}^{(N)} = \sum_{i=1}^{N} \frac{I\left(\widehat{y}_{sd}[i] = \widehat{y}_{rd}[i]\right)}{N}, \tag{A.2}$$

is an uniform minimum variance unbiased estimator if i.i.d samples are received [35].

Using the estimator (A.2) as $\hat{p}$, the method of moments and (A.1) suggests the estimator

$$\widehat{e}_{sr} = \frac{1 - e_{sd} - e_{rd} + 2e_{sd}e_{rd} - \widehat{p}}{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}} = e_{sr}(\hat{p}) \tag{A.3}$$

if both e_{rd} and e_{sd} are known.

A.3 Proof of Lemma 2.1

The joint probability density function (pdf) of the N i.i.d. random variables $W^{(1)}, \cdots, W^{(N)}$ are given by

$$f_{\mathbf{W}}(W[1], \cdots, W[N]; p) = \prod_{i=1}^{N} f(W^{(i)}; p) = p^t(1-p)^{N-t}. \tag{A.4}$$

where $W[i] = I\left(\widehat{y}_{sd}[i] = \widehat{y}_{rd}[i]\right)$, $\sum_{i=1}^{N} I(\widehat{y}_{sd}^{(i)} = \widehat{y}_{rd}^{(i)}) = t$, and the marginal pdf $f(W^{(i)}; p)$ is a Bernoulli distribution with parameter p.

Taking derivative of the function $\ln f_{\mathbf{W}}(W[1], \cdots, W[N]; p)$ with respect to p, we obtain

$$\frac{\partial \ln f_{\mathbf{W}}(W[1], \cdots, W[N]; p)}{\partial p} = \frac{t}{p} - \frac{N-t}{1-p} = \frac{N}{p(1-p)}\left(\frac{t}{N} - p\right). \tag{A.5}$$

In [50], it was shown that an unbiased estimator $\widehat{p}(W[1], \cdots, W[N]) \stackrel{def}{=} \widehat{p}$ for p attains the CRLB for all p if and only if

$$\frac{\partial \ln f_{\mathbf{W}}(W[1], \cdots, W[N]; p)}{\partial p} = C(p)\left[\widehat{p}(W^{(1)}, \cdots, W^{(N)}) - p\right]. \tag{A.6}$$

where $C(p)$ is a function of p and is independent of the observations. Comparing the above two equations, we conclude that $C(p) = \frac{N}{p(1-p)}$ and the estimator $\widehat{p}(W[1], \cdots, W[N]) = \frac{t}{N}$, which is equivalent to (2.9), is efficient. Furthermore, if $\widehat{p}$ is an efficient estimator for p and $h(p)$ is a linear function of p, then $h(\widehat{p})$ is also an efficient estimator for $h(p)$ [51]. Since e_{sr} is linear function of p, we prove the first part of the Lemma.

To prove the second part, we first notice that $\widehat{p}$ is a zero of (A.5) and thus an ML estimator. Now if we let $h(p)$ be $e_{sr}(p)$ defined by (2.8), which is a one-to-one function of p, then the invariance property of ML estimators [50] implies that $e_{sr}(\widehat{p})$ is also an ML estimator for e_{sr}. As we have shown that $\widehat{e}_{sr}$ is unbiased and achieves the CRLB, it is an UMVU estimator. The corresponding CRLB is given by [50],

$$\text{CRLB} = \frac{\left(\frac{\partial e_{sr}(p)}{\partial p}\right)^2}{C(p)} = \frac{p(1-p)}{N(1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd})^2}. \tag{A.7}$$

A.4 Proof of Lemma 2.2

Without the soft-limiting effect, which can be ignored for reasonable large sample size, the convergence speed of the estimator $\widehat{p}$ can be estimated by Chernoff's inequality [52]

$$\Pr(|\widehat{p}-p| \geq \varepsilon_1 p/N) \quad \leq \quad 2\exp\left(-\min(\varepsilon_1^2/4, \varepsilon_1/2)p\right)$$

where N is the sample size and ε_1 is an arbitrary positive number. Using the following identity

$$\frac{\widehat{p}-p}{1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}} \quad = \quad e_{sr} - \widehat{e}_{sr}$$

and setting $C_1 = 1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}$ and $\varepsilon = \frac{\varepsilon_1 p}{NC_1} > 0$, we then obtain (2.13).

A.5 Derivation of (2.18)

Notice that a basic nonlinear system (2.17) can be represented as

$$\begin{bmatrix} 1-Q_1-Q_0+2Q_0Q_1 \\ 1-Q_1-Q_2+2Q_1Q_2 \\ 1-Q_0-Q_2+2Q_0Q_2 \end{bmatrix} = \begin{bmatrix} p_{01} \\ p_{12} \\ p_{02} \end{bmatrix} \tag{A.8}$$

where $\widehat{Q}_0 = \widehat{e}_{sd}$, $p_{01} = p_{sr_1}$, and $p_{02} = p_{sr_2}$.

