Power allocation for distortion outage minimization in wireless sensor networks

by

Chih-Hong Wang

Submitted in total fulfilment of
the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Electronic Engineering
The University of Melbourne
Australia

February, 2011

The University of Melbourne

Australia

Abstract

Power allocation for distortion outage minimization in wireless sensor networks

by Chih-Hong Wang

This thesis presents energy-efficient power allocation algorithms for wireless sensor networks used in distributed estimation. We focus on problems that require minimizing the distortion outage probability when fading is present. Optimal and suboptimal power allocation schemes have been obtained and analyzed for a number of different wireless sensor network models and network assumptions.

We first look at power allocation for a clustered wireless sensor network where the clusterheads send their observations to the fusion center through orthogonal multi-access channels. Assuming full CSI (channel state information) is available, we obtain optimal power allocation that minimizes distortion outage probability subject to a long-term average power constraint for any arbitrary fading distributions. We then consider partial CSI at the transmitter via limited feedback and propose suboptimal power allocation schemes with low computational complexity for Nakagami-m fading channels. Simulation results show significant power gain for just a few bits of feedback.

We then consider a wireless sensor network where sensors send their measurements to the fusion center using coherent multi-access assuming full CSI is available. We present three power allocations - equal power allocation, short-term optimal power allocation and long-term optimal power allocation - and analyze theoretically the diversity order of estimation outage, given in terms of the number of sensors in the network, that these three power allocation can achieve for Rayleigh fading. Simulation results are given for performance comparison.

Finally we consider power allocation for a wireless sensor network where the sen-

sor data are spatially correlated. We obtained necessary and sufficient conditions for finding the optimal power allocation. Simulation results show significant power gain can be achieved when the correlation between the sensor data is exploited. Various other aspects of the problem such as finding the closed form power allocations for any general correlation matrix and asymptotic outage performance analysis are still open problems. Some ideas are provided for future work.

This is to certify that

(i) the thesis comprises only my original work,

(ii) due acknowledgment has been made in the text to all other material used,

(iii) the thesis is less than 100,000 words in length, exclusive of table, maps, bibliographies, appendices and footnotes.

Signature_____

Date_____

Acknowledgments

I would like to thank my supervisors Professor Subhrakanti Dey, Professor Jamie Evans and Dr Brian Krongold for their continuing guidance, support and effort which have contributed greatly to the contents of this thesis. I would like to also thank Dr Alex Leong for his collaboration, help, source of ideas and many great research skills demonstrated by him.

Many thanks to the ARC Special Research Center for Ultra-Broadband Information Networks (CUBIN) and staffs and colleagues in CUBIN for providing a friendly and enjoyable working environment and many great memories.

I would like to thank my family and friends for their encouragement, support and understanding. Financial support was provided by the Australian Research Council and an Australian Postgraduate Award. I am also grateful for the funding provided by my parents for the last few months of candidature.

List of Publications

Conference Papers

1. C. Wang and S. Dey. Power allocation for distortion outage minimization in clustered wireless sensor network. In *Proc. IWCMC'08*, pages 395-400, Crete Island, Greece, August 2008.

2. C. Wang and S. Dey. Distortion outage minimization in Rayleigh fading using limited feedback. In *Proc. GLOBECOM'09*, Honolulu, Hawaii, USA, Nov-Dec 2009.

3. C. Wang, A. S. Leong and S. Dey. On diversity orders of distortion outage for coherent multi-access channels. In *Proc. ISIT'11*, Saint Petersburg, Russia, Jul-Aug 2011.

Journal Papers

1. C. Wang and S. Dey. Distortion outage minimization in Nakagami fading using limited feedback. *EURASIP Journal on Advances in Signal Processing*, in press, (accepted on June 21, 2011).

2. C. Wang, A. S. Leong and S. Dey. Distortion outage minimization and diversity order analysis for coherent multi-access. *IEEE Trans. Signal Processing*, in press, (accepted on August 2, 2011).

Contents

List of Figures

List of Acronyms

AWGN	Additive White Gaussian Noise
c.d.f.	cumulative distribution function
CH	Clusterhead
CSI	Channel State Information
CSIR	CSI at the receiver
CSIT	CSI at the transmitter
EPA	Equal Power Allocation
EPPC	Equal Power Per Clusterhead
EPPR	Equal Power Per Region
ES	Exhaustive Search
FC	Fusion Center
FDMA	Frequency-Division Multiple Access
KKT	Karush-Kuhn-Tucker
LT-OPA	Long-Term OPA
MAC	Multi Access Channel
MIMO	Multiple-Input Multiple-Output
MMSE	Minimum Mean Square Error
NZPOR	Non-Zero Power in Outage Region
OPA	Optimal Power Allocation
p.d.f.	probability density function
SLA	Straight-Line Approximation
SPSA	Simultaneous Perturbation Stochastic Approximation
ST-OPA	Short-Term OPA
WSN	Wireless Sensor Network
ZPOR	Zero Power in Outage Region

Chapter 1

Introduction

This chapter first gives the problem statement of the thesis. This includes a general introduction and gives the motivations that lead to the formulation of the problem. A survey of related literatures is given next. The author then defines the scope of the research problem as well as the aim of the research. Next the contribution of thesis is presented, and this chapter concludes by giving the outline structure of the thesis.

1.1 Problem statement

Wireless sensor network (WSN) consists of many sensor nodes distributed over a geographical area where the sensors can communicate with other sensors through wireless channels to accomplish certain tasks. WSNs have many potential applications that have already been implemented and those yet to emerge as newer technologies become available. They can be used in environmental and wildlife habitat monitoring, in tracking targets for defense applications, monitoring chemical/poisonous gas level in factories, in healthcare products and many other areas of human life. A real example of WSN developed in Australia is a smart sensor network that consists of 120 sensor nodes that monitor the quality of drinking water in south-east Queensland [1].

Wireless sensor network is a type of ad-hoc network, which is characterized by any type of networks that can be set up wirelessly without the use of infrastructure [2]. Hence, many network configurations exist for WSNs, for example, sensors in a WSN can configure themselves into clusters, elect clusterheads, perform cooperative transmission by acting as relay nodes and many more. An example of a wireless

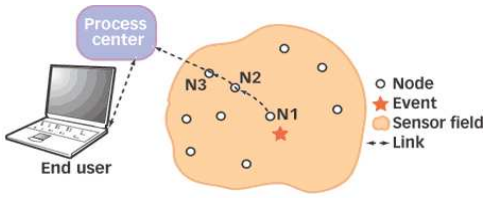

Figure 1.1: An example of a WSN.

sensor network structure is given in Fig. 1.1 [3]. However this ad-hoc nature of WSNs also brings many design challenges that are often unique for different applications.

Sensors are usually cheap, mass-manufactured, battery-operated devices that have limited energy and processing and communication capabilities. They are usually deployed in large numbers with high redundancies in wireless sensor networks. As technology advances, it is expected that the sensor to become smaller in size. A picture of a typical sensor is shown in Fig. 1.2 [4], depicting the smallness in physical size of the sensor node. Replacement of batteries is usually costly and unnecessary, since sensors are expected to be deployed once only and the environment in which the sensors are deployed can be hazardous or hard to reach. Energy is consumed in WSNs (and in ad-hoc networks in general) mainly in one of the three modes of operation - transmitting, receiving and in 'listening' mode, and energy is consumed differently in different modes of operation [5]. Generally energy is consumed in circuit operation (e.g. in DAC (digital to analog converter), ADC (analog to digital converter) and frequency synthesizer) and in transmission. For long-range applications the transmission energy dominates, while in short-range applications, circuit energy and transmission energy consumptions are comparable to each other [5].

As ad-hoc networks (and hence WSNs) are critically dependent on the rate of energy consumption, how to efficiently manage the energy/power consumption of sensors is a problem that is particularly crucial for wireless sensor networks. Power control is a strategy that is used to manage the transmission power of sensors in WSNs. Power control in WSNs can be centralized or distributive or a mixture of both. In centralized power control, a central processing unit requires to collect all of the network parameters and compute the power allocation which is then sent back

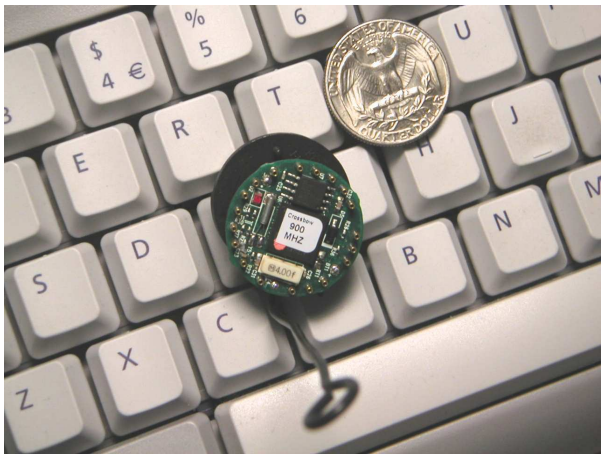

Figure 1.2: A picture of Mica2dot sensor node manufactured by Crossbow Technology Inc..

to each sensors. Centralized power control thus requires greater communication overheads. In distributed power control, the sensors are capable of computing the power based on regional information, such as information gathered from its neighbor sensors. In partial centralized/distributive power allocation schemes, only some information (some key parameters) are exchanged between the sensors and the central processing unit, and both sensors and central processing unit perform computation.

The issue of energy consumption in WSNs is of great importance. This motivates the author to conduct research in obtaining the derivation and theoretical analysis of power control algorithms in WSNs.

1.2 Literature review

Recently many works in the literature have proposed and studied cross-layer optimization to maximize the lifetime of energy-constrained wireless sensor networks and wireless networks in general. The traditional design approach for wireless networks has been based on the OSI (open system interconnection) model framework that sub-divides a communication system into seven layers based on its functionalities. Although the OSI model has simplified the protocol design of communication systems greatly, the layers are usually designed independently and does not take into

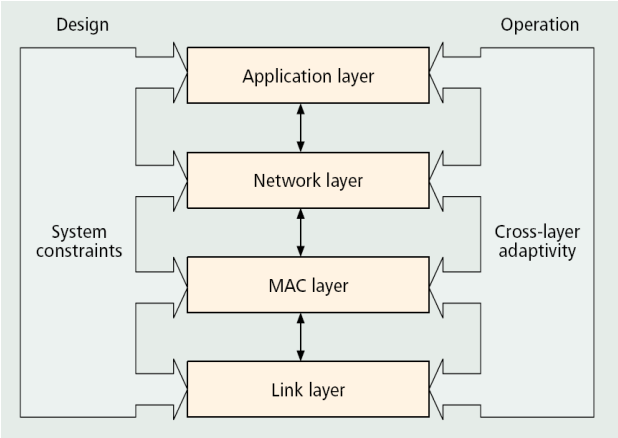

Figure 1.3: A graphical illustration of an adaptive cross-layer design and operation.

account of parameters used by other layers. This design method is hence regarded as sub-optimal for wireless communication systems [6, 7] and, as noted in [6], system components such as medium access control protocols, radio link control, radio resource management schemes and routing algorithms can all benefit in some degree from the awareness of the the variations in the channel in time and frequency. This leads to the concept of cross-layer optimization where the layers in the wireless communication system should be designed jointly by taking into account of global system constraints and the relevant parameters that are exchanged between layers [7]. See Fig. 1.3 for a graphical illustration of an adaptive cross-layer design and operation [7]. These parameters may include the channel state information, QoS-related parameters, network resource information and the traffic pattern offered by each layer to others, see [6]. Although cross-layer optimization can bring higher efficiency in wireless communication systems, it also comes with extra signaling required to extract and process the relevant parameters, and greater computation complexity. This trade-off should be investigated in a case-by-case basis to see if the overall gain from the cross-layer optimization outweighs the cost [6].

Cross-layer optimization has been studied in many areas and applications of wireless networks and is shown to improve the lifetime of such networks. The authors in [8] studied minimum-energy scheduling problem in wireless data networks that is

adaptive jointly to backlog and channel conditions subject to a packet transmission deadline over multi-access channels, broadcast channels, and channels with fading. They have deviced a heuristic online algorithm that can achieve significant energy savings by such joint adaptation. Work in [9] studied the problem of energy efficient transmission scheduling for delay constrained (both a deadline constraint and an average delay constraint) wireless networks. In particular, they investigated the jointly use of channel coding and electro-chemical mechanisms in batteries (batteries allow them to recover energy during idle periods) to achieve significant energy savings. Cross-layer optimization method is applied to wireless bio-sensor networks in [10] where the authors propose to jointly optimize over congestion control and medium access control, and jointly over power and bio-effects control. The authors in [11] investigated a cross-layer optimization problem that computes the optimal transmission powers, rates, and link schedule that maximize the network lifetime in interference-limited wireless sensor networks. They observed the advantage of using transmission strategies such as multi-hop routing, load balancing, interference mitigation and frequency reuse to increase the lifetime of energy-constrained networks. [12] studied adaptive transmission rate and power based on perfect channel state information and the buffer occupancy and illustrated the trade off between average power and average delay. The authors also gave some discussion on the architectural issues relating to cross-layer design.

Cooperative diversity can be utilized to improve the performance through the use of spatial diversity gain and MIMO (multiple input multiple output) techniques in wireless sensor networks. Here multiple sensors (terminals) in a network cooperate to form a virtual antenna array realizing spatial diversity in a distributed fashion [13]. The authors in [13] studied the simplest form of cooperative network - a single-relay fading channel, where the relay terminal uses either amplify-and-forward or decode-and-forward. They showed that based on an outage capacity analysis, full spatial diversity is achieved by certain time-division multiple-access-based cooperative protocols provided that appropriate power control is employed. Various cooperative diversity and protocols for multiple source-relay networks have also been studied in [14–17]. In [14] the authors studied power allocation in wireless

networks under decode-and-forward cooperative diversity and assumed that only the mean channel gains are available at the transmitters. They showed that using a near-optimal solution, significant performance gain can be achieved over other schemes in the literature. Multi-hop, cluster-based sensor network with cooperative distributed MIMO channels is studied in [18] for minimizing the end-to-end outage probability by deriving the optimum time and power allocated to each cluster subject a energy constraint. Energy-efficiency of MIMO and cooperative MIMO techniques are also investigated in [5] in sensor networks having throughput and delay requirements. In the energy consumption model the authors considered transmission energy as well as the circuit energy consumption. Cross-layer optimization based on cooperative MIMO techniques with rate adaption is considered in [19] and achieves significant improvement in energy and delay performances.

Recent result in [20] demonstrating the asymptotic optimality of uncoded analog forwarding of measurements by multiple sensors as opposed to separate source channel coding have motivated a lot of researchers to investigate multi-sensor estimation (distributed estimation) problems and related energy/power efficiency issues within this uncoded transmission framework. In [21] an optimal power allocation scheme is obtained for analog forwarding based transmission through an AWGN (additive white Gaussian noise) multi-access channel in an inhomogeneous Gaussian sensor network. [22] looks at estimation diversity and energy efficiency in distributed sensing. It shows that the estimation diversity increases in the order of the number of sensors and derives optimal power allocation schemes for minimum distortion under power constraint and minimum power under distortion constraint using orthogonal multi-access protocol. The same group of authors later in [23] studied optimal linear decentralized estimation in coherent MAC (multi-access channel) for scalar and vector sources under bandwidth and power constraints. Minimum energy problem with correlated sensor noise is studied in [24]. The aforementioned works are based on static channels and do not explicitly take into account fading channels, for which meeting a strict distortion constraint may not be always possible. This motivates us to study distortion outage, similar to the concept of capacity outage in delay-sensitive applications such as voice communications over wireless channels [25–27],

and seek to find the power allocation policies that minimize distortion outage probability when channel fading is present.

1.3 Aim and scope

We consider WSNs used in distributed estimation. This type of WSN usually has sensors deployed over a geographical area to monitor some physical phenomenon of interest which we call the *source*. The sensors take measurements of the source independently and transmit their measurements wirelessly to a central processing unit, or commonly known as the fusion center (FC). After having received all the measurements (or some of the measurements) from the sensors, the fusion center then tries to reconstruct, using the appropriate algorithms, the actual value of the physical phenomenon that is being observed. We call the algorithm that is used in reconstructing the source the *estimator* and the reconstructed quantity the *estimate* of the source.

There are a number of uncertainties in this network that makes the measurements received at the fusion center distorted; they are the measurement noise when the sensors take measurements, the channel fading which is a fundamental property of the wireless channel, as well as the channel noise. The estimate of the source hence is a random quantity as a function of all these uncertainties. In estimation theory, the estimator should be chosen (or constructed) such that it gives the minimum *distortion*, which is simply given by the mean square error or the variance of the estimate, given that the estimator is unbiased. In this work we will use *distortion outage probability* or simply *outage probability* as a performance measure. The distortion outage probability is given by the probability that the distortion exceeds a given threshold D_{max}.

Motivated by results in [20–22], the aim of this thesis is to investigate power allocation schemes and their performances relating to distortion outage probability in wireless sensor networks used in distributed estimation within the framework of amplify-and-forward transmission with different multi-access protocols and various sensor network assumptions.

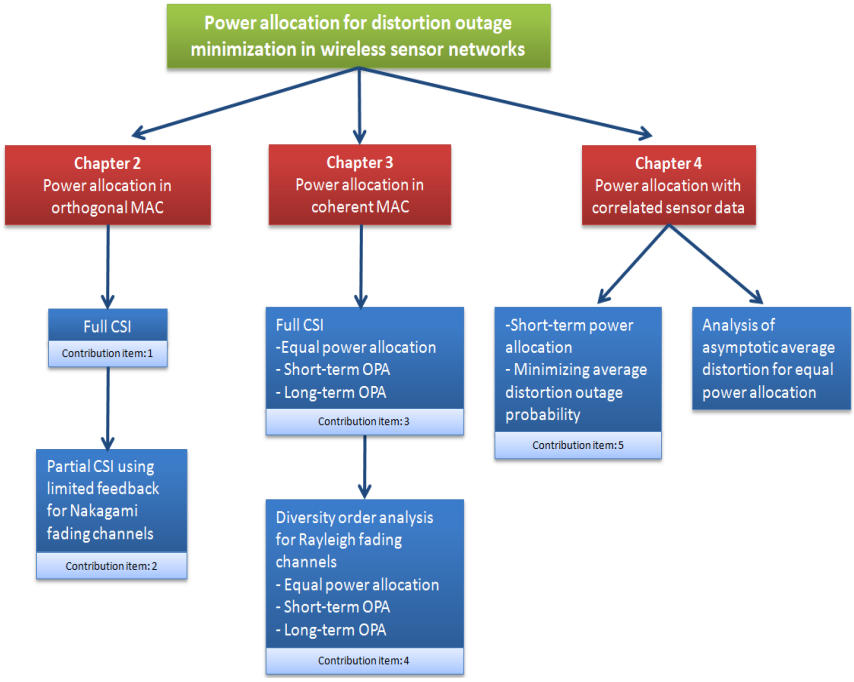

Figure 1.4: Block diagram of thesis structure and contribution items.