Representing Q_1 and Q_2 in terms of Q_0 through the first and third equations in (A.8) yields

$$Q_1 = (1 - Q_0 - p_{01})/(1 - 2Q_0)$$

$$Q_2 = (1 - Q_0 - p_{02})/(1 - 2Q_0)$$

Substituting these two equation into the second equation in (A.8), we have

$$1 - \frac{1 - Q_0 - p_{01}}{1 - 2Q_0} - \frac{1 - Q_0 - p_{02}}{1 - 2Q_0} + 2\frac{1 - Q_0 - p_{01}}{1 - 2Q_0}\frac{1 - Q_0 - p_{02}}{1 - 2Q_0} = p_{12}$$

$$\Rightarrow (1 - p_{12})(1 - 2Q_0)^2 - (2 - 2Q_0 - p_{01} - p_{02})(1 - 2Q_0) + 2(1 - Q_0 - p_{01})(1 - Q_0 - p_{02}) = 0$$

$$\Rightarrow (1/2 - b/2)(2a)^2 + (-2a - c/2 - d/2)2a + 2(-a - c/2)(-a - d/2) = 0$$

$$\Rightarrow -2ba^2 + (-c - d + c + d)a + cd/2 = 0$$

$$\Rightarrow a = \pm\sqrt{\frac{cd}{4b}}$$

where $a = (Q_0 - 1/2)$, $b = 2p_{12} - 1$, $c = 2p_{01} - 1$, and $d = 2p_{02} - 1$. Because $Q_0 < 1/2$, we have

$$Q_0 = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{cd}{b}} = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{(2p_{01} - 1)(2p_{02} - 1)}{2p_{12} - 1}}$$

The symmetric nature of (A.8) gives

$$Q_i = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{(2p_{ij} - 1)(2p_{ik} - 1)}{2p_{jk} - 1}}$$

where $i, j, k \in \{0, 1, 2\}$.

Appendix B

B.1 Derivation of (4.16)

To derive (4.16), notice that we have (3.1)

$$e = \frac{1}{2}\left(1 - \sqrt{\frac{\overline{\gamma}}{\overline{\gamma}+1}}\right)$$

and inverse formula (3.2)

$$\overline{\gamma}^{(w)} = \frac{(2e^{(w)} - 1)^2}{1 - (2e^{(}(w)) - 1)^2}$$

Substitute (3.2) into (3.1) and $\overline{\gamma} = a^{(w)}\overline{\gamma}^{(w)}$, we have

$$\begin{aligned} e &= \frac{1}{2}\left(1 - \sqrt{\frac{\overline{\gamma}}{\overline{\gamma}+1}}\right) \\ &= \frac{1}{2}\left(1 - \sqrt{\frac{a^{(w)}\overline{\gamma}^{(w)}}{a^{(w)}\overline{\gamma}^{(w)}+1}}\right) \\ &= \frac{1}{2}\left(1 - \sqrt{\frac{a^{(w)}(1-2e^{(w)})^2}{1-(1-2e^{(}(w)))^2 + a^{(w)}(1-2e^{(w)})^2}}\right) \end{aligned}$$

which is (4.16).

B.2 Proof of Lemma 4.1

Let $\widetilde{p}^{(w)}$ be the average count based estimate of $p^{(w)}$–the SMP of the noise-injected SD and RD link outputs–then we have, from (2.8), the conversion rule

$$\widehat{p}^{(w)} = \mathcal{D}_o + \frac{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}}{1 - 2e_{sd}^{(w)} - 2e_{rd}^{(w)} + 4e_{sd}^{(w)}e_{rd}^{(w)}}\widetilde{p}^{(w)} \tag{B.1}$$

where

$$\begin{aligned}\mathcal{D}_o =& \frac{(1-e_{sd}-e_{rd}+2e_{sd}e_{rd})(1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)})}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}} \\ &- \frac{(1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd})(1-e_{sd}^{(w)}-e_{rd}^{(w)}+2e_{sd}^{(w)}e_{rd}^{(w)})}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}\end{aligned} \tag{B.2}$$

As $\widehat{p}^{(w)}$ is a linear function of $\widetilde{p}^{(w)}$, the ML estimate of $p^{(w)}$, it is an ML estimator of p. Furthermore, $\widetilde{p}^{(w)}$ is a sample mean estimator; its variance is equal to $\mathrm{var}\left[\widetilde{p}^{(w)}\right] = \frac{p^{(w)}(1-p^{(w)})}{N}$. Similarly, the variances of $\widehat{p}^{(w)}$ and $\widehat{p}$ are respectively given by

$$\begin{aligned}\mathrm{Var}\left[\widehat{p}^{(w)}\right] =& \frac{p^{(w)}(1-p^{(w)})}{N}\left(\frac{1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}\right)^2 \\ \mathrm{Var}\left[\widehat{p}\right] =& \frac{p(1-p)}{N}\end{aligned}$$

Invoking the inequalities, $0 \le p < p^{(w)} \le 0.5$ or $1 \ge p > p^{(w)} \ge 0.5$, $e_{sd} \le e_{sd}^{(w)}$ and $e_{rd} \le e_{rd}^{(w)}$, we have $p(1-p) \le p^{(w)}(1-p^{(w)})$ and $\left(\frac{1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}\right)^2 \ge 1$. Hence

$$\begin{aligned}\mathrm{Var}\left[\widehat{p}\right] =& \frac{p(1-p)}{N} \\ \le& \left(\frac{1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}\right)^2 \frac{p^{(w)}(1-p^{(w)})}{N} \\ =& \mathrm{Var}\left[\widehat{p}^{(w)}\right]\end{aligned}$$

In other words, as far as estimating p is concerned, the noise-injection method does not help.

B.3 Proof of (4.20)

Following [53], we have the approximation for MSEE reduction ratio

$$\gamma \approx \frac{\int_0^\infty f(y_{rd})dy_{rd}}{\int_0^\infty W(y_{rd})f(y_{rd})dy_{rd}}, W(y_{rd}) \triangleq \frac{f(y_{rd})}{f^*(y_{rd})} \tag{B.3}$$

where $f^*(y_{rd})$ and $f(y_{rd})$ are Gaussian pdf's with the same mean $\sqrt{P_r|h_{rd}|^2}$ but distinct variances $a_{rd}^{(w)}\sigma_d^2$, σ_d^2, respectively.

After some calculations, we have

$$\int_0^\infty W(y_{rd})f(y_{rd})dy_{rd} = \frac{a_{rd}^{(w)}}{\sqrt{2a_{rd}^{(w)}-1}} Q\left(\sqrt{\frac{(2a_{rd}^{(w)}-1)P_r|h_{rd}|^2}{a\sigma_d^2}}\right) \tag{B.4}$$

Since $Q(y) \approx \frac{\exp(-y^2/2)}{y\sqrt{2\pi}}$, for large y, we obtain

$$\gamma \approx \frac{2a_{rd}^{(w)}-1}{\left(\sqrt{a_{rd}^{(w)}}\right)^3} \exp\left[-\frac{(1-a_{rd}^{(w)})P_r|h_{rd}|^2}{2a_{rd}^{(w)}\sigma_d^2}\right] \tag{B.5}$$

The approximation, $2a_{rd}^{(w)} - 1 \approx 2a_{rd}^{(w)}$, yields

$$\gamma \approx \frac{2}{\sqrt{a_{rd}^{(w)}}} \exp\left[-\frac{(1-a_{rd}^{(w)})P_r|h_{rd}|^2}{2a_{rd}^{(w)}\sigma_d^2}\right] \tag{B.6}$$

which is maximized when $a_{rd}^{(w)} = P_r|h_{rd}|^2/\sigma_d^2$.

B.4 Proof of Theorem 4.2

The basic idea behind the derivation is the following theorem, called the Delta method.

Theorem B.1. (The Delta method) Suppose that $Y_n = (Y_{n1}, \cdots, Y_{nk})$ is a sequence of random vector such that

$$\sqrt{N}(Y_n - E[Y_n]) \sim N(0, \mathbf{C})$$

Let $g_i : \mathbb{R}^k \to \mathbb{R}$ and let

$$\mathbf{J} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_k} \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial y_1} & \cdots & \frac{\partial g_k}{\partial y_k} \end{pmatrix}$$

If $\mathbf{J}$ is not zero at $E[Y_n]$, then

$$\sqrt{N}(g(Y_n) - g(E[Y_n])) \sim N(0, \mathbf{JCJ}^T)$$

Based on this theorem, we need to evaluate the variance of $\widehat{e}^{(w)}$ and then the derivative of the transform (4.16). It can be shown that the variance of $\widehat{e}^{(w)}$ is

$$Var[\widehat{e}^{(w)}] = \frac{e^{(w)}(1-e^{(w)})}{N}$$

By (3.1), we have

$$\begin{aligned} Var[\widehat{e}^{(w)}] &= \frac{\left(1-\sqrt{\frac{\overline{\gamma}}{\overline{\gamma}+a^{(w)}}}\right)\left(1+\sqrt{\frac{\overline{\gamma}}{\overline{\gamma}+a^{(w)}}}\right)}{4N} \\ &= \frac{a^{(w)}}{4N(\overline{\gamma}+a^{(w)})} \end{aligned}$$

where $\overline{\gamma} = \text{SNR}$. Moreover, the derivative of the transform (4.16) is

$$\frac{a^{(w)}}{[1+(a^{(w)}-1)(1-2e^{(w)})^2]^3} = \frac{a^{(w)}}{\left[1+(a^{(w)}-1)\frac{\overline{\gamma}}{\overline{\gamma}+a^{(w)}}\right]^3}$$

Hence, the variance (MSEE) of $\widehat{e}$ is

$$\frac{a^{(w)}}{\left[1+(a^{(w)}-1)\frac{\overline{\gamma}}{\overline{\gamma}+a^{(w)}}\right]^3} \frac{a^{(w)}}{4N(\overline{\gamma}+a^{(w)})} = \frac{(\overline{\gamma}+a^{(w)})^2}{4Na^{(w)}(1+\overline{\gamma})^3} \tag{B.7}$$

Hence, the reduction factor is

$$\gamma = \frac{(\overline{\gamma}+a^{(w)})^2}{a^{(w)}(1+\overline{\gamma})^2} \tag{B.8}$$

To find the minimum value of γ, we solve the equation $\frac{\partial\gamma}{\partial a^{(w)}} = 0$ and we have

$$a^{(w)} = \overline{\gamma} \tag{B.9}$$

and

$$\gamma_{min} = \frac{4\overline{\gamma}}{(1+\overline{\gamma})^2} \tag{B.10}$$

Solving the equation $\gamma_{min} = 1$ for SNR gives one repeated root 1. Since $a(w) > 1$, SNR must be greater than 1 in order to get benefit.