1.4 Contribution of thesis

A block diagram showing the organization and the contribution items of the thesis is shown in Fig. 1.4. The thesis has the following contribution items:

1. We consider fading channels in a clustered wireless sensor network used in distributed estimation and obtain optimal power allocation scheme that minimizes distortion outage probability with full CSI (channel state information) under amplify-and-forward transmission and orthogonal multi-access protocol. We show that the optimal power allocation schemes can be implemented in a mixture of centralized and distributed manner. Computer simulations show that the optimal power allocation scheme achieves significant power gain over both equal power allocation and power allocation based on *channel statistics*.

2. We extend the problem in item 1 and study a more practical scenario where we

consider partial CSI at the clusterheads (transmitters). We study the power allocation with limited feedback and obtain, after using a number of useful approximations, a sub-optimal scheme that is simple to implement (low in computational complexity). Simulation results show that outage performance with limited feedback achieves significant power gain over no CSI, and in some cases some power gain over power allocations obtained by simultaneous perturbation stochastic approximation.

3. We consider coherent multi-access protocol in a wireless sensor network used in distributed estimation with fading channels under amplify-and-forward transmission framework. We obtain the optimal power allocation that minimizes the distortion outage probability.

4. We study the diversity order of estimation outage for the coherent multi-access protocol case for equal power allocation, short-term optimal power allocation and long-term optimal power allocation schemes. We obtain the asymptotic expressions for these three power allocation schemes and show that the diversity orders increase in the order of $N \log N$ for both equal power and short-term optimal power allocations, where N is the number of sensors. For long-term optimal power allocation, we show that it is possible to drive the outage probability to zero with a finite long-term total power constraint.

5. We study the problem when the sensor data are spatially correlated. We consider a wireless sensor network using orthogonal multi-access and amplify-and-forward protocols. We obtain the optimal power allocation that minimizes the distortion outage probability subject to a long-term average power constraint and show that significant energy can be saved by incorporating correlation into the network model. The higher the correlation amongst the sensor data, the more power gain the network can achieve.

1.5 Organization of thesis

The thesis is organized as follows. In Chapter 2 we investigate the power allocation algorithm for a clustered wireless sensor network where sensors within a cluster transmit their measurement to their clusterhead using amplify-forwarding scheme. The clusterheads then transmit, also using amplify-forwarding scheme, the aggregated measurements from the sensors to the fusion center where an estimator is used to reconstruct the source. We assume that the clusterheads transmit through orthogonal multi-access Rayleigh-fading channels. We seek to find the optimal power allocation that minimizes the distortion outage probability for a given average power constraint. In this problem we assumed that full CSI is available at both the sensors and the fusion center. In order to implement the optimal power allocation algorithm to real systems it would require feedback channels (from fusion center to sensors) having zero noise, zero delay and infinite bandwidth. This motivates us to study power allocation with limited feedback, where only partial or quantized CSI is available at the sensors. The study of power allocation with limited feedback is given in the second half of Chapter 2.

In Chapter 3 we investigate power allocation algorithms that minimizes outage probability in wireless sensor networks where sensors transmit their signals to the fusion center via a coherent multi-access channel. Motivated by the results in diversity gain for orthogonal channels obtained in [22], we analyze the diversity order for the coherent MAC case. We give theoretical analysis on three different power allocation schemes - equal power allocation, short-term optimal power allocation and long-term optimal power allocation, and obtain their diversity order of estimation outage as the number of sensors gets large.

Measurements of sensors in a wireless sensor network may be spatially correlated. This correlation structure may be exploited to further improve the energy efficiency in wireless sensor networks. This motivates us to investigate the power allocation algorithms for wireless sensor networks with correlated data. This work is given in Chapter 4.

We conclude the thesis in Chapter 5 by giving concluding remarks and future

work.

Chapter 2

Power allocation in orthogonal MAC: full CSI and limited feedback

2.1 Introduction

Wireless sensor networks have many useful applications such as in environmental and wildlife habitat monitoring, in tracking targets for defense applications, in aged health care and many other ares of human life. Wireless sensor networks usually involve large numbers of sensor nodes that are distributed geographically to monitor certain physical phenomenon and collect measurements which are then sent to a central processing unit (often called a FC) via the wireless channel. The fusion center computes estimates of the samples of the physical phenomenon from the noisy measurements collected by the sensors. Energy consumption is a unique and important issue in wireless sensor networks performing such distributed estimation tasks because sensors are expected to be deployed once only and their batteries are often irreplaceable due to high cost. Due to random fading in wireless transmission, the quality of the estimation at the FC, measured by a distortion measure, becomes a random variable. In delay-limited settings, instead of minimizing a long term average distortion (or expected distortion for ergodic fading channels), it is more appropriate to minimize the probability that the distortion for each estimate exceeds a certain threshold, the so-called distortion outage probability. This is similar to the idea of minimizing the information outage probability in block-fading wireless communication channels in the information theoretic context [27]. Optimal power allocation at the sensor transmitters for such outage minimization under various power constraints is an important problem from the point of view of reducing energy consumption in sensor networks. In this chapter we first look at optimal power

allocation problem for distortion outage minimization in a clustered sensor network with full CSI at both the clusterheads (transmitters) and the receiver (FC). We then study a more practical power allocation problem where the sensors only have quantized CSI by using finite rate channel feedback broadcast by the FC to the transmitters at the clusterheads (CHs) for the various clusters.

Motivated by the recent result in [20] showing that uncoded analog-forwarding of measurements by multiple sensors asymptotically outperforms the separate source channel coding, many works in the literature have investigated multi-sensor estimation problems and related energy/power efficiency issues within this uncoded transmission framework [21–24]. Energy efficiency in wireless sensor networks has been studied in the context of power optimization problems where optimal power allocations (minimizing total power subject to a distortion constraint or minimizing distortion subject to a total power constraint) have been found for sensor networks with orthogonal MAC [22] and coherent MAC [21]. However these optimal power allocation strategies are based on static channels and do not explicitly take into account of fading channels, for which meeting a strict distortion constraint may not be always possible. In [27], optimal power allocation schemes were obtained for minimizing the information outage probability over block-fading channels subject to peak and/or long-term average power constraints. In this chapter we will use the same techniques as in [27] and obtain the optimal power allocation that minimizes the distortion outage probability for a clustered wireless sensor network with orthogonal MAC.

The optimal power allocation that minimizes the distortion outage probability, however, assumes perfect CSI at both transmitter and receiver. In practice, perfect CSI at the transmitter (CSIT) relies on instantaneous channel feedback from the FC, which is difficult to implement due to the limited bandwidth, delay and error in the feedback channel. Motivated by these constraints, many work in the literature have looked at power control in the field of MIMO beamforming systems with partial CSIT using limited feedback [28, 29]. The optimal power allocation scheme for systems employing limited feedback is in general complex and hence difficult to obtain. [30] studied average reliable throughput minimization over slow fading channels. They

found properties of optimal power allocation policy that aid in the design of power allocation algorithms. A suboptimal power allocation scheme is proposed in [31] for a single user system with multiple transmit antennas and single receive antenna with finite rate feedback power control. These suboptimal power allocation schemes, although not optimal, can provide significant gains over no-CSIT even for small number of feedback bits. A recent paper [32] studies the effect of partial CSIT in a distributed estimation problem over a multiaccess channel where various forms partial CSI are assumed to be available at the sensor transmitters, and their effect on minimization of distortion or estimation error is investigated. Finally, a related performance criterion in distributed estimation, called *distortion exponent*, measures the *slope* of the average end-to-end distortion on a log-log scale at high SNR [33]. This metric is similar to that of diversity gain studied in this paper (also termed as estimation diversity in [22]), which looks at the rate of diminishing of outage probability at high SNR rather than distortion.

In this chapter we study a wireless sensor network where sensors are organized into clusters. Each cluster has an elected clusterhead. The sensors observe a Gaussian random source and send their observed (noisy) information to the clusterhead (CH) by uncoded analog transmission using distributed beamforming. The cluster heads then transmit the combined received signal to the FC using amplify-and-forward and an orthogonal multi-access (e.g. FDMA (frequency-division multiple access)) where the channel is subject to random fading. It is assumed that the channel from each clusterhead to the FC is subject to independent and ergodic block fading where each fading block is long enough for all transmissions within the clusters and between the CHs and the FC to be completed and an estimate of the random Gaussian source to be computed at the FC. The details of this clustered wireless sensor network model is given in Section 2.2.

We present two main problems in this chapter. In the first problem, given in Section 2.3, we design an optimal power allocation scheme at the CH's (based on full CSI at the FC and the CH's) to minimize the distortion outage probability. The optimal power allocation scheme can be applied to any arbitrary fading distribution in general. In our work we have assumed Rayleigh fading and simulation

results showed that optimal power allocation achieves significant power gain in outage performance over equal power allocation. The optimal power allocation scheme obtained here assumes that full instantaneous CSIT is available via feedback from the FC, which requires an error-free, delay-less, infinite-bandwidth feedback channel and extra communication overheads, and this can be impractical in real systems. Hence we study some sub-optimal power allocation algorithms based on the statistics of the fading channels, by minimizing some upper bounds of the outage probability as obtaining an explicit expression for the outage probability proves to be difficult. It is seen that these statistical power allocation schemes do not fare well compared to the performance of the full CSI based algorithm, thus mandating the need for power allocation algorithms based on finite rate channel feedback.

In Section 2.4 we present our second problem which is to find power control schemes in Nakagami-m fading using limited feedback. The feedback system works as follows. We assume that an optimal power codebook (to be designed) is pre-computed at the FC based on the fading channel distributions and the average power constraint, and stored at the FC as well as the CH transmitters. In real time, under the assumption of perfect CSI at the FC, an index (corresponding to a region in a multi-dimensional space that the channel vector belongs to) is computed and broadcast to all CHs so that they can use the corresponding transmit power from their pre-computed power codebooks. This index can only take a limited number of values due to the finite rate constraint on the feedback channel. In general, solving for the globally optimal power codebook is difficult due to the non-convexity of the associated optimization problems and the difficulty of exactly computing the probability of the channel vector belonging to a specific region defined by the index of the power codebook. We obtain a power allocation scheme that is low in computational complexity after imposing a constraint on the structure of the power codebook and using various useful approximations for computing the probability of the channel vectors belonging to the multi-dimensional quantized regions and distributions of average sum power across the various regions. We also study the asymptotic behavior of the outage probability and obtain an approximate expression of the diversity gain (as the number of feedback bits goes to infinity) achieved by

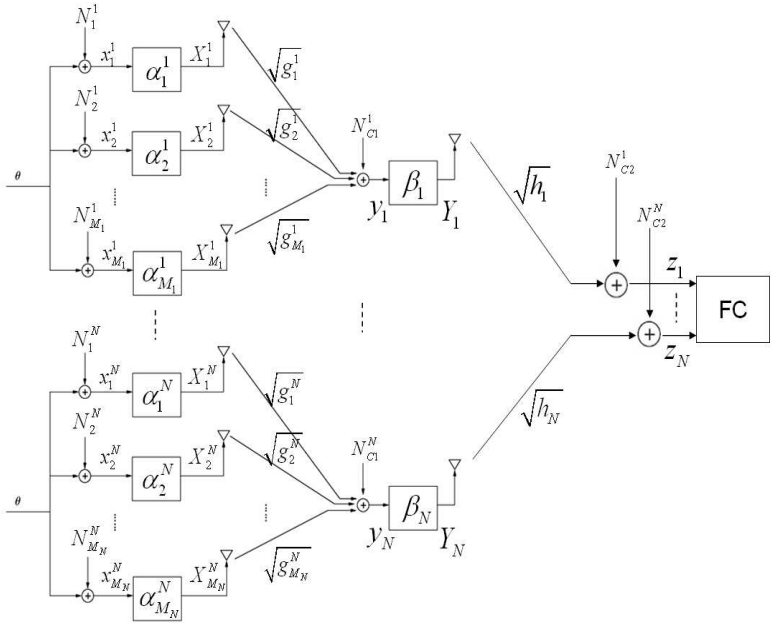

Figure 2.1: Schematic diagram of a wireless sensor network.

the proposed power allocation scheme.

2.2 Sensor network model

A schematic diagram of the wireless sensor network studied in this chapter is shown in Fig. 2.1. N clusters of sensors are distributed around a single point source $\theta[k]$ that is to be measured. Here $k = 0, 1, 2 \ldots$ denotes discrete time instants. We assume that $\theta[k]$ is an independent and identically distributed (i.i.d.) Gaussian, band-limited random process of zero mean and variance σ_θ^2. Each cluster contains M_i sensors (the subscript i denotes the index of the cluster and $i = 1, \ldots, N$) that observe the source and send their measurements to a pre-selected clusterhead. The observed sample $x_m^i[k]$ of the mth sensor in the ith cluster at time k is given as

$$x_m^i[k] = \theta[k] + N_m^i[k] \tag{2.1}$$

where $N_m^i[k]$ is the measurement noise which is i.i.d., Gaussian distributed of zero mean and variance $(\sigma_m^i)^2$. We assume that $\theta[k]$ is independent of $N_m^i[k]$, $\forall k, i, m$. The measurement noise variance is assumed to be proportional to the square of the

distance from the source to the sensor (although the noise model at the sensors can be appropriately adjusted depending on the sensing model of the sensors). A more realistic model may include spatial correlation of the observed signal amongst sensors, which is studied in Chapter 4. For simplicity, this chapter treats the observed signal as being spatially independent of each other. We also assume that CHs are selected by some chosen protocol, and that CHs are capable of transmitting with greater power than the sensors, since the transmission distance between a CH and the fusion center is larger in general.

Motivated by recent results showing asymptotic optimality of uncoded analog transmission from multiple sensors observing a Gaussian source [20], we assume that the sensors within a cluster simply amplify-and-forward (using equal power) their observations to CH via a multi-access channel using distributed beamforming so that the signals arriving at CH can be added up coherently. This is referred to as the first stage of transmission. Although distributed beamforming may not be easy to implement (see [34] for details of implementing distributed beamforming), specially in the case of large number of sensors within each cluster, it has been shown that even under the presence of random phase errors, the average loss in performance is not significant unless the variance of the phase errors is severely large. The signal received at the nth CH is given as

$$y_i[k] = \sum_{m=1}^{M_i} \left[\alpha_m^i \sqrt{g_m^i} \left(\theta[k] + N_m^i[k] \right) \right] + N + C1^i[k] \tag{2.2}$$

where α_m^i is the power gain factor, $\sqrt{g_m^i}$ is the channel gain of the first stage of transmission and N_{C1}^i is the zero-mean AWGN channel noise of variance $(\sigma_{C1}^i)^2$. We also assume that the channels between sensors and CHs are static, where the channel gains are assumed to be proportional to the inverse of the square of the transmission distance. We also assume that the signal received at each CH is not interfered by any signals from other clusters (which can be achieved by a time-division multi access protocol where each cluster operates in a different time slot). For simplicity, we let CHs also use the amplify-and-forward scheme to transmit $y_i[k]$ to FC using an orthogonal multi access protocol such as FDMA. We refer the transmission of

signals from CHs to FC as the second stage of transmission. We assume that full (instantaneous) CSI is available at both the transmitters (CHs) and the receiver (FC) (which can be obtained by delay-less and error-free feedback from FC once FC has estimated the channels using pilot tones). We do not consider the effects of channel estimation errors or power consumptions due to channel estimation in this work. The signal received at FC from the ith CH is given as

$$z_i[k] = \beta_i \sqrt{h_i} y_i[k] + N^i_{C2}[k] \tag{2.3}$$

where β_i is the power gain factor, $\sqrt{h_i}$ is the channel gain of the second stage of transmission and N^i_{C2} is the zero mean AWGN channel noise of variance $(\sigma^i_{C2})^2$.The received signal vector is given as $\mathbf{z} = \mathbf{s}\theta + \mathbf{v}$ where

$$\begin{aligned}
\mathbf{z} &= [z_1[k], \ldots, z_N[k]]^T \\
\mathbf{s} &= \left[\beta_1 \sqrt{h_1} \sum_{m=1}^{M_1} \alpha^1_m \sqrt{g^1_m}, \ldots, \beta_N \sqrt{h_N} \sum_{m=1}^{M_N} \alpha^N_m \sqrt{g^N_m} \right]^T \\
\mathbf{v} &= \left[\beta_1 \sqrt{h_1} \left(\sum_{m=1}^{M_1} \alpha^1_m \sqrt{g^1_m} N^1_m[k] + N^1_{C1}[k] \right) + N^1_{C2}[k], \right. \\
&\quad \left. \ldots, \beta_N \sqrt{h_N} \left(\sum_{m=1}^{M_N} \alpha^N_m \sqrt{g^N_m} N^N_m[k] + N^N_{C1}[k] \right) + N^N_{C2}[k] \right]^T
\end{aligned}$$

where ^{T} denotes transposition.

In what follows, we suppress the time index k for simplicity. The fusion center now needs to reconstruct the value of θ given the vector of measurements $\mathbf{z}$. Assuming that the FC has prior knowledge of the p.d.f. (probability density function) of θ, it is well-known from estimation theory that the optimal estimator to use in this case is the MMSE (minimum mean square error) estimator given as $\hat{\theta} = \frac{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{z}}{\frac{1}{\sigma^2_\theta} + \mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}}$ where $\mathbf{C}$ is a diagonal matrix with its ith diagonal element given as $C_{ii} = \beta_i^2 h_i \left(\sum_{m=1}^{M_i} (\alpha^i_m)^2 g^i_m (\sigma^i_m)^2 + (\sigma^i_{C1})^2 \right) + (\sigma^i_{C2})^2$. The variance of $\hat{\theta}$, also referred to as *distortion*, is given as $\mathrm{var}(\hat{\theta}) = \left[\frac{1}{\sigma^2_\theta} + \mathbf{s}^T \mathbf{C}^{-1} \mathbf{s} \right]^{-1}$. From Fig. 2.1 we may also obtain $X^i_m = \alpha^i_m(\theta + N^i_m)$ and $Y_i = \beta_i \left(\sum_{m=1}^{M_i} \sqrt{g^i_m} X^i_m + N^i_{C1} \right)$. Define q_i as the total power of sensors in the ith cluster and P_i the power of the ith CH. We then obtain

$q_i = \sum_{m=1}^{M_i} \left(\alpha_m^i\right)^2 \left(\sigma_\theta^2 + (\sigma_m^i)^2\right)$ and $P_i = \beta_i^2 \left(\sum_{m=1}^{M_i} \left(\alpha_m^i\right)^2 \left(\sigma_\theta^2 + (\sigma_m^i)^2\right) g_m^i + (\sigma_{C1}^i)^2\right)$. With the above model and expressions, we are now ready to formulate the problems.

2.3 Power control schemes in fading channels with full CSI

2.3.1 Problem formulation

We are primarily interested in obtaining the optimal power allocation that minimizes the total power of sensors and CHs subject to a distortion constraint at the FC, i.e.,

$$\begin{aligned} \underset{P_i, q_i}{\text{minimize}} \quad & \sum_{i=1}^{N} (P_i + q_i) \\ \text{subject to} \quad & \text{var}[\hat{\theta}] \leq D_{\max} \\ & q_i \geq 0 \quad \forall i \\ & P_i \geq 0 \quad \forall i. \end{aligned} \tag{2.4}$$

The optimization problem (2.4) can be easily shown to be non-convex. In order to avoid this difficulty, we assume that the sensors within clusters have fairly limited functionality and have only a few adjustable transmission power levels (e.g. low, medium and high transmission power). With this assumption we drop the optimization variable q_i and assume that it is fixed at a value within a finite set of a small number of elements. We investigate the effect of q_i later via simulations in Section 2.3.3. Furthermore we assume all sensors within a cluster transmit with equal power (q_i/M_i). Hence the expressions for sensor power gain, CH power and distortion become

$$\alpha_m^i = \sqrt{\frac{q_i}{M_i \left[\sigma_\theta^2 + (\sigma_m^i)^2\right]}} \tag{2.5}$$

$$P_i = \beta_i^2 C_i \tag{2.6}$$

$$\text{var}\left[\hat{\theta}\right] = \sigma_\theta^2 \left(1 + \sum_{i=1}^{N} \frac{\beta_i^2 h_i U_i}{\beta_i^2 h_i V_i + (\sigma_{C2}^i)^2}\right)^{-1} \tag{2.7}$$

where $C_i = (q_i/M_i)\sum_{m=1}^{M_i} g_m^i + (\sigma_{C1}^i)^2$, $U_i = (q_i/M_i)\left(\sum_{m=1}^{M_i}\sqrt{g_m^i/\left(1+(\gamma_m^i)^{-1}\right)}\right)^2$, $V_i = (q_i/M_i)\sum_{m=1}^{M_i}(g_m^i(\gamma_m^i)^{-1})/\left(1+(\gamma_m^i)^{-1}\right) + (\sigma_{C1}^i)2$ and $\gamma_m^i = \sigma_\theta^2/(\sigma_m^i)^2$.

We now solve this optimization problem for static (modeling only distance based attenuation) and fading (modeling random channel variations in addition to distance based attenuations) channels in the second stage of transmission (CH's to FC), and describe the corresponding problem formulations in the following two subsections respectively. Note that the fading channel gain, $\sqrt{h_i}$ is assumed to be i.i.d. Rayleigh-distributed, and hence the signal power gain, h_i is i.i.d. exponentially distributed (although the analysis can be extended to any other fading distribution). The fading channel power gain is modeled as

$$h_i = \zeta_i f_i \tag{2.8}$$

where ζ_i is the mean channel gain and f_i is i.i.d. exponentially distributed with unity mean (any non-unity mean value of f_i can be absorbed into ζ_i). The mean channel gain is assumed to be equal to the inverse of the transmission distance squared.

Static channel

In this section, we assume that the channel gains are static and distance-based, and are given by ζ_i. The optimization problem becomes

$$\begin{aligned} \min_{\beta_i^2} \quad & \sum_{i=1}^{N} \beta_i^2 C_i \\ \text{s.t.} \quad & \sigma_\theta^2 \left(1 + \sum_{i=1}^{N} \frac{\beta_i^2 \zeta_i U_i}{\beta_i^2 \zeta_i V_i + (\sigma_{C2}^i)^2}\right)^{-1} \leq D_{max} \\ & \beta_i^2 \geq 0,\ i = 1, \ldots, N. \end{aligned} \tag{2.9}$$

The solution of a variation to this problem (for a best linear unbiased estimator (BLUE) instead of the MMSE estimator) can be found in [22], and we just state it below as it will be useful in later sections where we solve the problem for fading

channels. The optimal power gain for problem (2.9) is given as

$$P_i^* = \begin{cases} 0, & i > N_1 \\ \frac{C_i G_i}{H_i}\left(\frac{1}{\sqrt{\eta_i^{-1}}\rho_0} - 1\right), & i \leq N_1 \end{cases} \tag{2.10}$$

where $G_i = U_i/V_i$, $H_i = \zeta_i U_i/(\sigma_{C2}^i)^2$, $\eta_i = H_i/C_i$ and $\rho_0 = D(N_1)/C(N_1)$. $D(i) = \sum_{j=1}^{i} G_j - (\sigma_\theta^2/D_{max} - 1)$ and $C(i) = \sum_{j=1}^{i} G_j/\sqrt{\eta_j}$. N_1 is given by ordering $\eta_1 \geq \ldots \geq \eta_N$ and finding $g(N_1) > 0$ and $g(N_1+1) \leq 0$, where $g(i) = 1 - D(i)/\left(\sqrt{\eta_i}C(i)\right)$, $i = 1, \ldots, N$.

Similarly, the solution for the dual problem given as

$$\begin{aligned} \min_{\beta_i^2} \quad & \sigma_\theta^2 \left(1 + \sum_{i=1}^{N} \frac{\beta_i^2 \zeta_i U_i}{\beta_i^2 \zeta_i V_i + (\sigma_{C2}^i)^2}\right)^{-1} \\ \text{s.t.} \quad & \sum_{i=1}^{N} \beta_i^2 C_i \leq \mathcal{P}_{tot} \\ & \beta_i^2 \geq 0, \ i = 1, \ldots, N \end{aligned} \tag{2.11}$$

can also be found in [22]. The optimal power allocation is given as

$$P_i^* = \begin{cases} 0, & i > N_1 \\ C_i v_i \left(\frac{1}{\sqrt{\xi_i^{-1}}c_0} - 1\right), & i \leq N_1 \end{cases} \tag{2.12}$$

where $v_i = (\sigma_{C2}^i)^2/\zeta_i V_i$, $\xi_i = \zeta_i U_i/C_i(\sigma_{C2}^i)^2$ and $c_0 = A(N_1)/B(N_1)$. $A(i) = \sum_{j=1}^{i} v_j \sqrt{\xi_j} C_j$ and $B(i) = \sum_{j=1}^{i} v_j C_j + \mathcal{P}_{tot}$. N_1 is given by ordering $\xi_1 \geq \ldots \geq \xi_N$ and finding $f(N_1) > 0$ and $f(N_1 + 1) \leq 0$, where $f(i) = \sqrt{\xi_i} B(i)/A(i) - 1$, $i = 1, \ldots, N$.

Fading channel

In this subsection we assume Rayleigh-faded channels between CH's and FC as given by (2.8) and define the probability of distortion outage. The probability of distortion outage is defined as the probability that the distortion exceeds some predefined threshold, D_{max}. We want to minimize this distortion outage probability subject to

a long term power constraint, stated as

$$\begin{aligned} \min \quad & \Pr\left(D(\mathbf{P}(\mathbf{h}),\mathbf{h}) > D_{max}\right) \\ \text{s.t.} \quad & E[\langle \mathbf{P}(\mathbf{h})\rangle] \leq \mathcal{P}_{av} \\ & \mathbf{P}(\mathbf{h}) \succeq 0 \end{aligned} \tag{2.13}$$

where $\mathbf{P}(\mathbf{h}) \triangleq [P_1(\mathbf{h}), \ldots, P_N(\mathbf{h})]^T$, $\mathbf{h} \triangleq [h_1, \ldots, h_N]^T$, $\langle \mathbf{x} \rangle \triangleq (1/M)\sum_{i=1}^{M} x_i$ where M is the length of the vector $\mathbf{x}$, $\Pr(x)$ denotes probability of the event x, $\succeq$ denotes component-wise inequality. and

$$D(\mathbf{P}(\mathbf{h}),\mathbf{h}) = \sigma_\theta^2 \left(1 + \sum_{i=1}^{N} \frac{P_i(\mathbf{h}) h_i U_i}{P_i(\mathbf{h}) h_i V_i + C_i (\sigma_{C2}^i)^2}\right)^{-1} \tag{2.14}$$

is the distortion achieved at FC as a function of the channel gains (random) and CH transmission power, which are also functions of the channel gains. Note that we assume instantaneous channel knowledge at FC (receiver) and at the transmitters (CH's) (where the transmitter CSI can be accurately obtained via feedback channels which are error-free and have zero delay).

2.3.2 Solution and optimal power allocation scheme

The problem given in (2.13) can be solved in the same way as in [27]. We first consider the following minimization problem given as

$$\begin{aligned} \min \quad & \langle \mathbf{P}(\mathbf{h}) \rangle \\ \text{s.t.} \quad & D(\mathbf{P}(\mathbf{h}),\mathbf{h}) \leq D_{max} \\ & \mathbf{P}(\mathbf{h}) \succeq 0. \end{aligned} \tag{2.15}$$

The above problem seek to find the minimum average transmission power required to meet the distortion constraint for FC to receive a single measurement from all CHs. We assume full instantaneous CSI is available, hence $\mathbf{h}$ is a known vector quantity and is regarded here simply as a constant vector.

We have the following lemma:

Lemma 2.3.1. *Without loss of generality, assume $h_1 \geq h_2 \geq \ldots \geq h_N$. With the*

knowledge of $\mathbf{h}$*, the solution for (2.15) has already been given in (2.10). Hence the ith optimal power is given as*

$$P_i^*(\mathbf{h}) = \frac{C_i G_i}{\bar{H}_i}\left[\frac{\sqrt{\bar{\eta}_i}}{\bar{\rho}_0(\mathbf{h}, N_1)} - 1\right]^+, \quad i = 1, \ldots, N \tag{2.16}$$

where N_1 *is a unique integer in* $\{1, \ldots, N\}$ *required to evaluate* $\bar{\rho}_0(\mathbf{h}, N_1)$*.* $\bar{H}_i$*,* $\bar{\eta}_i$ *and* $\bar{\rho}_0(\mathbf{h}, N_1)$ *are defined similarly to the static channel case except that* ζ_i *is now replaced by* h_i *and the explicit dependence on* $\mathbf{h}$ *is shown. Note also that* $[x]^+$ *denotes* $\max(x, 0)$*.*

One can also obtain the following Lemma which is necessary to find the final optimal solution of problem (2.13).

Lemma 2.3.2. *The optimal power function,* $\mathbf{P}^*(\mathbf{h}) \triangleq (P_1^*(\mathbf{h}), ..., P_N^*(\mathbf{h}))$ *, is a continuous function of* $\mathbf{h}$*. Furthermore,* $\langle \mathbf{P}^*(\mathbf{h}) \rangle$ *is a non-increasing function of* h_i*,* $i = 1, \ldots, N$*.*

The proof of this lemma, as well as other lemma and theorems, can be found in the appendix.

We define two regions, $\mathcal{R}(s)$ and $\overline{\mathcal{R}}(s)$ and the boundary surface $\mathcal{B}(s)$ for some non-negative s as in [27]:

$$\begin{aligned} \mathcal{R}(s) &= \{\mathbf{h} \in \mathbb{R}_+^N : \langle \mathbf{P}(\mathbf{h}) \rangle < s\} \\ \overline{\mathcal{R}}(s) &= \{\mathbf{h} \in \mathbb{R}_+^N : \langle \mathbf{P}(\mathbf{h}) \rangle \leq s\} \\ \mathcal{B}(s) &= \{\mathbf{h} \in \mathbb{R}_+^N : \langle \mathbf{P}(\mathbf{h}) \rangle = s\} \end{aligned} \tag{2.17}$$

We then define two average power sums as

$$\begin{aligned} P(s) &= \int_{\mathcal{R}(s)} \langle \mathbf{P}(\mathbf{h}) \rangle dF(\mathbf{h}) \\ \overline{P}(s) &= \int_{\overline{\mathcal{R}}(s)} \langle \mathbf{P}(\mathbf{h}) \rangle dF(\mathbf{h}) \end{aligned} \tag{2.18}$$

where $F(\mathbf{h})$ denotes the joint c.d.f. (cumulative distribution function) of $\mathbf{h}$. Finally,

the power sum threshold s^* and the weight w^* are given as

$$\begin{aligned} s^* &= \sup\{s : P(s) < \mathcal{P}_{av}\} \\ w^* &= \frac{\mathcal{P}_{av} - P(s^*)}{\overline{P}(s^*) - P(s^*)} \end{aligned} \tag{2.19}$$

With the above lemma and definitions we can now present the solution to (2.13). The proof follows using similar techniques as in [27] and is excluded.

Theorem 2.3.1. *The solution of problem (2.13) is given as*

$$\hat{\mathbf{P}}(\mathbf{h}) = \begin{cases} \mathbf{P}^*(\mathbf{h}), & \text{if } \mathbf{h} \in \mathcal{R}(s^*) \\ \mathbf{0}, & \text{if } \mathbf{h} \notin \overline{\mathcal{R}}(s^*) \end{cases} \tag{2.20}$$

while if $\mathbf{h} \in \mathcal{B}(s^*)$, $\hat{\mathbf{P}}(\mathbf{h}) = \mathbf{P}^*(\mathbf{h})$ *with probability* w^* *and* $\hat{\mathbf{P}}(\mathbf{h}) = 0$ *with probability* $1 - w^*$.

The optimal power allocation scheme states that if the channel condition is above some threshold then CHs transmit with power allocation given by (2.16), or else none should transmit to save power.

Remark 2.3.1. Note that the solution given in (2.20) is of a general form, which can be applied to both continuous and discontinuous fading distributions. If the fading distribution is continuous (which is true for this problem), then the probability that $\mathbf{h} \in \mathcal{B}(s^*)$ is zero, hence discarding the need for randomization at the boundary.

Remark 2.3.2. Note also that while the computations necessary to implement the above solutions are carried out at FC (such as those of s^* (based on $\mathcal{P}_{av}$) and $\bar{\rho}_0(\mathbf{h}, N_1)$) and the decision whether CH's should transmit or not transmit can be broadcast by FC, the optimal power allocation for individual CH can be easily implemented in a distributed fashion (in the case where the CH's transmit). The fusion center has to just broadcast the quantity $\bar{\rho}_0(\mathbf{h}, N_1)$ to all CH's and CH's can then update their transmission power according to (2.16) which only involves local variables at CH's (apart from $\bar{\rho}_0(\mathbf{h}, N_1)$).

2.3.3 Simulation results

Two sensor network topologies are simulated in MATLAB. Topology A has six clusters deployed equally spaced around the source and Topology B deploys six clusters on one side of the source only as shown in Fig. 2.2. Topology B models environments where it is difficult or impossible to deploy sensors in certain parts of the landscape, for example, when the source is located at the edge of a cliff. The sensors in each cluster are organized in four equally spaced concentric circles and the number of sensors in each circle are 6,12,18 and 24 from the smallest to the biggest circle respectively. All clusters have a radius of 40m. All sensors transmit with $q_i/M_i = 1m$W in topology A. In topology B sensors transmit with $1.33m$W, $1m$W and $0.67m$W in the two clusters closest to the source, two clusters second closest and two clusters farthest away from the source respectively. The clusterheads are located at the center of each cluster for simplicity. CHs are 100m and 60m apart from the next closest CH in topology A and B respectively. The fusion center is located 500m away from the source in both topologies. The channel noise variances are set to $(\sigma_{C1}^i)^2 = 10^{-12}$ Watts and $(\sigma_{C2}^i)^2 = 10^{-10}$ Watts for $i = 1, \ldots, 6$ in the first and second stage of transmission respectively in both topologies. Source variance is set to $\sigma_\theta^2 = 1$ Watt.

Static channel

Fig. 2.3 shows total power consumption, $\sum_{i=1}^{N}(P_i + q_i)$, versus total sensor power within clusters, $\sum_{i=1}^{N} q_i$, in topology A. In this simulation only the total sensor power of one of the six clusters is varied. As $\sum_{i=1}^{N} q_i$ increases, more power is allocated to the sensors, and hence signals received at CHs have a lower distortion. Therefore, CHs can transmit with less power to achieve the same distortion. However total power starts to increase after some point of $\sum_{i=1}^{N} q_i$ since allocating extra q_i cannot bring down the distortion anymore and this power is wasted. Asymptotic analysis

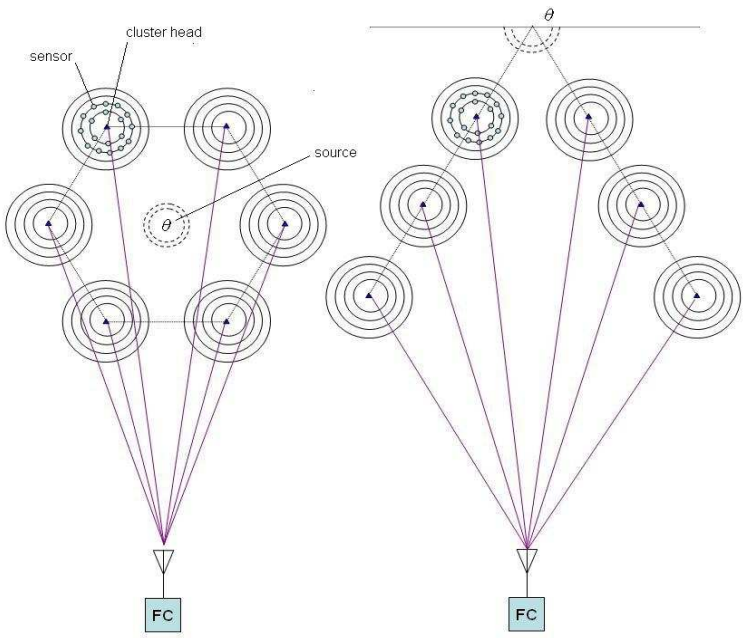

Figure 2.2: Wireless sensor network topologies. Left: topology A. Right: topology B.

shows that as q_i goes to infinity $\forall i$, distortion is given as

$$\lim_{q_i \to \infty} D = \sigma_\theta^2 \left(1 + \sum_{i=1}^{N} \frac{\left(\sum_{m=1}^{M_i} \sqrt{\frac{g_m^i}{1+(\gamma_m^i)^{-1}}} \right)^2}{\sum_{m=1}^{M_i} \frac{g_m^i (\sigma_m^i)^2}{1+(\sigma_m^i)^{-1}}} \right)^{-1} . \tag{2.21}$$

This is in fact the same expression as the minimum distortion achievable at FC if $\beta_i^2 \to \infty \ \forall i$, and characterizes the feasible set of the distortion constraint for the optimization problem in (2.9).