B.5 Proof of Lemma 4.3

The analysis presented here follows that of [33] with three major distinctions: (i) we do not use the small ER assumption $e_i^{(w)} \ll 1$, (ii) we have the equal ER constraint, and (iii) we need to consider the ER conversion (4.17).To evaluate the performance, the covariance matrix $\mathbf{C}$ and Jacobian matrix $\mathbf{J}$ must be calculated based on this theorem. The formulas of $\mathbf{C}$ and $\mathbf{J}$ are derived in the remaining of this section.

Assuming independent links, we can show that the covariance matrix of pairwise matching indicators $I(\hat{y}_k[t] = \hat{y}_j[t])$ for the noise injected network is

$$\mathbf{C} = \begin{pmatrix} p_{12}(1-p_{12}) & p_{123} - p_{12}p_{13} & p_{123} - p_{12}p_{23} \\ p_{123} - p_{12}p_{13} & p_{13}(1-p_{13}) & p_{123} - p_{13}p_{23} \\ p_{123} - p_{12}p_{23} & p_{123} - p_{13}p_{23} & p_{23}(1-p_{23}) \end{pmatrix} \tag{B.11}$$

where

$$\begin{aligned} p_{kl} &= \left(1 - e_k^{(w)}\right)\left(1 - e_l^{(w)}\right) + e_k^{(w)} e_l^{(w)} \\ p_{klm} &= \left(1 - e_k^{(w)}\right)\left(1 - e_l^{(w)}\right)\left(1 - e_m^{(w)}\right) + e_k^{(w)} e_l^{(w)} e_m^{(w)} \end{aligned}$$

The three-link network induces the nonlinear system (2.17) whose solution is given by (2.18). It is easier to compute the associated inverse Jacobian matrix for such a nonlinear mapping.

$$\mathbf{J}^{-1} = \begin{pmatrix} \left(2e_2^{(w)} - 1\right) & \left(2e_1^{(w)} - 1\right) & 0 \\ \left(2e_3^{(w)} - 1\right) & 0 & \left(2e_1^{(w)} - 1\right) \\ 0 & \left(2e_3^{(w)} - 1\right) & \left(2e_2^{(w)} - 1\right) \end{pmatrix}$$

Using the constraint $e_1^{(w)} = e_2^{(w)} = e_3^{(w)} = \epsilon$, we obtain the Jacobian and covariance matrices as

$$\mathbf{J} = \frac{1}{2(2\epsilon - 1)} \begin{pmatrix} -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}, \ \mathbf{C} = \begin{pmatrix} z_1 & z_2 & z_2 \\ z_2 & z_1 & z_2 \\ z_2 & z_2 & z_1 \end{pmatrix}$$

where $z_1 = 2\epsilon - 6\epsilon^2 + 8\epsilon^3 - 4\epsilon^4$ and $z_2 = \epsilon - 5\epsilon^2 + 8\epsilon^3 - 4\epsilon^4$.

The covariance matrix for the estimation error is thus given by

$$\mathbf{JCJ}^T = \frac{1}{4(2\epsilon-1)^2}\begin{pmatrix} 3z_1-2z_2 & 2z_2-z_1 & 2z_2-z_1 \\ 2z_2-z_1 & 3z_1-2z_2 & 2z_2-z_1 \\ 2z_2-z_1 & 2z_2-z_1 & 3z_1-2z_2 \end{pmatrix}$$

The variance of the estimator $\widehat{\epsilon}$ can be approximated by

$$\mathrm{Var}\left[\widehat{\epsilon}\right] \approx \frac{3z_1-2z_2}{4(2\epsilon-1)^2 N} = \frac{4\epsilon-8\epsilon^2+8\epsilon^3-4\epsilon^4}{4(2\epsilon-1)^2 N} \tag{B.12}$$

and the variance of $\widehat{e}_i$ can be approximated by ([54] pp. 242)

$$\mathrm{Var}\left[\widehat{e}_i\right] \approx \left(\frac{dg_i(\epsilon)}{d\epsilon}\right)^2 \mathrm{Var}\left[\widehat{\epsilon}\right] = \frac{\left(a_i^{(w)}\right)^2}{\left(a_i^{(w)}+2\epsilon-2a_i^{(w)}\epsilon\right)^4}\frac{\epsilon-2\epsilon^2+2\epsilon^3-\epsilon^4}{(2\epsilon-1)^2 N} \tag{B.13}$$

where $g_i(x) = x/\left(a_i^{(w)}+2x-2a_i^{(w)}x\right)$ is the noncoherent conversion rule.

B.6 Proof of Theorem 4.4

Taking into account the constant noise-injected link ER constraint, we express the average bit error rates for BFSK and DPSK as

$$P_b^{bfsk} = a_i^{(w)}\left(2a_i^{(w)}+\mathrm{SNR}_i\right)^{-1} = \epsilon \tag{B.14}$$

$$P_b^{dpsk} = a_i^{(w)}\left[2\left(a_i^{(w)}+\mathrm{SNR}_i\right)\right]^{-1} = \epsilon \tag{B.15}$$

Using (B.15) and omitting the superscript (w) for simplicity, we obtain

$$a_i + 2\epsilon - 2a_i\epsilon = \frac{a_i\left(\mathrm{SNR}_i+1\right)}{a_i+\mathrm{SNR}_i}, \tag{B.16}$$

which, along with *Lemma 4.3*, gives

$$Var(\widehat{e}_i) \approx \frac{a_i^2(a_i+\mathrm{SNR}_i)^4}{(a_i\mathrm{SNR}_i+a_i)^4}\left[\frac{3a_i^4+12a_i^3\mathrm{SNR}_i+16a_i^2\mathrm{SNR}_i^2+8a_i\mathrm{SNR}_i^3}{16(a_i+\mathrm{SNR}_i)^4 N}\frac{(a_i+\mathrm{SNR}_i)^2}{\mathrm{SNR}_i^2}\right]$$

The MSEE reduction ratio γ is thus given by

$$\gamma = \frac{1}{a_i^2}\left[\frac{3a_i^6+18a_i^5\mathrm{SNR}_i+43a_i^4\mathrm{SNR}_i^2+52a_i^3\mathrm{SNR}_i^3+32a_i^2\mathrm{SNR}_i^4+8a_i\mathrm{SNR}_i^5}{(1+\mathrm{SNR}_i)^2(3+12\mathrm{SNR}_i+16\mathrm{SNR}_i^2+8\mathrm{SNR}_i^3)}\right] \tag{B.17}$$

Using the change of variable $q_i = a_i/\mathrm{SNR}_i$, we find that the condition $\frac{\partial \gamma}{\partial a_i} = 0$ is equivalent to

$$6q_i^5 + 27q_i^4 + 43q_i^3 + 26q_i^2 - 4 = 0 \tag{B.18}$$

Since the only positive rational root is $q_i \approx 0.30855316 \equiv t_1$, (B.15) suggests that we inject noise such that

$$\epsilon = \frac{t_1 \mathrm{SNR}_i}{2(t_1+1)\mathrm{SNR}_i} = 0.1179 \tag{B.19}$$

Furthermore, the minimum achievable MMSE reduction ratio is given by

$$\begin{aligned}
&\gamma_{min} = \gamma|_{a_i = t_1 \mathrm{SNR}_i} \\
=&\frac{78.622\mathrm{SNR}^4}{8\mathrm{SNR}_i^5 + 32\mathrm{SNR}_i^4 + 52\mathrm{SNR}_i^3 + 43\mathrm{SNR}_i^2 + 18\mathrm{SNR}_i + 3} \\
\approx&9.8277\frac{\mathrm{SNR}_i^2}{(1+\mathrm{SNR}_i)^3}
\end{aligned}$$

Solving the equation $\gamma|_{a_i = t_1 \mathrm{SNR}_i} = 1$ for SNR_i gives one positive repeated root 3.24092. Since $a_i > 1$, SNR_i must be greater than $1/t_1 = 3.24093$ in order that noise-injection to become beneficial.

Employing a similar approach for a BFSK based network, we conclude that

$$\begin{aligned}
&\gamma_{min} = \gamma|_{a_i = t_1' \mathrm{SNR}_i} \\
=&\frac{19.655\mathrm{SNR}_i^4}{\mathrm{SNR}_i^5 + 8\mathrm{SNR}_i^4 + 26\mathrm{SNR}_i^3 + 43\mathrm{SNR}_i^2 + 36\mathrm{SNR}_i + 12} \\
\approx&19.655\frac{\mathrm{SNR}_i^2}{(2+\mathrm{SNR}_i)^3}
\end{aligned} \tag{B.20}$$

where $t_1' = 0.15427658$, and noise-injection is beneficial only if $\mathrm{SNR}_i \geq 6.4828$.

B.7 Proof of Lemma 4.6

We begin with the simpler case where the network only consists of PLs 1,2 and VL 1 whose outputs are y_1, y_2 and $y_1^{(v)}$. The probability that two PLs and the VL all give

identical decision can be decomposed as

$$\begin{aligned}
&\Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}\right) \\
=&\frac{1}{2}\left[\Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}|x=1\right) + \Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}|x=-1\right)\right] \\
\stackrel{def}{=}&p_{12(v1)}
\end{aligned} \tag{B.21}$$

The binary symmetric nature of both PLs gives

$$\begin{aligned}
&\Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}|x=-1\right) = \Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}|x=1\right) \\
=&\Pr\left(\widehat{y}_2 = 1|x=1\right)\Pr\left(\widehat{y}_1 = \widehat{y}_1^{(v)} = 1|x=1\right) \\
+&\Pr\left(\widehat{y}_2 = -1|x=1\right)\Pr\left(\widehat{y}_1 = \widehat{y}_1^{(v)} = -1|x=1\right) \\
=&(1-e_2)p_{cm} + e_2 p_{em}
\end{aligned} \tag{B.22}$$