Fig. 2.4 shows how the number of sensors per cluster affects sum power while keeping q_i fixed. As the number of sensors per cluster increases (while keeping q_i fixed), more observations are transmitted to CH. This lowers the distortion at CH and hence CHs need less power to meet the distortion requirement at the fusion center. The expression of taking the limit of distortion as the number of sensors,

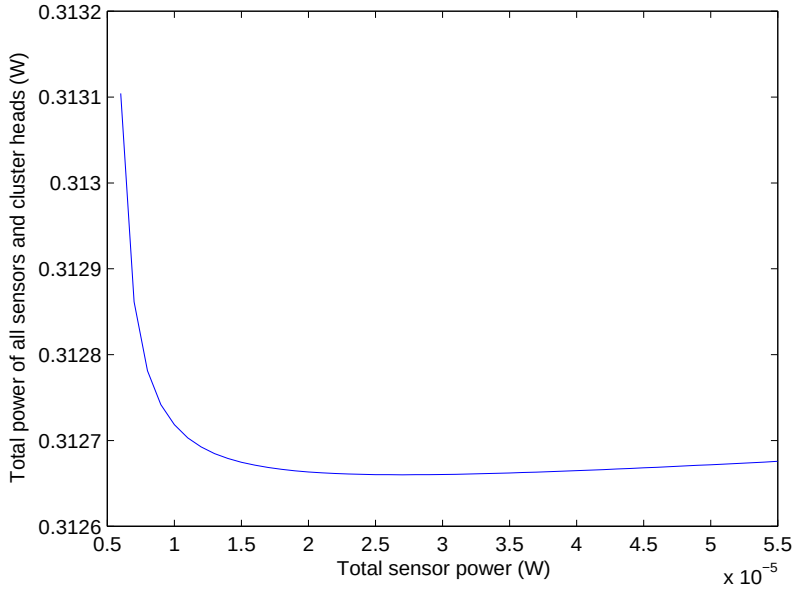

Figure 2.3: Total clusterhead power versus distortion (Topology A).

M_i, goes to infinity is given as

$$\lim_{M_i \to \infty} \sigma_\theta^2 \left(1 + \sum_{i=1}^{N} \frac{\left(\sum_{m=1}^{M_i} \sqrt{\frac{g_m^i}{1+(\gamma_m^i)^{-1}}} \right)^2}{\sum_{m=1}^{M_i} \frac{g_m^i (\sigma_m^i)^2}{1+(\sigma_m^i)^{-1}}} \right)^{-1} . \tag{2.22}$$

Numerical analysis shows that the quantity (distortion) in (2.22) decreases like $1/M_i$ which conforms with the asymptotic analysis given in [20].

Fading channel

In this section, the channels between CH's and FC are modeled as Rayleigh-faded channels. The following results are obtained over 1,000,000 realizations of exponentially-distributed channel power gains of mean equal to the inverse of the distance squared for each average power given. The distortion requirement is set to 0.0043, which is a hundred times the minimum achievable distortion.

Fig. 2.5 shows $\mathcal{P}_{av}$ versus s^* (the sum power threshold that determines whether CH's should transmit or not). This graph allows us to obtain the value of s^* that corresponds to a given $\mathcal{P}_{av}$.

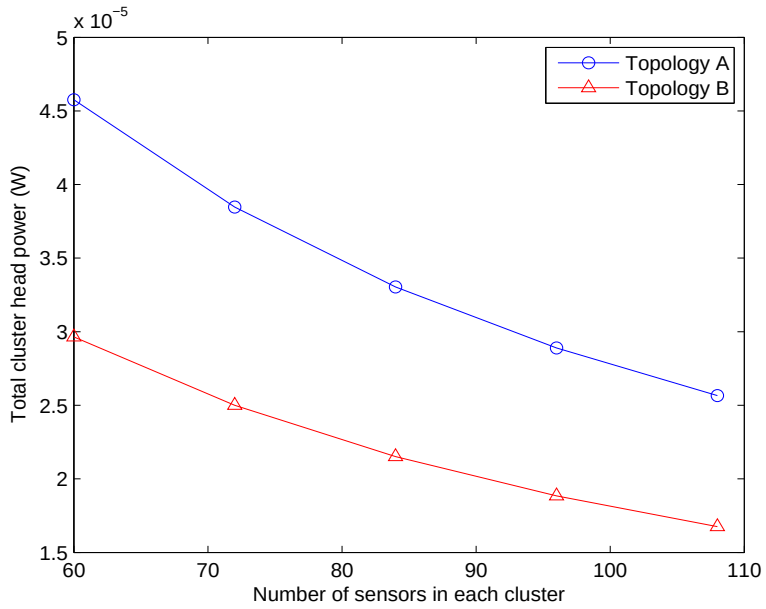

Figure 2.4: Total clusterhead power versus number of sensors per cluster.

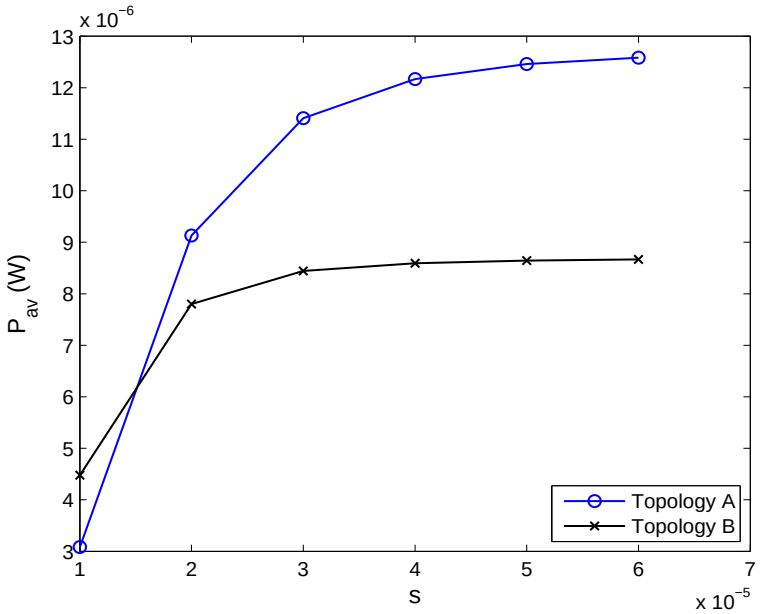

Figure 2.5: $\mathcal{P}_{av}$ against s.

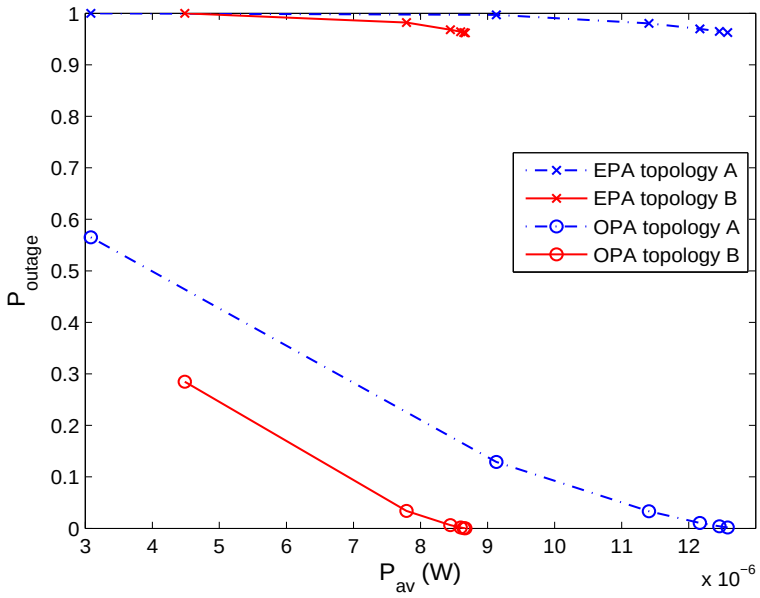

Figure 2.6: P_{outage} against $\mathcal{P}_{av}$.

Fig. 2.6 shows the distortion outage probability against average power for optimal power allocation (OPA) and equal power allocation (EPA). EPA allocates all CHs with equal transmitting power, which equals $\mathcal{P}_{av}$. As shown in this figure, optimal power allocation scheme performs significantly better than EPA scheme for both network topologies. In the (simplified) problem formulation we assumed that the sensors can only transmit with a finite number of power levels and hence q_i is no longer a variable of optimization. Here we investigate the effect of q_i on the outage performance via simulation. Fig. 2.7 shows how the outage probability varies with q_i using optimal power allocation in topology A (essentially total power consumed by sensor transmissions in all clusters are kept at the same value q_i). D_{max} is set to a hundred times the minimal achievable distortion. As q_i increases, the outage probability obviously decreases. However,the effect of lowering the outage probability by increasing q_i quickly saturates when q_i reaches around -70dBW; any q_i higher than this power level does not lower the outage probability significantly. This is because adjusting q_i only affects the first stage of transmission and the resulting distortion achieved at CHs. The saturation level outage probability then depends on the channel conditions in the second stage of transmission (clusterheads to the fusion center). One can similarly plot the outage probability versus $\mathcal{P}_{av}$ for various

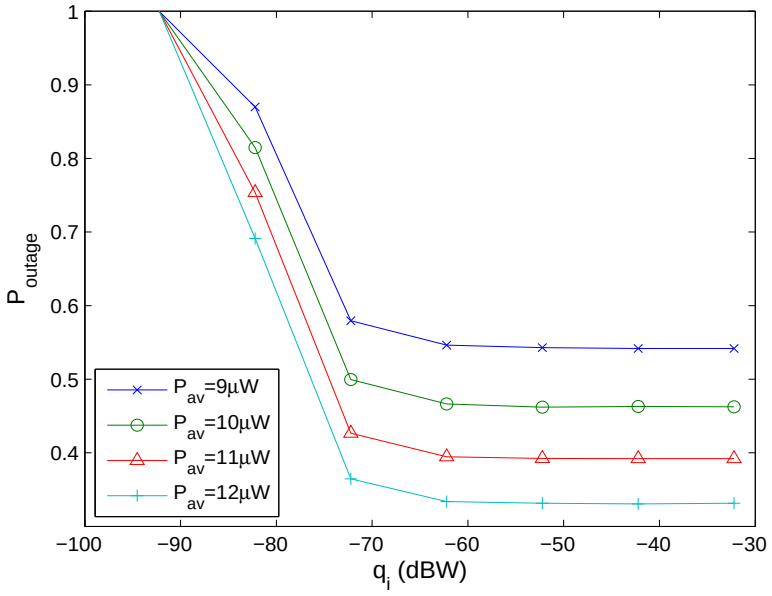

Figure 2.7: P_{outage} against q_i with different long-term average power using optimal power allocation in topology A ($D_{max} = 0.0043$).

values of q_i. While increasing average transmit power for the clusterheads decreases the outage probability, increasing q_i beyond a certain level does not result in any significant reduction in the outage probability.

Power allocation based on statistics of the fading channels

As acquiring full instantaneous channel knowledge at the clusterhead transmitters can be costly, here we look at some optimal power allocation methods based on statistical knowledge of the fading channels between the clusterheads and the fusion center. Since the fading statistics do not necessarily vary rapidly with time, this requires very little overhead communication between the clusterheads and the fusion center. It is however difficult to obtain an explicit expression of the outage probability for $N > 1$ (an observation which was also made in [22]). Hence we choose to minimize an upper bound on the outage probability by minimizing the expected distortion, which is motivated by Markov's inequality $\Pr(D > D_{max}) = \Pr(D \geq D_{max}) \leq \frac{E[D]}{D_{max}}$. To simplify the analysis even further, we obtain an approximation to the expected

distortion by obtaining a lower bound on it, which is given by

$$\begin{aligned}
&\min E\left[\frac{1}{\sigma_\theta^2}+\frac{1}{\sigma_\theta^2}\sum_{i=1}^{N}\frac{U_i\beta_i^2h_i}{V_i\beta_i^2h_i+\delta_i}\right]^{-1}\\
&\quad\geq \min\left(\frac{1}{\sigma_\theta^2}+\frac{1}{\sigma_\theta^2}\sum_{i=1}^{N}\frac{U_i\beta_i^2E[h_i]}{V_i\beta_i^2E[h_i]+\delta_i}\right)^{-1}\\
&\quad\equiv \sigma_\theta^2(1+\max\sum_{i=1}^{N}\frac{U_i\beta_i^2E[h_i]}{V_i\beta_i^2h_iE[h_i]+\delta_i})\\
&\quad\geq \sigma_\theta^2(1+\max E\left[\sum_{i=1}^{N}\frac{U_i\beta_i^2h_i}{V_i\beta_i^2h_ih_i+\delta_i}\right])
\end{aligned}\tag{2.23}$$

where $\delta_i = (\sigma_{C2}^i)^2$. The inequalities in the above analysis follow from Jensen's inequality due to convexity of the distortion function with respect to the channel gains. Although the distortion outage probability may not be strictly upper bounded by this lower bound on the expected distortion, it provides a heuristic for obtaining a statistical power allocation scheme.

We can now solve an optimization problem by minimizing the lower bound of expected distortion given by the last line of (2.23) as

$$\begin{aligned}
\max\quad & E\left[\sum_{i=1}^{N}\frac{U_i\beta_i^2h_i}{V_i\beta_i^2h_i+\delta_i}\right]\\
\text{s.t.}\quad & \sum_{i=1}^{N}\beta_i^2C_i\leq\mathcal{P}_{tot}\\
& \beta_i^2\geq 0\quad\forall i.
\end{aligned}\tag{2.24}$$

For Rayleigh-faded channels, the objective function can be expressed as

$$\begin{aligned}
E\left[\sum_{i=1}^{N}\frac{U_i\beta_i^2h_i}{V_i\beta_i^2h_i+\delta_i}\right] &= \frac{U_i}{V_i}-\frac{U_i\delta_i}{V_i}E\left[\frac{1}{V_i\beta_i^2h_i+\delta_i}\right]\\
&= K_{1i}-\frac{U_i\delta_i}{V_i}\int_0^\infty\frac{\lambda_ie^{-\lambda_ih_i}}{V_i\beta_i^2h_i+\delta_i}dh_i\\
&= K_{1i}-\frac{K_{2i}}{\beta_i^2}e^{\frac{K_{3i}}{\beta_i^2}}E_1\left(\frac{K_{3i}}{\beta_i^2}\right)
\end{aligned}$$

where $K_{1i}=U_i/V_i$, $K_{2i}=U_i\lambda_i\delta_i/V_i^2$, $K_{3i}=\lambda_i\delta_i/V_i$ and $E_1(z)=\int_z^\infty e^{-t}/t\,dt$. Hence

the optimization problem is given as

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} \left(\frac{K_{2i}}{\beta_i^2} e^{\frac{K_{3i}}{\beta_i^2}} E_1\left(\frac{K_{3i}}{\beta_i^2}\right) - K_{1i} \right) \\ \text{s.t.} \quad & \sum_{i=1}^{N} \beta_i^2 C_i \leq \mathcal{P}_{tot} \\ & \beta_i^2 \geq 0 \quad \forall i. \end{aligned} \tag{2.25}$$

It can be easily shown that this problem is a standard convex optimization problem and by solving the KKT (Karush-Kuhn-Tucker) conditions and letting $z_i = K_{3i}/{\beta_i}^2$, we get the following set of nonlinear equations

$$\begin{cases} \mu > 0 \\ \sum_{i=1}^{N} \frac{K_{3i} C_i}{z_i} - \mathcal{P}_{tot} = 0 \\ z_i^2 \left[e^{z_i} E_1(z_i)(1+z_i) - 1 \right] = \frac{C_i K_{3i}^2}{K_{2i}} \mu, \quad 0 \leq z_i < \infty \\ z_i^2 \left[e^{z_i} E_1(z_i)(1+z_i) - 1 \right] \leq \frac{C_i K_{3i}^2}{K_{2i}} \mu, \quad z_i = \infty \end{cases} \tag{2.26}$$

where μ is the Lagrangian multiplier. The optimal power values can be obtained by solving the above set of nonlinear equations numerically by using provably convergent fixed point iterative methods.

We can also look at minimizing the lower bound on expected distortion given by the third line of (2.23), which is equivalent to problem (2.11). Fig. 2.8 shows the outage probability achieved by problem (2.24) (heuristic method 1) and problem (2.11) (heuristic method 2) for Topology A. Clearly, the sub-optimal statistical power allocation methods based on minimizing the upper bounds on the outage probability do not fare well compared to the performance of the optimal power allocation method based on full CSI at the clusterhead transmitters. A similar observation was also made in [29] in the context of outage probability performance of beamforming in multiple antenna systems. This motivates the need for optimal power allocation for distortion outage minimization based on quantized or finite rate channel feedback from the fusion center to the clusterheads, a topic that is begin investigated in the following section.

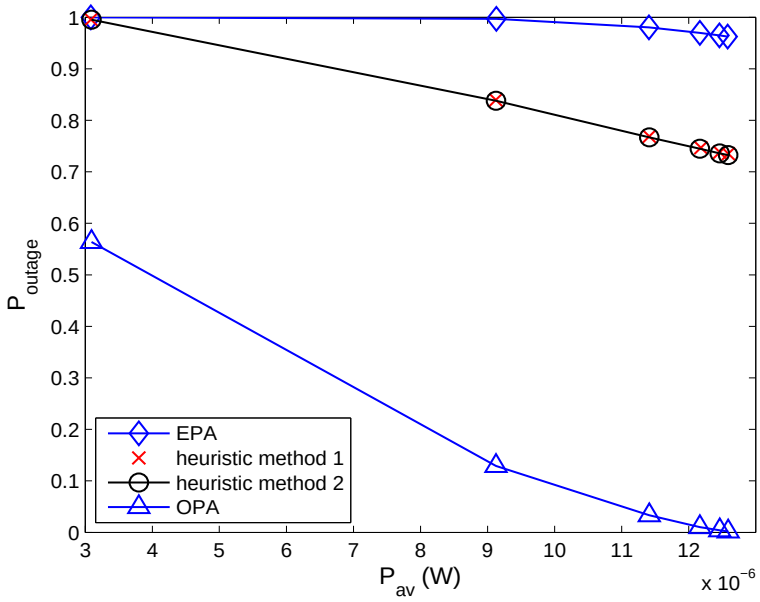

Figure 2.8: Performance of heuristic methods that use knowledge of channel statistics (topology A).