Based on the normalized model for link 1, $y_1 = hx + w$, where $x \in \{\pm 1\}$, h is Raleigh distributed, and w is a zero mean Gaussian random variable with variance $\mathrm{var}(w) = N_0/2 = 1/2\mathrm{SNR}_1$, we obtain

$$\begin{aligned}
&p_{em} \\
=&\Pr\left(\widehat{y}_1 = \widehat{y}_1^{(v)} = -1|x=1\right) = \Pr\left(\widehat{y}_1 = \widehat{y}_1^{(v)} = 1|x=-1\right) \\
=&\int_h \Pr\left(-h+w>0, -h+w+w_v>0|x=-1,h\right) f(h)dh \\
=&\int_h \Pr\left(n > h\sqrt{\frac{2}{N_0}}, m > h\sqrt{\frac{2}{a_1^{(v)}N_0}}\middle| x=-1,h\right) f(h)dh
\end{aligned} \tag{B.23}$$

where $m = (w+w_v)/\sqrt{(a_1^{(v)}N_0)/2}$, $n = w/\sqrt{N_0/2}$, w_v is a zero mean real Gaussian random variable with variance $(a_1^{(v)}-1)N_0/2$, and $E[nm] = 1/\sqrt{a_1^{(v)}}$.

The first integrand of (B.23) can be expressed as a standard bivariate Gaussian

distribution function $Q(x, y; \rho)$ which, in turn, yields the Craig form as [37, (4.17)]

$$\begin{aligned}
&\Pr\left[n > h\sqrt{\frac{2}{N_0}}, m > h\sqrt{\frac{2}{a_1^{(v)}N_0}}\middle| x = -1, h\right] \\
=&Q\left(h\sqrt{\frac{2}{N_0}}, h\sqrt{\frac{2}{a_1^{(v)}N_0}}; \rho\right) \\
=&\frac{1}{2\pi}\int_0^{\tan^{-1}\left(\frac{\sqrt{a_1^{(v)}-1}}{1-\rho\sqrt{a_1^{(v)}}}\right)} \exp\left(-\frac{2h^2}{2N_0\sin^2\Phi}\right) d\Phi \\
+&\frac{1}{2\pi}\int_0^{\tan^{-1}\left(\frac{1}{\sqrt{a_1^{(v)}-1}}\right)} \exp\left(-\frac{2\rho^2h^2}{2N_0\sin^2\Phi}\right) d\Phi
\end{aligned} \tag{B.24}$$

where $\rho = \frac{1}{\sqrt{a_1^{(v)}}}$ is the correlation coefficient.

Using the method described in [37, ch.5] and the identity [37, (5.A.11)]

$$\int\left(1+\frac{c}{\sin^2\Phi}\right)^{-1} d\Phi = \Phi - \sqrt{\frac{c}{c+1}}\tan^{-1}\left[\frac{\tan\Phi}{\sqrt{\frac{c}{c+1}}}\right]$$

we obtain

$$\begin{aligned}
&\int_h Q\left(h\sqrt{\frac{2}{N_0}}, h\sqrt{\frac{2}{a_1^{(v)}N_0}}, \rho\right) f(h)dh \\
=&\frac{1}{2\pi}\int_0^{\pi/2}\left(1+\frac{1}{N_0\sin^2\Phi}\right)^{-1} d\Phi + \frac{1}{2\pi}\int_0^{\tan^{-1}\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)}\left(1+\frac{\rho^2}{N_0\sin^2\Phi}\right)^{-1} d\Phi \\
=&\frac{1}{4}\left(1-\sqrt{\frac{\mathrm{SNR}_1}{1+\mathrm{SNR}_1}}\right) + \frac{1}{2\pi}\left[\tan^{-1}\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)\right. \\
&\left.-\sqrt{\frac{\rho^2\mathrm{SNR}_1}{1+\rho^2\mathrm{SNR}_1}}\tan^{-1}\left(\frac{\rho\sqrt{1+\rho^2\mathrm{SNR}_1}}{\sqrt{(1-\rho^2)\rho^2\mathrm{SNR}_1}}\right)\right] \\
=&\frac{e_1}{2} + \frac{1}{2\pi}\left[\tan^{-1}\left(\frac{1}{\sqrt{a_1^{(v)}-1}}\right) - (1-2e_1^{(v)})\tan^{-1}\left[\frac{\left(1-2e_1^{(v)}\right)^{-1}}{\sqrt{(a_1^{(v)}-1)}}\right]\right] \\
\stackrel{def}{=}&p_{em}(e_1, a_1^{(v)})
\end{aligned} \tag{B.25}$$

Invoking the relation [37, (6.42)]

$$Q(-x, -y; \rho) = 1 - Q(x) - Q(y) + Q(x, y; \rho) \qquad x, y \geq 0$$

and (B.25), we express the conditional correct (pairwise) SMP as

$$\begin{aligned}
&\Pr\left(\widehat{y}_1 = \widehat{y}_1^{(v)} = 1 | x = 1\right) = \Pr\left(\widehat{y}_1 = \widehat{y}_1^{(v)} = -1 | x = -1\right) \\
=&\int_h \Pr\left(w > -h, m > -h | x = -1, h\right) f(h) dh \\
=&1 - e_1 - e_1^{(v)} + p_{em}(e_1, a_1^{(v)}) \stackrel{def}{=} p_{cm}(e_1, a_1^{(v)})
\end{aligned} \tag{B.26}$$

Summarizing (B.21)—(B.26), we then obtain

$$p_{12(v1)} = e_2 p_{em}(e_1, a_1^{(v)}) + (1 - e_2) p_{cm}(e_1, a_1^{(v)}) \tag{B.27}$$

which is (4.27) in the main text. The other probabilities, (4.30)-(4.33), can be similarly derived with the aid of the following two identities [37, (6.42)]:

$$Q(x, y, \rho) = Q(x) - Q(x, -y, -\rho),\ x \geq 0, y < 0 \tag{B.28}$$

$$Q(x, y, \rho) = Q(y) - Q(-x, y, -\rho),\ x < 0, y \geq 0 \tag{B.29}$$

Bibliography

[1] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490-530, 2005.