2.4 Power control schemes in Nakagami-m fading using limited feedback

In the previous section (Section 2.3) we have presented the optimal power allocation algorithm that minimizes the distortion outage probability with the assumption that CSIT and CSIR (CSI at the receiver) are known. In this section we will assume that only partial CSIT is available. The type of partial CSIT considered here is in the form of quantized (or rate-limited) feedback.

Following the assumption made in Section 2.3.1 that all sensors within a cluster transmit with equal power (q_n/M_n), the expressions for sensor power gain, CH power and distortion have already been obtained in Section 2.3.1 and are given in (2.5), (2.6) and (2.7) respectively.

We assume that the channels between the CHs and the FC are stationary ergodic and subject to independent Nakagami-m block-fading, and hence the channel power gain $h_n \in \Re^+$ is distributed according to a gamma distribution with a mean equal to the inverse of the square of the transmission distance. In other words, the p.d.f.

of $h_i, i = 1, 2, \ldots, N$, is given as

$$f_i(h_i) = \frac{(m_i\lambda_i)^{m_i} h_i^{m_i-1}}{\Gamma(m_i)} e^{-m_i\lambda_i h_i}, \quad i = 1, \ldots, N \tag{2.27}$$

where $\frac{1}{\lambda_i}$ is the mean channel power gain and $m_i \geq 0.5$ is a real parameter that indicates the severity of the fading. $\Gamma(\cdot)$ is the Gamma function defined as $\Gamma(m) = \int_0^\infty t^{m-1}e^{-t}dt$. We include a subscript i in $f_i(\cdot)$ because the distributions are independent but not identical. For the special case of Rayleigh-fading, the channel power gain is exponentially distributed given by $f_i(h_i) = \lambda_i e^{-\lambda_i h_i}$ which can be easily obtained by substituting $m_i = 1$ into (2.27).

2.4.1 Problem formulation

In wireless sensor networks with rate-limited feedback links, only a finite set of power values can be transmitted from the receiver (FC) to the transmitters (CHs). We denote the collection of this finite set of power values as a power codebook $\mathcal{P}^{(N,L)}$ where N and L are the number of CH transmitters and the number of power levels respectively. It is often more practical to convert L into R binary bits using the relationship $L = 2^R$ and refer to the unit of feedback resolution in terms of bits. For an R-bit broadcast feedback channel and N clusters in the network, we quantize the vector channel space $\Re_+^N$ into L regions. Denote the regions as $\mathcal{R}_j^{(N)}$ and the power codeword associated with the j-th quantized region as $\mathcal{P}_j^{(N)} \in \mathcal{P}^{(N,L)}$, $j = 1, \ldots, L$. Furthermore, the j-th region power codeword $\mathcal{P}_j^{(N)} = [P_{1,j}, \ldots, P_{N,j}]^T$ contains a set of N power values specifying the CH transmit powers. We assume that CHs and FC know this (pre-computed) power codebook, since this power codebook can be computed offline, purely based on the channel statistics and the available average power. We will first present the single-cluster network problem formulation as it is simple and provides some useful intuitions and properties that will be useful later in formulating the multi-cluster problem.

Power allocation with quantized CSIT with a single cluster ($N = 1$)

arbitrary power codebook $\mathcal{P}^{(1,L)} = [P_{1,1}, \ldots, P_{1,L}]^T$ assigned deterministically to L quantization regions in $h_1 \in \Re^+$, that is whenever h_1 belongs to the j-th quantization region, the CH uses the transmission power $P_{1,j}$ with probability one. Without loss of generality, we assume that $P_{1,1} > \ldots > P_{1,L} \geq 0$. Before we define the quantization regions, we need to first state a property that the optimal quantizer (one that minimizes the outage probability) possesses. Note that when $N = 1$ it can be easily shown that the distortion and the outage probability are monotonically decreasing functions of power. These two properties are the same as the problems studied in [35–37] and hence it can be easily shown in a similar fashion that the optimal (deterministic) index mapping achieving minimum outage probability also has a circular structure (one that wraps around) as in [35–37]. It is straightforward to show that, for a given fading block, in the case of non-outage, the index is assigned to the minimum power that can meet the distortion threshold, and in the case of outage, which occurs when none of the power in the power codebook can meet the distortion threshold, the index is assigned to the smallest power. We now introduce a set of channel thresholds defining the boundaries of the quantized channel regions as an alternative for defining the problem instead of power simply because it is easier to define the cumulative distribution function (c.d.f.) for the fading distribution and the outage probability in terms of the channel thresholds. However through out this chapter we may use channel thresholds and power interchangeably, depending on whichever is more convenient in the given context. The channel thresholds are one-to-one functions of the quantized power values, given as $s_{1,j} = \phi_1/P_{1,j}$ where $\phi_1 = C_1(\sigma_{C2}^1)^2\gamma_{th}/(U_1 - V_1\gamma_{th})$ and $\gamma_{th} = \sigma_\theta^2/D_{max} - 1$. For notational completeness we denote $\mathcal{S}^{(1,L)} = \{s_{1,1}, \ldots, s_{1,L}\}$ (the superscript '1' denotes $N = 1$ and L denotes that there are L power feedback levels or quantization regions). Denote the regions as $R_j^{(1)}$, $j = 1, \ldots, L$ (the superscript indicates $N = 1$). The circular index mapping allows us to naturally define $R_j^{(1)} = [s_{1,j}, s_{1,j+1})$, $j = 1, \ldots, L-1$, $R_L^{(1)} = \{[0, s_{1,1}), [s_{1,L}, \infty]\}$ and the outage region $R_{out}^{(1)} = [0, s_{1,1})$. Note that $R_{out}^{(1)} \subseteq R_L^{(1)}$. Let $F_1(x) \triangleq \Pr\{0 < h_1 \leq x\}$ denote the cdf of the channel

gain for $N = 1$. Note that the outage probability is then simply given by $F_1(s_{1,1})$. The problem of minimizing the outage probability subject to a long-term average power constraint can then be formulated as

$$\begin{aligned} \min \quad & F_1(s_{1,1}) \\ \text{s.t.} \quad & \sum_{j=1}^{L-1} P_{1,j}\left[F_1(s_{1,j+1}) - F_1(1, s_{1,j})\right] + P_{1,L}(1 - F_1(s_{1,L}) + F_1(s_{1,1})) \leq P_{av} \\ & 0 < s_{1,j} < s_{1,j+1} \quad \forall j = 1, 2, \ldots, L-1 \end{aligned} \tag{2.28}$$

Power allocation with quantized CSI when $N \geq 2$

We begin by first illustrating the complexity in the structure of quantization regions for $N \geq 2$ through an example. Fig. 2.9 shows the quantization regions of a suboptimal solution for $N = 2$ and $L = 4$ obtained by using iterative Lloyd's algorithm incorporating a simulation based randomized optimization method called SPSA (simultaneous perturbation stochastic approximation [38]), where the first step of the algorithm finds the optimal channel partitions for a given set of quantized power values and the second step uses SPSA to find a locally optimal set of quantized power values for these channel partitions using SPSA . These two steps are iterated until a satisfactory convergence criterion is met. For more details on this algorithm and SPSA as a stochastic optimization tool, see Section 2.4.2) where we provide this SPSA based algorithm that has a superior performance compared to our quantized power allocation algorithms, but at the cost of a high computational complexity. We can see from Fig. 2.9 the irregularity in the way the regions can be formed already for $N = 2$ and $L = 4$. In the general case with N (≥ 2) cluster network with L-level power feedback, the optimal quantizer is unknown. Hence in order to make the quantized power allocation problem for distortion outage minimization analytically tractable, we impose a restriction on the ordering of the powers. This restriction gives the quantization regions a certain structure which can be exploited for analytical tractability, at the cost of a small performance loss.

Recall that the power codewords of a (N,L) power codebook are given by $\mathcal{P}_j^{(N)} = [P_{1,j}, \ldots, P_{N,j}]^T$, $j = 1, \ldots, L$. We assume the restriction in ordering of the power

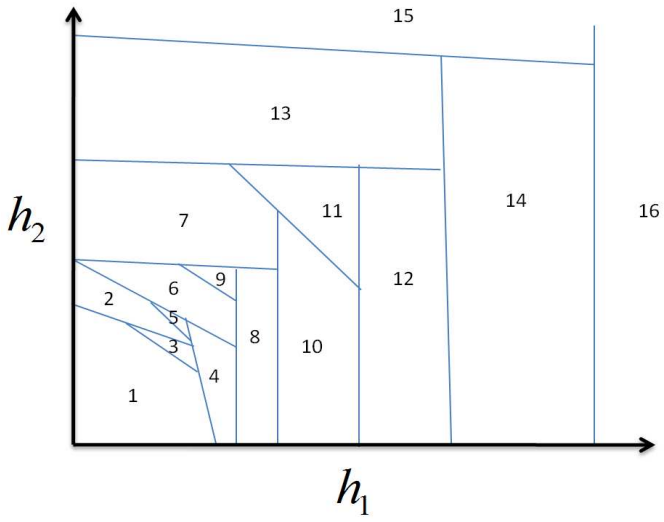

Figure 2.9: Quantization regions when $N = 2$, $L = 4$ using Lloyd's algorithm with SPSA.

codeword given as $\mathcal{P}_1^{(N)} \succ \ldots \succ \mathcal{P}_L^{(N)}$ where $\succ$ denotes component-wise inequality. We first show, in a similar way to [37], that the optimal (deterministic) index mapping that achieves the minimum outage probability for $N \geq 2$ also has a circular structure. The component-wise inequality of the power codeword implies that $\Lambda_1 > \ldots > \Lambda_L$ where $\Lambda_j = \sum_{i=1}^{N} P_{i,j}$, $j = 1, \ldots, L$. Note also that distortion and the outage probability are monotonically decreasing functions of $P_{i,j}$. We are interested in finding an index mapping scheme that achieves the minimum outage probability subject to a long-term average power constraint. We first consider the set of channel gains which are not in outage with a non-zero probability measure: $\mathcal{S} = \{\mathbf{h} : D(\mathcal{P}_1^{(N)}, \mathbf{h}) \leq D_{max}\}$. The optimal index mapping strategy for a channel $\mathbf{h}$ in this set is for the receiver to feed back an index i such that $D(\mathcal{P}_i^{(N)}, \mathbf{h}) \leq D_{max}$ and $D(\mathcal{P}_{i+1}^{(N)}, \mathbf{h}) > D_{max}$. Denote by $\mathcal{I}$ the set of channel realizations that get assigned to the index i. Now assume the contrary, that it is optimal to feed back some $j \neq i$ for $\mathbf{h} \in \mathcal{H} \subseteq \mathcal{I}$ where $\mathcal{H}$ has a non-zero probability measure. If $j < i$, construct a new scheme that maps all elements of $\mathcal{H}$ to i instead. The newly constructed scheme clearly uses less average power since $\Lambda_i < \Lambda_j$ while the outage probability remains the same. If $j > i$, we see that an outage also occurs for $\mathbf{h} \in \mathcal{H}$. Thus

the corresponding outage has increased, which is a contradiction to the assumption that $j \neq i$ is optimal. Now consider the set of channels in outage, namely $\{\mathbf{h} : D(\mathcal{P}_1^{(N)}, \mathbf{h}) > D_{max}\}$ with a non-zero probability measure. It is easy to see that the optimal feedback index should be L since it is the one that results in the smallest average power consumption while achieving the same outage probability, since $\Lambda_L < \Lambda_j \ \forall j < L$.

To illustrate the structure of the quantization regions under the above-mentioned restriction on the quantized power values, we give an example of an $N = 2$ network with $R = \log_2 L$-bit feedback in Fig. 2.10. Similar to the $N = 1$ case, we quantize the channel space into L regions according to a circular quantization structure. The regions are defined as $\mathcal{R}_j^{(N)} = \{\mathbf{h} : D(\mathcal{P}_j^{(N)}, \mathbf{h}) \leq D_{max} \cap D(\mathcal{P}_{j+1}^{(N)}, \mathbf{h}) > D_{max}\}$ for $j = 1, \ldots, L-1$ and $\mathcal{R}_L^{(N)} = \{\mathbf{h} : D(\mathcal{P}_1^{(N)}, \mathbf{h}) > D_{max} \cup D(\mathcal{P}_L^{(N)}, \mathbf{h}) \leq D_{max}\}$. Denote the boundaries that divide the channel space into L regions as $\mathcal{B}_j(\mathbf{s}_j^{(N)})$ for $j = 1, \ldots, L$, where $\mathbf{s}_j^{(N)} = \{s_{1,j}, \ldots, s_{N,j}\} \in \mathcal{S}^{(N,L)}$. The circular quantizer structure implies that there should only exist a single outage region given by $\mathcal{R}_{out}^{(N)} = \{\mathbf{h} : D(\mathbf{h}, \mathcal{P}_1^{(N)}) > D_{max}\} \subseteq \mathcal{R}_L^{(N)}$. It also implies that $s_{i,j} = \phi_i / P_{i,j}$ where $\phi_i = C_i(\sigma_{C2}^i)^2 \gamma_{th}/(U_i - V_i\gamma_{th})$. In order to ensure no outage exists outside the set $\mathcal{R}_{out}^{(N)}$ defined above, the distortion must be constant and equal to D_{max} on all the boundaries between any two quantized regions. This allows us to easily write down the expressions that define the boundaries $\mathcal{B}_j(\mathcal{P}_j^{(N)}) : D_{max} = \sigma_\theta^2 \left(1 + \sum_{i=1}^N \frac{P_{i,j} h_i U_i}{P_{i,j} h_i V_i + C_i(\sigma_{C2}^i)^2}\right)^{-1}$ after substituting $P_{i,j} = C_i \beta_{i,j}^2$. We also call the boundaries as distortion curves for this reason.

With this quantizer structure, we are interested in minimizing the distortion outage probability subject to a long-term average power constraint in the vector channel quantization space. Defined $F_N(\mathbf{s}_j^{(N)}) \triangleq \Pr(\mathbf{h} \prec B_j)$ where the set $\{\mathbf{h} \prec B_j\} \triangleq \{\mathbf{h} : D(\mathbf{h}, \mathcal{P}_j^{(N)}) > D_{max}\}$. The quantized power allocation problem for outage minimization for this quantizer structure for N-clusters and R-bit feedback

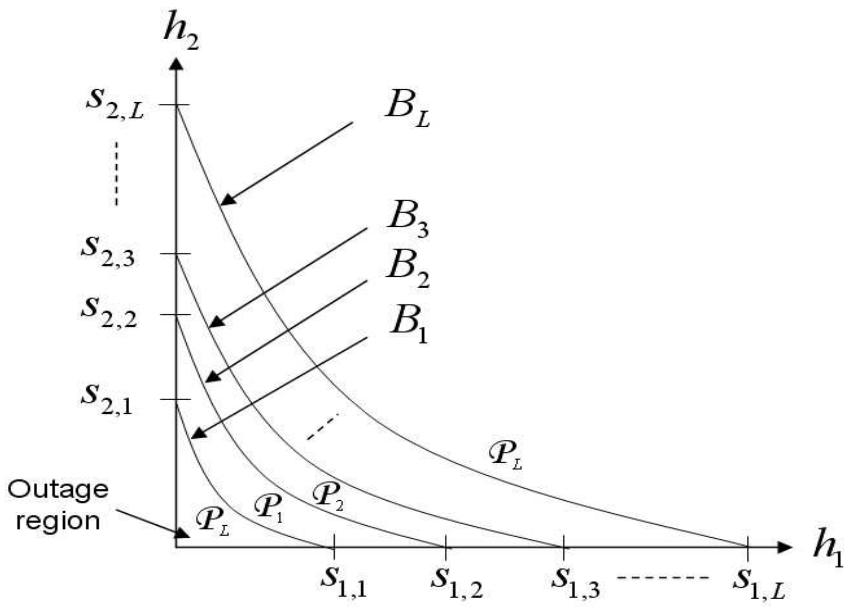

Figure 2.10: Vector channel quantization regions formed by a series of distortion curves for a 2-cluster network.

is given by

$$\begin{aligned} \min \quad & F_N(\mathbf{s}_1^{(N)}) \\ \text{s.t.} \quad & \sum_{j=1}^{L-1} \Lambda_j \left[F_N(\mathbf{s}_{j+1}^{(N)}) - F_N(\mathbf{s}_j^{(N)}) \right] + \Lambda_L \left[1 - F_N(\mathbf{s}_L^{(N)}) + F_N(\mathbf{s}_1^{(N)}) \right] \leq N P_{av} \\ & 0 \leq s_{i,j} \leq s_{i,j+1} \quad \forall i, j. \end{aligned} \tag{2.29}$$

where $\Lambda_j = \sum_{i=1}^{N} P_{i,j}$ denotes the elementwise sum of the power codeword $\mathcal{P}_j^{(N)}$.

2.4.2 Solution and optimal power allocation scheme

Problem (2.29) is non-convex in general, but we can find a locally optimal solution using the standard Lagrange multiplier based optimization technique and the associated Karush-Kuhn-Tucker (KKT) necessary optimality conditions. Note that it can be easily shown that the second constraint in (2.29) is satisfied with a strict inequality. We therefore discard this constraint in what follows as it will not affect

the result. The Lagrangian is given by

$$F_N(\mathbf{s}_1^{(N)})+\mu\left[\sum_{j=1}^{L-1}\Lambda_j(F_N(\mathbf{s}_{j+1}^{(N)})-F_N(\mathbf{s}_j^{(N)}))+\Lambda_L(1-F_N(\mathbf{s}_L^{(N)})+F_N(\mathbf{s}_1^{(N)}))-NP_{av}\right] \tag{2.30}$$

where μ is the Lagrange multiplier. For ease of viewing, we write the partial derivatives of the cdf $F_N(\mathbf{s}_j^{(N)})$ and the sum power function Λ_j with respect to any of its variables in $\mathbf{s}_j^{(N)}$ or $\mathcal{P}_j^{(N)}$ as $\partial F_N(\mathbf{s}_j^{(N)})/\partial \mathbf{s}_j^{(N)}$, $\partial \Lambda_j/\partial \mathbf{s}_j^{(N)}$, $\partial F_N(\mathcal{P}_j^{(N)})/\partial \mathcal{P}_j^{(N)}$ and $\partial \Lambda_j/\partial \mathcal{P}_j^{(N)}$

Single-cluster network $(N = 1)$

In this case, the cdf $F_1(s_{1,j})$ can be obtained by integrating (2.27) from 0 to $s_{1,j}$. For Nakagami-m fading, the cdf is given by the regularized lower incomplete Gamma function defined as $F_1(s_{1,j}) = \gamma(m\lambda s_{1,j}, m)/\Gamma(m)$ where $\gamma(x,m) = \int_0^x t^{m-1}e^{-t}dt$ is the incomplete Gamma function.