[2] S. M. Kay, "Noise enhanced detection as a special case of randomization," *IEEE Signal Process. Lett.*, vol. 15, pp. 709-712, 2008.

[3] S. M. Kay, "Can detectability be improved by adding noise?," *IEEE Signal Process. Lett.*, vol. 7, no. 1, pp. 8-10, Jan. 2000.

[4] S. M. Kay, J. H. Michels, H. Chen, and P. K. Varshney, "Reducing probability of decision error using stochastic resonance," *IEEE Signal Process. Lett.*, vol. 13, no. 11, pp. 695-698, Nov. 2006.

[5] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part I-Fixed detectors," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3172-3184, Jul. 2007.

[6] H. Chen, P. K. Varshney, and J. H. Michels "Noise enhanced parameter estimation," *IEEE Trans. Signal Processing*, vol. 56, pp. 5074-5081, Oct. 2008.

[7] G. P. Harmer, B. R. Davis, and D. Abbott, "A review of stochastic resonance: Circuits and measurement," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 2, pp. 299-309, Apr. 2002.

[8] A. Longtin, "Stochastic resonance in neuron models," *J. Statist. Physics*, vol. 70, no. 1, pp. 309V327, Jan. 1993.

[9] S. Mitaim and B. Kosko, "Adaptive stochastic resonance," *Proc. IEEE*, vol. 86, pp. 2152-2183, 1998.

[10] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, no. 1, pp. 223-287, 1998.

[11] S. Kay, "Generalizing Stochastic Resonance by the Transformation Method," Available: http://www.ele.uri.edu/faculty/kay/New%20web/downloadable%20files/transform_method.pdf.

[12] Z. Gingl, R. Vajtai, and L. B. Kiss, "Signal-to-noise ratio gain by stochastic resonance in a bistable system," *Chaos. Solitons & Fractals,* 11, pp. 1929-1932, 2000.

[13] G. O. Balkan and S. Gezici, "CRLB based optimal noise enhanced parameter estimation using quantized observations," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 477-480, May 2010.

[14] A. R. Bulsara and A. Zador, "Threshold detection of wideband signals: A noise-induced maximum in the mutual information," *Phys. Rev. E*, vol. 54, no. 3, pp. R2185-R2188, Sep. 1996.

[15] M. Stemmler, "A single Spike suffices: The simplest form of stochastic resonance in model neurons," *Network: Computation in Neural Systems,* vol. 7, pp. 687-716, 1996.

[16] J. J. Collins, C. C. Chow, A. C. Capela, and T. T. Imhoff, "Aperiodic stochastic resonance," *Phys. Rev. E*, vol. 54, no. 5, pp. 5575-5584, Nov. 1996.

[17] M. D. McDonnell and D. Abbott. "What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology." *PLoS Computational Biol.*, 5(5): e1000348. doi:10.1371, May 2009.

[18] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062-3080, Dec. 2004.

[19] I. Olabarrieta, J. Del Ser, "Enhanced sensing error probability estimation for iterative data fusion in the low SNR regime," in *Int. ITG Workshop Smart Antennas (WSA)*, Bremen, Germany, 2010.

[20] J. Del Ser, J. Garcia-Frias, and P. M. Crespo, "Iterative concatenated zigzag decoding and blind data fusion of correlated sensors," in *Int. conf. Ultra Modern Telecom. & Workshops (ICUMT)*, St. Petersburg, Russia, 2009.

[21] A. Razi, F. Afghah, and A. Abedi, "On minimum number of wireless sensors required for reliable binary source estimation.pdf," in *Proc. IEEE Wireless Comm. Networking Conf. (WCNC).*, Cancun, Quintana Roo, pp. 1852-1857, Mar. 2011.

[22] A. Razi, F. Afghah, and A. Abedi, "Binary source estimation using a two-tiered wireless sensor network," *IEEE Commun. Lett.*, vol. 4, pp. 449-451, Apr. 2011.

[23] F. Daneshgaran, M. Laddomada, and M. Mondin, "Iterative joint channel decoding of correlated sources," *IEEE inform. Theory*, vol. 51, pp. 2721-2731, Jul. 2005.

[24] J. N. Laneman and G. W. Wornell, "Energy-efficient antenna sharing and relaying for wireless networks," in *Proc. IEEE Wireless Comm. Networking Conf. (WCNC)*, Chicago, IL, pp. 7-12, Sept. 2000.