For Rayleigh fading channels the cdf has a simple closed form expression given as $F_1(s_{1,j}) = 1 - e^{-\lambda s_{1,j}}$ and the KKT conditions (2.28) for $m = 1$ and $P_{1,j} > 0$ are given as

$$\begin{aligned}
&\frac{\lambda e^{-\lambda s_{1,i+1}}}{s_{1,i}} - \frac{e^{-\lambda s_{1,i+1}} - e^{-\lambda s_{1,i+2}}}{s_{1,i+1}^2} - \frac{\lambda e^{-\lambda s_{1,i+1}}}{s_{1,i+1}} = 0, \quad i = 1, \ldots, L-2, \\
&\frac{\lambda e^{-\lambda s_{1,L}}}{s_{1,L-1}} - \frac{1 - e^{-\lambda s_{1,1}} + e^{-\lambda s_{1,L}}}{s_{1,L}^2} - \frac{\lambda e^{-\lambda s_{1,L}}}{s_{1,L}} = 0 \\
&\sum_{i=1}^{L-1} \frac{e^{-\lambda s_{1,i}} - e^{-\lambda s_{1,i+1}}}{s_{1,i}} + \frac{1 - e^{-\lambda s_{1,1}} + e^{-\lambda s_{1,L}}}{s_{1,L}} = \frac{P_{av}}{\phi}.
\end{aligned} \tag{2.31}$$

Note that the last KKT condition relates to the long-term average power constraint which must be met with equality as implied by the optimality condition. Problem (2.31) then can be solved by fixed point iterative methods for solving nonlinear equations or any other suitable nonlinear equation solver. The corresponding equations for Nakagami-m fading can be also solved similarly, we do not include them here to avoid repetition.

Multi-cluster network ($N \geq 2$)

The KKT conditions of (2.29) for $N \geq 2$ and $P_{i,j} > 0$ are given as

$$\begin{aligned}
&\frac{\partial \Lambda_j}{\partial s_{i,j}} \Big/ \frac{\partial F_N(\mathbf{s}_j^{(N)})}{\partial s_{i,j}} = \frac{\partial \Lambda_j}{\partial s_{k,j}} \Big/ \frac{\partial F_N(\mathbf{s}_j^{(N)})}{\partial s_{k,j}} \quad \forall i,k \in \{1,\ldots,N\}, \forall j = 1,\ldots,L \\
&\sum_{j=1}^{L-1} \Lambda_j (F_N(\mathbf{s}_{j+1}^{(N)}) - F_N(\mathbf{s}_j^{(N)})) + \Lambda_L (1 - F_N(\mathbf{s}_L^{(N)}) + F_N(\mathbf{s}_1^{(N)})) = N P_{av} \qquad (2.32) \\
&\mathbf{0} \prec \mathbf{s}_1 \prec \mathbf{s}_2 \prec \ldots \prec \mathbf{s}_L.
\end{aligned}$$

In general, computing the cdfs, namely $F_N(\mathbf{s}_j^{(N)})$ for $N > 1$, involves evaluating multi-dimensional integrals as a function of the distortion curves and cannot be expressed in closed forms. We can however approximate the distortion curve by a straight line (or a hyperplane if $N > 2$) that passes through the same points as the distortion curve does at the axes, shown as the straight line above the distortion curve in Fig. 2.12. We call this approximation the outer-straight-line approximation and denote the ith plane as $\bar{B}_i$. We can also construct another straight line/hyperplane that is parallel to $\bar{B}_i$ and is tangent to B_i, shown by the straight line below the distortion curve in Fig. 2.12. We call this the inner-straight-line approximation and denote the ith plane as $\underline{B}_i$. Simulation results show that these two approximations give very comparable outage performances, hence the rest of the chapter will be based on the outer-straight-line approximation (referred in this chapter simply as the straight-line approximation (SLA)). A visual illustration comparing the actual outage region and the SLA approximation for $N = 3$ is shown in Fig. 2.11. However it is difficult to illustrate what the regions would look like for $N > 3$. The approximated cdf function obtained by SLA is now defined as $\bar{F}_N(\mathbf{s}_j) \triangleq \Pr(\mathbf{h} \prec \bar{B}_j)$. In the literatures, a number of different expressions of the same cdf function exist for Nakagami-m fading. In [39, 40] the cdf is expressed in the form of iterative equations. [41] provides an expression which approximates the multivariate cdf by an equivalent scalar lower regularized incomplete Gamma function. In [42], the cdf is expressed in an integral form. In [43], the cdf is given in the

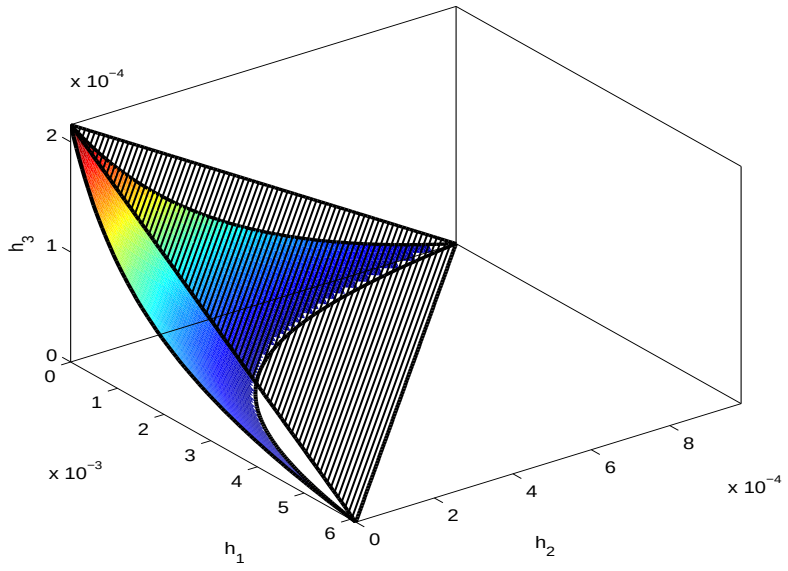

Figure 2.11: Exact outage region and SLA approximation in $\Re_+^3$.

form of an 'infinite-sum-series' representation

$$\bar{F}_N(\mathbf{P}_j, \mathbf{m}) = \frac{\prod_{i=1}^{N} \left(\frac{m_i}{\tilde{\mu}_i}\right)^{m_i}}{\Gamma\left(1 + \sum_{i=1}^{N} m_i\right)} \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \frac{\left[\prod_{i=1}^{N} (m_i)_{n_i} \left(-\frac{m_i}{\tilde{\mu}_i}\right)^{n_i} \frac{1}{n_i!}\right]}{\left(1 + \sum_{i=1}^{N} m_i\right)_{n_T}} \tag{2.33}$$

where $(\alpha)_k = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$, $n_T = \sum_{i=1}^{N} n_i$, $\tilde{\mu}_i = \frac{P_{i,j}}{\phi_i \lambda_i}$ and $P_{i,j} > 0 \ \forall i, j$. The partial derivative of the cdf is given as

$$\frac{\partial \bar{F}_N}{\partial P_{i,j}} = \frac{1}{\phi_i \lambda_i} \left(-\frac{m_i}{\tilde{\mu}_{i,j}} \bar{F}_N - \prod_{k=1}^{N} \left(\frac{m_k \gamma_{th}}{\tilde{\mu}_{k,j}}\right)^{m_k} \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \frac{n_i}{\tilde{\mu}_i} \frac{\prod_{k=1}^{N} \left[(m_k)_{n_k} \left(-\frac{m_k \gamma_{th}}{\tilde{\mu}_k}\right)^{n_k} \frac{1}{n_k!}\right]}{\left(1 + \sum_{k=1}^{N} m_k\right)_{n_T}} \right) \tag{2.34}$$

The KKT conditions shown in (2.32) constitute a set of nonlinear equations, where the number of equations grows exponentially as the number of feedback bits increases. In this section we develop a number of suboptimal algorithms by combining some existing and some newly derived (by us) approximations for special cases of high and low average power, respectively. For moderate to large number of feedback bits we use an exiting approximation called equal average power per region (EPPR)

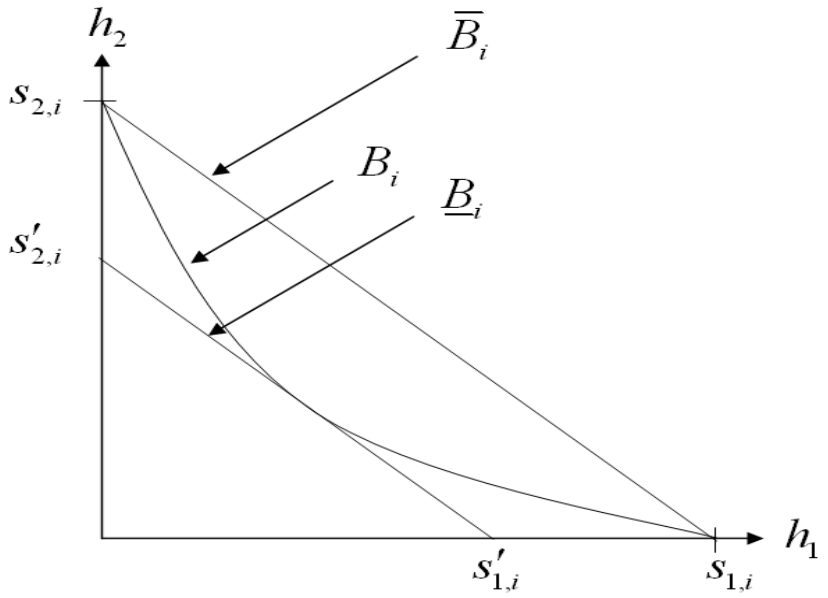

Figure 2.12: Inner and outer straight-line approximations.

derived in [28, 31] using the *Mean Value Theorem* of real analysis. However, before we can write down the problem formulation using this EPPR approximation, we must deal with the issue of whether we should allocate power in the outage region or not. Although it is counter-intuitive to envisage allocating power in the outage region and indeed when full channel information is available, the optimal solution is to not allocate any power in the outage region. This is not true however when quantized channel information is available, as shown in [31, 36] and it is optimal to use the smallest power from the power codebook in the outage region. With a nonzero power in the outage region (NZPOR), the channel space is quantized into L regions including $L-1$ non-outage regions and the Lth region containing a non-outage region as well as an outage region due to the circular nature mentioned earlier. It may be near-optimal however to allocate zero power in the outage region (ZPOR), in the case of very low average power as also noted in [37]. In this case, there would be L regions with $L-1$ non-outage regions and the Lth region containing only the outage region. We observe through simulations that using the EPPR approximation, ZPOR performs nearly optimally when the available average power is very low (the actual threshold below which ZPOR performs near-optimally depends on N and R). This algorithm with EPPR+ZPOR has the added advantage of low complexity of

implementation, as will be evident below. We now provide the problem formulations using EPPR approximation for NZPOR and ZPOR respectively given as

$$\begin{aligned} \min \quad & \bar{F}_N(\mathbf{s}_1) \\ \text{s.t.} \quad & \Lambda_j(\bar{F}_N(\mathbf{s}_{j+1}^{(N)}) - \bar{F}_N(\mathbf{s}_j^{(N)})) = \tfrac{NP_{av}}{L}, \quad j = 1, \ldots, L-1 \\ & \Lambda_L(1 - \bar{F}_N(\mathbf{s}_L) + \bar{F}_N(\mathbf{s}_1)) = \tfrac{NP_{av}}{L} \\ & \mathbf{0} \prec \mathbf{s}_1 \prec \mathbf{s}_2 \prec \ldots \prec \mathbf{s}_L. \end{aligned} \tag{2.35}$$

$$\begin{aligned} \min \quad & \bar{F}_N(\mathbf{s}_1) \\ \text{s.t.} \quad & \Lambda_j(\bar{F}_N(\mathbf{s}_{j+1}^{(N)}) - \bar{F}_N(\mathbf{s}_j^{(N)})) = \tfrac{NP_{av}}{L-1}, \quad j = 1, \ldots, L-2 \\ & \Lambda_{L-1}(1 - \bar{F}_N(\mathbf{s}_{L-1}^{(N)})) = \tfrac{NP_{av}}{L-1} \\ & \mathbf{0} \prec \mathbf{s}_1 \prec \mathbf{s}_2 \prec \ldots \prec \mathbf{s}_{L-1}. \end{aligned} \tag{2.36}$$

The following lemma shows that at high average power and using SLA, one can further simplify the optimal power allocation scheme.

Lemma 2.4.1. *Based on SLA, for Nakagami-m fading with* $\mathbf{m} = [m_1, \ldots, m_N]^T$ *being the fading parameter of each channel, as* $P_{av} \to \infty$*, it is asymptotically optimal to transmit with* $P_{i,j} = \frac{m_i}{m_k} P_{k,j}$*,* $i, k \in \{1, \ldots, N\}$*,* $j = 1, \ldots, L$*. If all the fading parameters are identical, it is asymptotically optimal to transmit with equal transmit power per CH for every quantization region, i.e.,* $P_{i,j} = P_{k,j}$ $\forall i, k \in \{1, \ldots, N\}$*,* $j = 1, \ldots, L$*.*

This proof, as well as proofs of other lemmas and theorems, can be found in the Appendix. Hence problem (2.35) and (2.36) can be further simplified at high average power by letting all CHs transmit with equal power in the case where all m_i are identical. We will abbreviate equal power per CH as EPPC. Each boundary can now be expressed as a function of a single scalar variable. For simplicity we use $P_{1,j}$ as the variable. Since $s_{i,j} = \phi_i / P_{1,j}$, we can also express channel thresholds belonging to the same boundary as a function of $s_{1,j}$ given as $s_{i,j} = (\phi_i/\phi_1)\, s_{1,j}$. When all channels from the CHs to the Fusion centre are independent and identically

distributed, using SLA, EPPR and EPPC, problem (2.35) becomes

$$\begin{aligned} \min \quad & \bar{F}_N(s_{1,1}) \\ \text{s.t.} \quad & P_j(\bar{F}_N(s_{1,j+1}) - \bar{F}_N(s_{1,j})) = \tfrac{P_{av}}{L}, \quad j = 1, \ldots, L-1 \\ & P_L(1 - \bar{F}_N(s_{1,L}) + \bar{F}_N(s_{1,1})) = \tfrac{P_{av}}{L} \\ & 0 < s_{1,1} < s_{1,2} < \ldots < s_{1,L}. \end{aligned} \tag{2.37}$$

For low values of the long-term average power, we solve problem (2.36) by using the nonlinear optimization toolbox 'fmincon' in MATLAB. and for large values long-term average power, we solve problem (2.37) using simple binary search algorithm through the constraint. The results are then combined and only the best are selected out of the outage performance obtained from these two problems. Note that the constraint on the component-wise ordering of the powers in problem (2.37) is automatically satisfied due to EPPC and EPPR approximations. In problem (2.36), we can preserve the power-ordering constraint by breaking down the problem into a series of inter-related sub-problems where we first solve for $\mathbf{s}_{L-1}$, and then solve for $\mathbf{s}_{L-2}$ and by following the same steps we can eventually solve for $\mathbf{s}_1$. Note that $\mathbf{s}_L$ has all its elements equal to positive infinity. The sub-problems are given as $\min \bar{F}_N(\mathbf{s}_{L-1})$ s.t. $\Lambda_{L-1}(1-\bar{F}_N(\mathbf{s}_{L-1}^{(N)})) = \frac{NP_{av}}{L-1}$ and $\min \bar{F}_N(\mathbf{s}_j)$ s.t. $\Lambda_j(\bar{F}_N(\mathbf{s}_{j+1}^{(N)})-\bar{F}_N(\mathbf{s}_j^{(N)})) = \frac{NP_{av}}{L-1}$, $j = 1, \ldots, L-2$. One can easily show that solving this series of sub-problems is the same as solving problem (2.36) by verifying the KKT conditions. At each sub-problem once $\mathbf{s}_{j+1}$ is obtained, we can solve for $\mathbf{s}_j$ by making sure that $\mathbf{s}_j \prec \mathbf{s}_{j+1}$, $j = 1, \ldots, L-2$.

Power allocation for quantized CSI using a simultaneous perturbation stochastic approximation (SPSA) algorithm

The vector channel quantization problem can be formulated as the classical vector quantization problem with a modified distortion measure, and the solution can be found by using the iterative Lloyd's algorithm incorporating SPSA [44]. Since results obtained using this method do not use any approximations, they can provide benchmarks for performance comparison. Lloyd's algorithm with SPSA can find a

locally optimal power codebook that minimizes the outage probability subject to a long-term average power constraint. The Lloyd iteration for codebook improvement involves two steps. In the first step, given the power codebook $\mathcal{P}^{(N,L)}$ one finds the optimal partition for the quantization cells using the nearest neighbor condition by solving the following optimization problem

$$\arg\min_{\mathcal{P}_j^{(N)}} \Lambda_j \quad \text{s.t.} \quad D\left(\mathbf{h}, \mathcal{P}_j^{(N)}\right) \leq D_{max} \tag{2.38}$$

Problem (2.38) can be solved numerically using Monte Carlo simulation for a given $\mathcal{P}^{(N,L)}$. Its solution contains a set of L regions or cells $\mathcal{R}_j^{(N)}$, $j = 1, \ldots, L$ in the vector channel space as well as the outage region $\mathcal{R}_{out}^{(N)} \subseteq \mathcal{R}_L^{(N)}$, where none of the power vectors in the power codebook can achieve the distortion constraint.

In the second step we find the improved power codebook. This involves solving the optimization problem

$$\begin{aligned} \min \quad & E\left[1(D\left(\mathbf{h}, \mathcal{P}_1^{(N)}\right) > D_{max}) | \mathbf{h} \in \mathcal{R}_{out}^{(N)}\right] \times \Pr(\mathbf{h} \in \mathcal{R}_{out}^{(N)}) \\ \text{s.t.} \quad & \sum_{j=1}^{L} \left(\Lambda_j \Pr(\mathbf{h} \in \mathcal{R}_j^{(N)})\right) \leq N P_{av} \end{aligned} \tag{2.39}$$

where $1(\cdot)$ is the indicator function. Because we do not have an explicit outage probability expression, we resort to using SPSA, a type of stochastic optimization algorithm, to numerically search for the new power codebook [45]. SPSA randomly chooses the search direction and iterates toward a locally optimal solution. Denote $\tilde{\mathcal{P}} = [\mathcal{P}_1^{(N)^T}, \cdots, \mathcal{P}_L^{(N)^T}]^T$ as the NL by 1 column vector. Define a loss function $J(\tilde{\mathcal{P}}) = \Pr(\mathbf{h} \in \mathcal{R}_{out}^{(N)}) + \bar{\lambda} \sum_{j=1}^{L} \left(\Lambda_j \Pr(\mathbf{h} \in \mathcal{R}_j^{(N)})\right)$ where $\bar{\lambda}$ is the Lagrangian multiplier. Since the loss function can be viewed as the objective function of an unconstrained optimization problem, we will have to obtain P_{av} numerically as a function of $\bar{\lambda}$. Once the new power codebook is found, we repeat step 1 and step 2 until the stopping criterion is met. The 2-sided SPSA algorithm used in this chapter can be summarized by the following steps [38]:

1. *Initialization and Coefficient Selection*: Set counter index $k = 0$. Use a random initial power codebook $\tilde{\mathcal{P}}_0$ and set non-negative coefficients a, c, A, α and γ

in the SPSA gain sequences as $a_k = a/(A+k+1)^{\alpha}$ and $c_k = c/(k+1)^{\gamma}$. For additional guidelines on choosing these coefficients, see [38].

2. *Generation of Simultaneous Perturbation*: Generate a NM-dimensional random perturbation column vector Δ_k. Each component of Δ_k are i.i.d Bernoulli ± 1 distributed with probability of 0.5 for each ± 1 outcome.

3. *Loss Function Evaluations*: Obtain two measurements of the loss function based on the simultaneous perturbations around the current power codebook $\tilde{\mathcal{P}}_k$: $J(\tilde{\mathcal{P}}_k + c_k\Delta_k)$ and $J(\tilde{\mathcal{P}}_k - c_k\Delta_k)$ with c_k and Δ_k are defined in Steps 1 and 2.

4. *Gradient Approximation*: Generate the simultaneous perturbation approximation to the unknown gradient given as $\hat{g}_k(\tilde{\mathcal{P}}_k) = \frac{J(\tilde{\mathcal{P}}_k+c_k\Delta_k)-L(\tilde{\mathcal{P}}_k-c_k\Delta_k)}{2c_k}\left[\Delta_{k,1}^{-1}, \Delta_{k,2}^{-1}, \ldots, \Delta_{k,NL}^{-1}\right]^T$ where $\Delta_{k,i}$ is the ith component of the Δ_k vector.

5. *Updating power codebook*: Use the standard stochastic approximation form $\tilde{\mathcal{P}}_{k+1} = \tilde{\mathcal{P}}_k - a_k\hat{g}_k(\tilde{\mathcal{P}}_k)$.

6. *Iteration or Termination*: Return to Step 2 with $k+1$ replacing k. Terminate the algorithm if there is little change in several successive iterations or the maximum allowable number of iterations has been reached.

Remark 1. SPSA is computationally intensive and requires tuning $\bar{\lambda}$ and all the coefficients whenever network parameter changes, such as any changes in the average power constraint or the number of feedback bits. Convergence can be slow and may settle to different local minima depending on the initial points chosen. Hence in the next section, we will only provide limited SPSA results (up to 4 bits of feedback) as a performance benchmark for our various approximate distortion outage minimization algorithms.

2.4.3 Asymptotic behavior of outage probability and diversity gain in quantized feedback

In this section we briefly present some results on the asymptotic behaviour of the distortion outage probability as the available long-term average power P_{av} goes to infinity. We also provide an approximation for the diversity gain (see definition below) which essentially indicates how fast the outage probability decays with increasing P_{av}. The asymptotic behavior of outage probability as $P_{av} \to \infty$ is given in the following Lemma.

Lemma 2.4.2. *Suppose the fading channels between the clusterheads and the fusion centre undergo independent Nakagami-m fading with the i-th clusterhead having a fading parameter of m_i. As $P_{av} \to \infty$, the asymptotic distortion outage probability achieved by the SLA based power allocation algorithm with quantized channel feedback of $R = \log_2 L$ bits is given by*

$$\lim_{P_{av}\to\infty} P_{outage} \approx \left(\frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} \right)^{Q^{L-1}+\cdots+Q+1} \times \left(\frac{LQ}{NP_{av}} \right)^{Q^L+\cdots+Q^2+Q} \tag{2.40}$$

where $Q = \sum_{i=1}^{N} m_i$. Note that $P_{outage} \approx \tilde{F}_N(s_{1,1})$ given by (6.22) in the Appendix.

The diversity gain d is defined as

$$d \triangleq -\lim_{P_{av}\to\infty} \frac{\log P_{outage}}{\log P_{av}} \tag{2.41}$$

Theorem 2.4.1. *Under the same conditions as in Lemma 2.4.2, the diversity gain achieved by the SLA based power allocation algorithm with quantized channel feedback of $R = \log_2 L$ bits is given by $d \approx Q^L + \cdots + Q^2 + Q$, where $Q = \sum_{i=1}^{N} m_i$.*

Remark 2. Note that there are a number of approximations (all of them analytically justified) that are used to derive the above results as can be seen in their proofs in the Appendix. We would like to remark here that it is because of this reason we express the asymptotic expressions as approximate relationships. Whether or not these limiting values hold exactly with equality is left open for future research.

2.4.4 Simulation results

Our simulation results are based on the topology given in Fig. 2.13. The topology for $N = 1$ (one cluster) is obtained by discarding all the clusters except the one on the top left. For $N = 2$ we keep the top left and the bottom right clusters. For $N = 6$ the topology is given as it is in Fig. 2.13. The sensors in each cluster are placed in four equally-spaced concentric circles and the number of sensors in each circle are 6, 12, 18 and 24 from the smallest to the biggest circle respectively. All clusters have a radius of 40m. All sensors transmit with a power of $q_n/M_n = 1m$W. The cluster heads are located at the center of each cluster for simplicity. CHs are 100m apart from the next closest CH (for the 6-cluster network). The FC is located 500m away from the source. The channel noise variances are set to $(\sigma_{C1}^n)^2 = 10^{-12}$ Watts and $(\sigma_{C2}^n)^2 = 10^{-10}$ Watts $\forall n$. The source variance is set to $\sigma_\theta^2 = 1$ Watt. The maximum distortion threshold D_{max} is set to 0.0043 (10% of the minimum achieveable distortion of the 6-cluster network). Recall that there are no expressions of the outage probability for $N \geq 2$ in closed form, hence we obtain the outage probability via Monte Carlo simulation over 1,000,000 channel realizations using the locally optimum power allocation (for $N = 1$ and the SPSA algorithm) and the strictly sub-optimal power allocation obtained via SLA and the various other approximations such as EPPC and EPPR etc. For very low average power values, the outage performance is obtained using the ZPOR algorithm.

We now present the simulation results for $N = 1, 2, 6$. In this section we give simulation results based on three different Nakagami-m fading parameters, i.e., $m = 0.5$ (severe fading), $m = 1$ (Rayleigh fading) and $m = 2$ (less severe fading). We assume that fading channels between CHs and FC have identical fading parameters ($m_i = m_k \; \forall i, k$) since the transmission distances between CHs and FC are relatively large compared with the inter-cluster distances.

The outage performance of the single cluster limited-feedback problem using EPPR approximation with Nakagami fading parameter $m = 0.5$ is shown in Fig. 2.14. Although in the single cluster network we are only quantizing a scalar channel space, its performance studies allow us obtain some fundamental but important

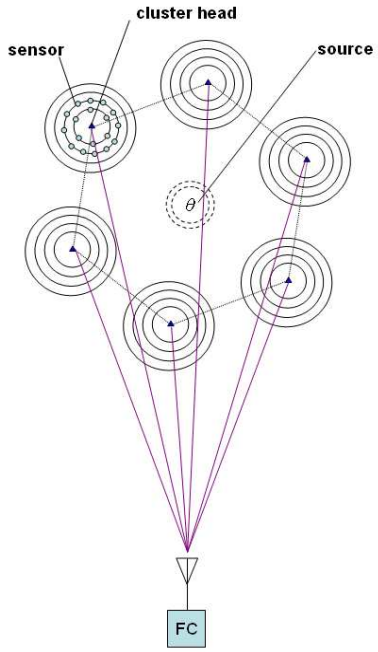

Figure 2.13: Wireless sensor network topology.

insights into the results for quantizing the multi-dimensional vector channel space. The outage performance using equal power allocation (EPA), allocating all CHs with the same powers, and optimal power allocation scheme for full CSI using (2.20) are also shown in the figure to provide performance benchmarks. Fig. 2.14 shows a progression of performance improvement from EPA which has no knowledge of CSIT, to partial CSIT with increase in feedback resolution from 1 bit to 6 bits, to full CSI (complete knowledge of CSIT). At $P_{outage} = 0.1$, a 1-bit feedback can achieve roughly half the power gain (in dB) than that of EPA relative to full-CSI. With $R = 6$, the outage performance is already very close to full CSI.

Fig. 2.15 gives some indications of how good the approximation methods (SLA and SLA+EPPR) are for $N = 2$, $R = 1$ and $m = 0.5, 1, 2$. The benchmark here is the optimal outage performance obtained using an exhaustive search (ES) method. The exhaustive search is used due to difficulties in obtaining the closed-form outage expressions for $N > 1$. ES is carried out over 100,000 search points in $\Re_+^2$. Fig. 2.15 shows both SLA and SLA+EPPR are good approximations at least under this topology setting as both give results that are closely matched to the optimal outage performance. In the remaining simulation results we will only be using the

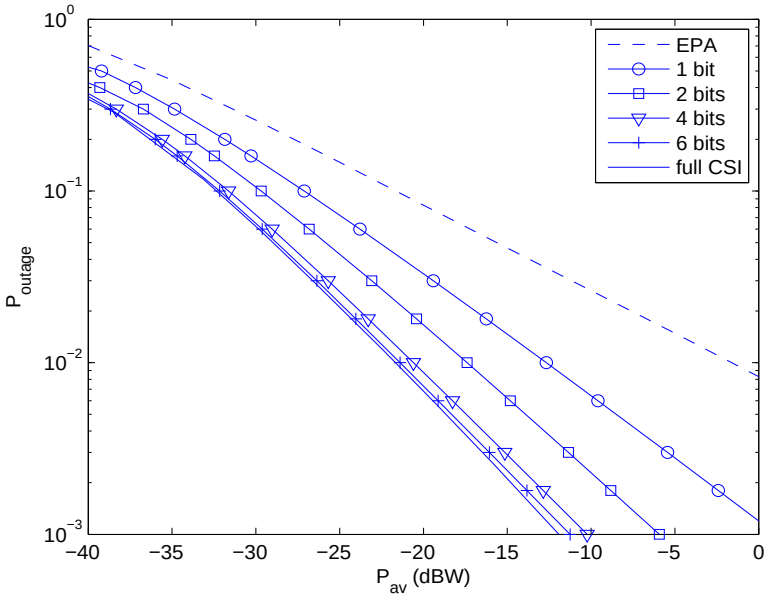

Figure 2.14: Outage performance of a single-cluster network employing EPA,1,2,4 and 6 feedback bits and optimal full-CSI power allocation for $m = 0.5$.

combination of SLA and EPPR.

Fig. 2.16 and Fig. 2.17 show the outage performance obtained by SPSA algorithm and EPA, SLA+EPPR and full CSI for $N = 2$ and $R = 1, 2, 4$ for $m = 0.5$ and 2 respectively. Comparing these two figures we find that larger average power is required in Fig. 2.16 to achieve the same outage probability due to more severe fading. We can also observe that as the number of CH increases from one to two, less average power is required to achieve the same outage probability due to diversity gain. For example, for $m = 0.5$, $R = 4$ and $N = 2$, the long-term average power required to achieve an outage performance of 0.1 is -39dBW, 7.4dB less than $N = 1$ with the same settings. Note also that the power gain gap between the 4-bit feedback and the full CSI has widened. This gap will become more prominent in the case $N = 6$. Also note that SPSA gives very similar results as to SLA+EPPR. The coefficients used in SPSA algorithm are roughly set to $c = 10^{-5}$, $A = 80$, $\alpha = 0.602$, $\gamma = 0.101$ and $a = 10^{-6} \cdot (A+1)^{\alpha}/(\text{mean magnitude of } \hat{g}_0)$. In step 2 of the Lloyd's algorithm outlined in Section 2.4.2 the probabilities are calculated by Monte Carlo simulation over 100,000 vector channel realizations.

The outage performance for $N = 6$, $R = 1, 2, 4$ obtained by using EPA, SLA,

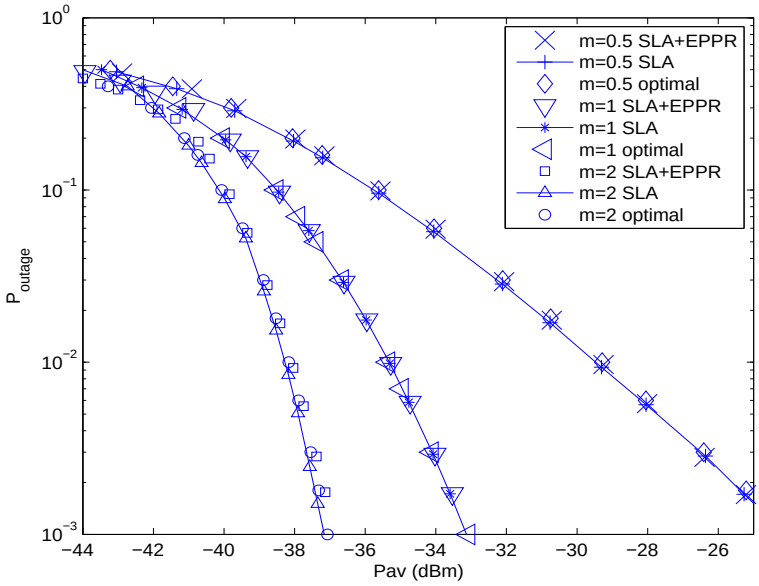

Figure 2.15: Comparison of the outage probability of a 2-cluster network 1-bit feedback. Figure shows optimal, SLA and SLA+EPPR for different Nakagami-m fading parameters.

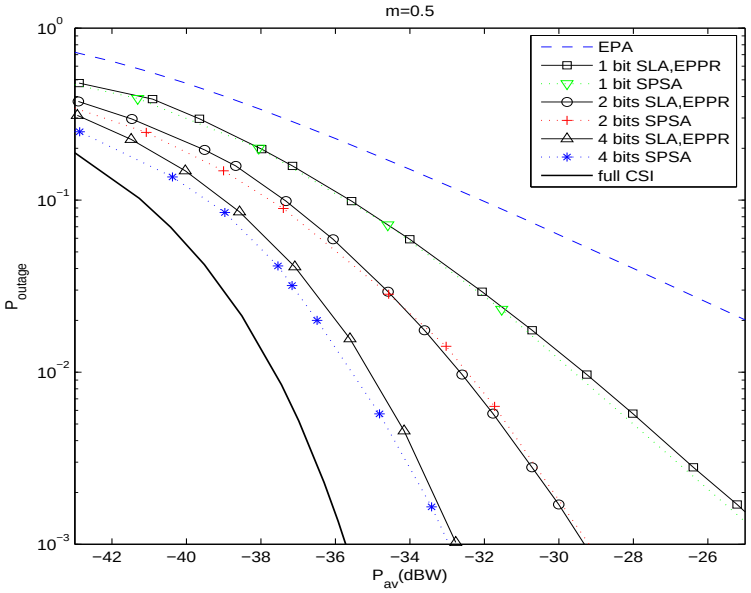

Figure 2.16: Outage performance of 1, 2 and 4-bit feedback, full CSI and EPA of the 2-cluster network for $m = 0.5$.

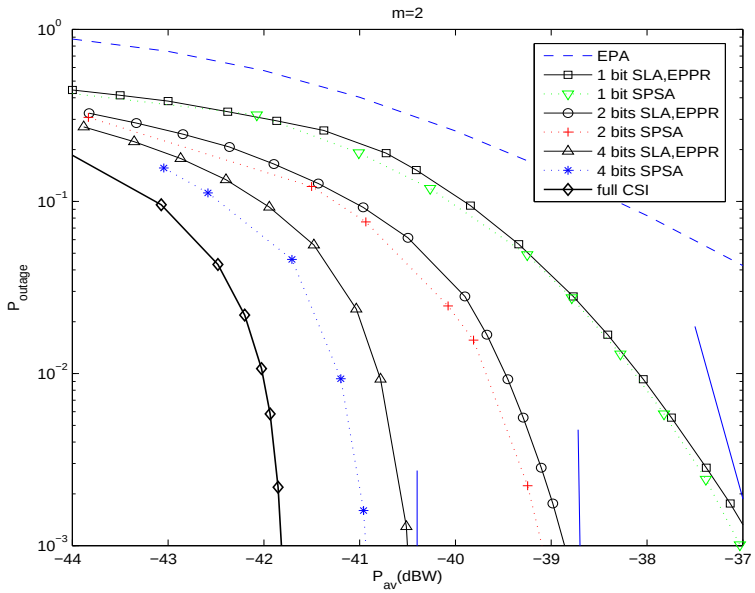

Figure 2.17: Outage performance of 1, 2 and 4-bit feedback, full CSI and EPA of the 2-cluster network for $m = 2$.

SPSA and full CSI for $m = 0.5$ and 1 are shown in Fig. 2.18 and Fig. 2.19 respectively. The parameters used in SPSA here are the same as for $N = 2$. Observe again the effect of diversity gain with the increased number of clusters. The gap between the 4-bit feedback and the full-CSI has widened. This may be due to the fact that the feedback resolution *per CH* decreases as N increases with a fixed R. Simulation results show that at $P_{outage} = 0.1$, having a 4-bit feedback can achieve half the power gain (in dB) than that of EPA relative to full CSI.

The diversity gains are also shown in Fig. 2.17 and Fig. 2.18 as solid straight lines just above the outage probability curves. From the definition of the diversity gain, we can see that it is simply given by the gradient of the outage probability as $P_{av} \to \infty$. Note that the straight lines are inserted in these figures to provide a visual description of the diversity gains by showing the *gradients*; they *do not* represent the actual outage performance. These straight lines indicate the constant slopes at which the outage curves should decrease as P_{av} gets very large.