[25] D. Chen and J. N. Laneman, "Modulation and demodulation for cooperative diversity in wireless systems," *IEEE Trans. Wireless Commun.*, vol. 5, pp.1785-1794, Jul. 2006.

[26] Q. Zhao and H. Li, "Differential modulation for cooperative wireless systems," *IEEE Trans. Signal Processing*, vol. 55, pp. 2273-2283, May 2007.

[27] J. Luo, R. S. Blum, L. J. Cimini, L. J. Greenstein, and A. M. Haimovich,"Link-failure probabilities for practical cooperative relay networks," in *Proc. IEEE Veh. Tech. conf. (VTC) spring*, Stockholm, Sweden, 2005.

[28] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw. J.*, vol. 38, no. 4, pp. 393-422, Mar. 2002.

[29] L. Ruiz-Garcia, L. Lunadei, P. Barreiro and J. Robla, "A Review of Wireless Sensor Technologies and Applications in Agriculture and Food Industry: State of the Art and Current Trends," *Sensors*, 9, pp. 4728-4750, 2009.

[30] D. Puccinelli and M. Haenggi, "Wireless sensor networks: applications and challenges of ubiquitous sensing," *IEEE Circuits and Systems Magazine*, vol. 3, no. 3, pp. 19-29, 2005.

[31] L. Y. Yu, N. Wang, and X. Q. Meng, "Real-time forest fire detection with wireless sensor networks," *Proc. Int. Conf. Wireless Commun., Networking, Mobile Comput.*, vol. 2, pp.1214-1217, 2005.

[32] A. Dixit, S. C. Douglas, and G. C. Orsak, "Blind estimation of channel BERs in a multi-receiver network," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP) 2004*, Montreal, QC, Canada, May 17-21, 2004, vol. 2, pp. ii-305-ii-308.

[33] J. P. Delmas and Y. Meurisse, "Performance analysis of optimal blind fusion of bits," in *IEEE Trans. signal processing*, vol. 55, pp. 1477-1485, Apr. 2007.

[34] B. Liu, A. Jeremić, and K. M. Wong, "Optimal distributed detection of multiple hypotheses using blind algorithm," *IEEE Trans. aerospace and electronic systems*, vol. 47, pp. 317-331, Jan. 2011.

[35] E. L. Lehmann and G. Casella, *Theory of point estimation*, 2nd ed. Springer, 1998.

[36] A. Dixit and G. C. Orsak, "Asymptotically optimal blind fusion of bit estimates," presented at the IEEE Signal Processing Workshop, Callaway Gardens, GA, Oct. 13-16, 2002.

[37] M. K. Simon and M. Alouini, *Digital Communication over Fading channels: A Unified Approah to Performance Analysis*, 2nd, John Wiley & Son, 2004.

[38] Y. B. Li and Y. L. Guan, "Modified Jakes' model for simulating multiple uncorrelated fading waveforms," in *Proc. 51st IEEE Vehicular Technology Conf.*, pp. 1819-1822, Tokyo, Japan, May 2000.

[39] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge University Press, 1998.

[40] P. J. Smith, M. Shafi, and H. Gao, "Quick simulation: a review of importance sampling techniques in communications systems," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 597-613, May. 1997.

[41] M. C. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 153-170, 1984.

[42] K. S. Shanmugam and P. Balaban, "A modified Monte-Carlo simulation technique for the evaluation of error rate in digital communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1916-1924, Nov. 1980.

[43] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*, 2nd, John Wiley & Sons, New York, , 2008.

[44] E. Agrell, J. Lassing, E. G. Ström, and T. Ottosson, "On the optimality of the binary reflected Gray code," *IEEE Trans. Inf. Theory,* vol. 50, pp. 3170-3182, Dec. 2004.

[45] W. T. A. Lopes, W. J. L. Queiroz, F. Madeiro and M. S. Alencar, "Exact bit error probability of M-QAM modulation over flat Rayleigh fading channels," in *Proc. Microwave and Optoelectronics Conf., 2007 (IMOC 2007)*, pp. 804-806, Oct. 29-Nov. 1 2007.

[46] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York: Cambridge University Press, 2007.

[47] Z. Xiong; A. D. Liveris and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Mag.*, Vol. 21, Issue 5, pp. 80-94, 2004.

[48] N. Patwari, A. O. Hero, III, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. Signal Processing*, Vol. 51, pp. 2137-2148, Aug. 2003.

[49] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-Part I: Gaussian case," *IEEE Trans. Signal Processing*, Vol. 54, pp. 1131-1143, Mar. 2006

[50] S. M. Kay, *Fundamentals of statistical signal processing, Vol. 1: Estimation theory* , Prentice Hall, 1998.

[51] T. A. Schonhoff and A. A. Giordano, *Detection and estimation theory and its applications* , Prentice Hall, 2006.

[52] T. Tao and V. Vu, *Additive Combinatorics*, New York, Cambridbe, 2006.

[53] K. S. Shanmugam and P. Balaban, "A modified Monte-Carlo simulation techniques for the evaluation of error rate in digital communication systems," *IEEE Trans. Commum.*, vol. 28, pp. 1916-1924, Nov. 1980.

[54] G. Casella and R. L. Berger, *Statistical inference*, Duxbury, 2nd ed., 2002.