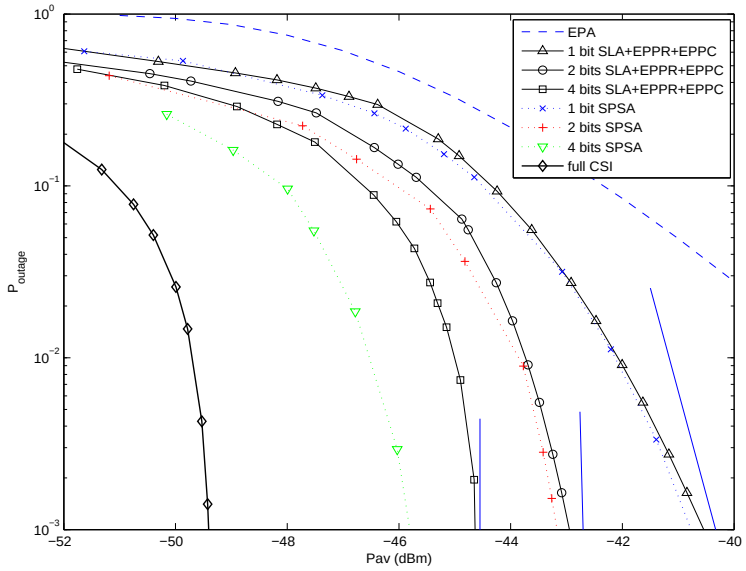

Figure 2.18: Outage performance of 1, 2 and 4-bit feedback, full CSI and EPA of the 6-cluster network for $m = 0.5$.

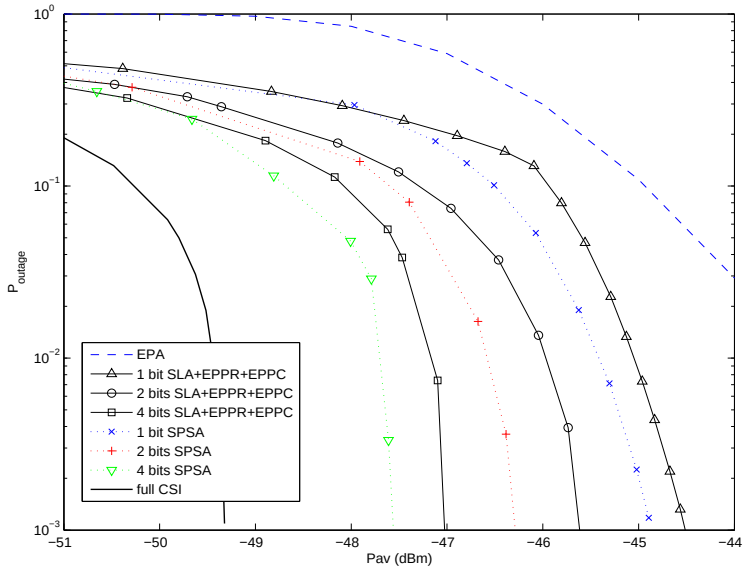

Figure 2.19: Outage performance of 1, 2 and 4-bit feedback, full CSI and EPA of the 6-cluster network for $m = 2$.

2.5 Conclusions

In this chapter we propose outage-minimizing power allocation algorithms for a clustered wireless sensor network using orthogonal multi-access protocol based on (1) having complete knowledge of the instantaneous CSI and (2) having partial CSIT. The sensors within a cluster observe a single Gaussian source and transmit their signals (corrupted by measurement noise) to the pre-elected clusterhead using analog-forwarding via a coherent multi-access channel (assuming distributed beamforming). The clusterheads then transmit the combined signals using amplify-forwarding via orthogonal multi-access channels with fading to the FC which computes an estimate of the source from all the signals it receives. The distortion (which is a function of the random channels between the CHs and the FC) is required to be less than a certain maximum threshold and can violate this condition only with a certain small probability, called the distortion outage probability.

Based on assumption (1), we have derived the optimal power allocation that minimizes the distortion outage probability subject to a long-term average power constraint for any arbitrary fading distributions. The optimal power allocation states that if the channel condition is above some threshold then the CHs should transmit with the solution of minimizing total power subject to a distortion constraint, or else none should transmit to save power. Simulation results show that when Rayleigh fading is considered, this power allocation scheme offers significant power gain over equal power allocation (where all cluster heads transmit with equal power) and some power allocation schemes based on the channel statistics.

Assumption (2) considers a more practical situation where the transmitters (CHs) have limited knowledge of the channel state information. We assume that there is a pre-computed power codebook which is available to both the CHs and the FC. Based on the channel state information, the FC finds the index of the corresponding power codebook and broadcast this index (in terms of fixed number of bits) to all the CHs which then transmit with the power specified by the power codebook. We investigate the power codebook that minimizes the outage probability subject to a long-term average power constraint in Nakagami-m fading channels. Essentially this involves

finding the quantization regions in the multi-dimensional channel state space and the corresponding CH transmission powers in the regions that would minimize the outage probability. We propose a number of low-complexity outage minimization algorithms with various levels of useful approximations. An extensive set of numerical results are presented to demonstrate the performance of these algorithms for different fading conditions (including Rayleigh fading) in Nakagami-m fading. An approximate expression of the diversity gain is also obtained in terms of the number of feedback bits and the Nakagami-m fading parameters for large power. Simulation results show that the diversity gain matched well against Monte-Carlo simulation results.

Future work includes finding better approximations or bounds to solve for the optimal power allocation that further improves the outage performance of the limited feedback case. Extending the work to other fading distributions should also be investigated.

Chapter 3

Power allocation in coherent MAC with full CSI and diversity order analysis

Wireless sensor networks have recently attracted research interests and practical implementations in many areas of human life due to the numerous applications WSNs can achieve such as in environmental monitoring, tracking in defense technology, monitoring chemical levels in factories, and health monitoring, just to name a few. WSNs normally consist of a large number of sensor nodes dispersed over some area to take measurements. The sensor nodes are battery operated devices that have sensing, computation and communication capabilities [5]. The sensors may be configured into various ad-hoc network structures depending on the protocol and the application being considered [7]. Examples of these such as forming clusters and electing cluster heads [18], cooperative transmission and cooperative diversity (relay nodes used to forward signals) [13–16,19] and multiple sensor transmission to achieve distributed beam-forming as in MIMO systems [5] show the flexibility of the WSNs and how various wireless communication technologies can be applied in WSNs.

One important issue in WSNs is the utilization of battery energy, since sensors rely on batteries to stay alive, and replacing batteries is considered expensive. Many works in the literature have considered energy-efficient protocols [46–50], power allocation schemes and cross-layer optimization [5, 11, 19] to optimize the use of energy in WSNs under various different network assumptions and protocols. In distributed estimation sensors independently collect data of some physical phenomenon and transmit their measurements to a central processing unit (a.k.a. the fusion center) where it tries to reconstruct the physical quantity from the sensor measurements. Recently [20] showed that in a Gaussian sensor network it is asymptotically optimal to transmit using uncoded analog forwarding of measurements by multiple sensors

as opposed to separate source channel coding. Later in [51] it was shown that in a Gaussian sensor network it is exactly optimal to transmit using uncoded analog forwarding of measurements by multiple sensors. Many works have since studied the power-allocation problems in multi-sensor estimation under the framework of analog-forwarding transmission.

In [21] the authors obtained the optimal power allocation of an inhomogeneous Gaussian wireless sensor network using analog amplify-and-forward through coherent MAC subject to a distortion constraint (a performance metric given by the variance of the reconstructed source). In the case of amplify-and-forward through orthogonal MAC, [22] solved the problem of minimizing power under distortion constraint and minimizing distortion under power constraint. The study of power allocation in distributed estimation for a vector source is given in [23] for coherent MAC and [52] for orthogonal MAC, which also studied power allocation with correlation in sensor data. Power allocation considering correlated sensor noise is studied in [24]. When fading channels are considered, distortion becomes a random variable as a function of the channel gains and it is not always possible to satisfy the distortion constraint. In such cases an *estimation outage* or *distortion outage* occurs [22]. This leads to the notion of *distortion outage probability*, which is defined as the probability that the distortion exceeds a given threshold D_{max}. The authors in [53] obtained the optimal power allocation that minimizes the distortion outage probability subject to a long-term average power constraint in a clustered WSN using amplify-and-forward orthogonal multi-access protocol.

The estimation diversity achieved by wireless sensor networks was first studied in [22] for equal power allocation in orthogonal multi-access channels with Rayleigh fading. They showed that such a network can achieve an estimation diversity on the order of the number of sensors in the network. In [54] it is shown that the diversity gain is unchanged in the presence of channel estimation error when compared against the perfect channel case. The study of outage scaling laws and diversity for distributed estimation over orthogonal multi-access channels is given in [55] for a large class of fading distributions. With a fixed power per sensor, the authors in [55] showed that the outage probability decays faster than exponentially in the number

of sensors and slower than $\exp(-K \log K)$, where K is the number of sensors.

In the previous chapter we studied outage-minimizing power allocation schemes for orthogonal MAC. In this chapter we will look at a WSN where multiple sensors take noisy measurements of a single i.i.d. Gaussian source and transmit, using amplify-and-forward, their noisy measurements to the fusion center through Rayleigh-faded channels with channel noise modeled by AWGN. We assume that the sensors transmit coherently to the FC so that the signals add up in phase at the FC [20]. Under this setting we consider three power allocation schemes - equal power allocation, short-term optimal power allocation (minimizing distortion) and long-term optimal power allocation (minimizing distortion outage probability) - and give theoretical analysis on the diversity order of distortion outage using these power allocation schemes. We show that the diversity order achieved by the equal power allocation and the short-term power allocation is $N \log N$, where N is the number of sensors. In the long-term optimal power allocation we show that we can drive the outage probability to zero using finite total power for $N > 1$. Using a lower bound on the total instantaneous power, we obtain an approximation for the minimum number of sensors in which the outage probability is driven to zero in the long-term optimal power allocation, for a given power constraint.

This chapter is organized as follows. In Section 3.1 we give the network model. We define and state the three different power allocations in Section 3.2, based on which we perform theoretical analysis to find their diversity orders of distortion outage in Section 3.3. Simulation results are given in Section 3.4, followed by concluding remarks in Section 3.5.

In this chapter, symbols in bold indicate that they are column vectors, e.g., $\mathbf{x} = [x_1, \ldots, x_N]^T$, where ^{T} denotes vector transposition. The arithmetic mean of a vector $\mathbf{x}$ of length N is denoted by $\langle \mathbf{x} \rangle \triangleq \sum_{i=1}^{N} x_i/N$. Given a random variable X, its p.d.f. and c.d.f. are denoted as $f_X(x)$ and $F_X(x)$ respectively, while $E[X]$ denotes its expectation.

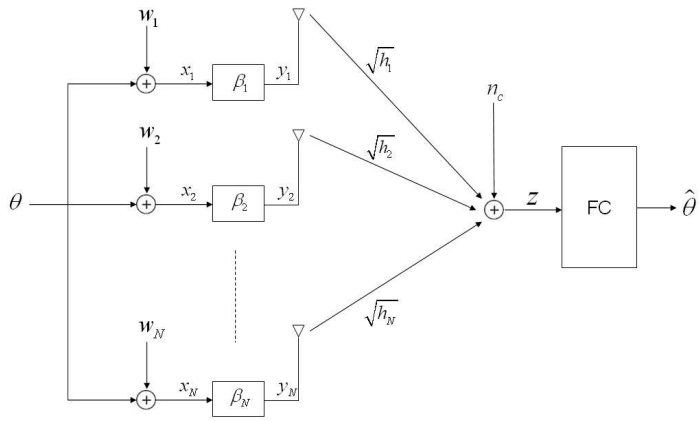

Figure 3.1: Schematic diagram of the wireless sensor network using coherent MAC scheme.

3.1 Network model

A schematic diagram of the wireless sensor network model is shown in Fig. 3.1. We assume that there are N sensors in the network and the sensors observe a single point Gaussian source, denoted by $\theta[k]$, which has zero mean and variance σ_θ^2, and is i.i.d. in time (k denotes the discrete time index). The measurements of the ith sensor at time k are given as

$$x_i[k] = \theta[k] + w_i[k]$$

where w_i is Gaussian with zero mean and variance σ_i^2 and denotes the sensor measurement noise. The sensors amplify and forward their signals to the fusion center via a coherent MAC channel [20] with a gain of $\beta_i[k]$. The transmitted signal is given as

$$y_i[k] = \beta_i[k]x_i[k].$$

We assume that the instantaneous channel gains, denoted as $\sqrt{h_i[k]}$, are time-varying random quantities that are i.i.d. over time (as in the block fading model). The channel noise is i.i.d. AWGN denoted as $n_c[k]$, with zero mean and variance σ_c^2. We assume that full CSI (channel state information including gain and phase)

is available at both the transmitters and the receiver. This implies that the FC is aware of all the values of $h_i[k]$ and the corresponding phase information while the i-th sensor has information of the gain and phase of its own channel to the FC, $\forall i, k$. Note that CSI at the receiver can be easily obtained by the use of pilot tone training from the transmitters, while CSI at the transmitter requires the FC to adopt some feedback mechanism to send the CSI back to the transmitters. We assume that this feedback mechanism is error-free, delay-less and has infinite bandwidth. Since the sensor transmitters are assumed have their channel phase information, they can individually cancel this phase at the transmitter and hence the signal received by the FC is given by[1]

$$z[k] = \sum_{i=1}^{N} \sqrt{h_i[k]}\beta_i[k]\theta[k] + \sum_{i=1}^{N} \sqrt{h_i[k]}\beta_i[k]w_i[k] + n_c[k]. \tag{3.1}$$

Remark 3. Note that in this chapter we are not claiming that such perfect synchronization at the sensor transmitters or in other words, distributed transmit beamforming is a realistic assumption. However our goal in this chapter is to derive the diversity order of distortion outage probability under this idealistic assumption. An analysis involving the case where the signals add up noncoherently at the FC will be interesting and is left for future work.

We define the transmission power of the ith sensor as $P_i[k] \triangleq E\left[y_i^2[k]\right]$, and obtain

$$P_i[k] = C_i\beta_i^2[k],$$

where $C_i = \sigma_\theta^2 + \sigma_i^2$.

It is well known that the optimal estimator for θ is the linear MMSE estimator [56], given as $\hat{\theta} = \frac{E[\theta z]}{E[z^2]}z$. The mean squared error or *distortion* D_k of this estimator,

[1]The coherent sum (3.1) requires distributed transmit beamforming [34] that may be difficult to achieve for large sensor networks. This model however is commonly studied, e.g. in works such as [20, 23, 51].

is given as

$$D_k = \left(\frac{1}{\sigma_\theta^2} + \left(\sum_{i=1}^{N} \sqrt{\frac{h_i[k]P_i[k]}{C_i}} \right)^2 \left(\sum_{i=1}^{N} \frac{h_i[k]P_i[k]\sigma_i^2}{C_i} + \sigma_c^2 \right)^{-1} \right)^{-1}. \tag{3.2}$$

Note that (3.2) gives the expression of the *instantaneous* distortion, i.e., it is a function of the channel realizations h_i, $\forall i, k$. Due to the randomness of the fading channels, the instantaneous distortion at the FC changes randomly over time. Such estimation networks usually impose a distortion threshold at the FC to guarantee acceptable estimation, and if the instantaneous distortion D_k exceeds the distortion threshold $D_{\max}$, a *distortion outage* event occurs. We define the *distortion outage probability*, or simply *outage probability*, as the probability that the distortion exceeds the maximum distortion threshold, expressed as $P_{outage} \triangleq \Pr(D_k > D_{max})$.

We would like to minimize the distortion outage probability by the use of *power control* or *power allocation*, by adapting the transmission power of the sensors $P_i[k]$. Under full CSI, $P_i[k](\mathbf{h}[k])$ will be assumed to be a function of the channel gains.

Remark: Due to the i.i.d. (in time) nature of the network model, we will drop the time index k from the rest of the chapter.

3.2 Full-CSI power control schemes

In the following subsections we introduce three different power control schemes for our proposed wireless sensor network model. We will give results on the diversity order of distortion outage achieved by these three schemes in Section 3.3.

Remark: In this chapter we assume that the power allocations are limited by a total power $\mathcal{P}_{tot}$ that is fixed as the number of sensors N varies, similar to the "total power constraint" of e.g. [57]. Analysis can also be carried out for the case where the total power $\mathcal{P}_{tot}$ scales linearly with the number of sensors N, but are omitted to avoid repetition.

3.2.1 Equal power allocation

A very simple power allocation scheme is to have all the sensors transmit with the same power. Given a fixed total power constraint $\mathcal{P}_{tot}$, the individual sensor power is then given as $P_i = \mathcal{P}_{tot}/N$, $\forall i$.

3.2.2 Short-term optimal power allocation

Since the transmitters have CSI, we can formulate a power control scheme that minimizes the distortion while satisfying a total power constraint in every transmission. We will call this power allocation the short-term optimal power allocation (ST-OPA). ST-OPA can be obtained by solving the following optimization problem

$$\begin{aligned} \min \quad & D(\mathbf{P}(\mathbf{h}), \mathbf{h}) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{N} P_i(\mathbf{h}) \leq \mathcal{P}_{tot},\ P_i(\mathbf{h}) \geq 0 \quad \forall i. \end{aligned} \tag{3.3}$$

Problem (3.3) has been solved in [23]. The short-term optimal power allocation of the ith sensor is given by

$$P_i^*(\mathbf{h}) = \mathcal{P}_{tot} c_i(h_i) \left(\sum_{j=1}^{N} c_j(h_j) \right)^{-1} \quad \forall i \tag{3.4}$$

where $c_i(h_i) = C_i h_i / \left(C_i + \mathcal{P}_{tot} h_i \sigma_i^2 / \sigma_c^2\right)^2$. From (3.4) we see that the optimal power of the ith sensor is computed by multiplying $\mathcal{P}_{tot}$ by a ratio that is bounded between zero and one, i.e., we divide up $\mathcal{P}_{tot}$ amongst the sensors by using this ratio. Also note that in coherent MAC the sensors will always transmit with non-zero powers, unlike in the case of orthogonal channels where some sensors may turn off and do not transmit [22].

3.2.3 Long-term optimal power allocation

We now consider imposing a long-term total power constraint to the wireless sensor network, where the total power usage is averaged over time. Since the problem now deals with an extra dimension in time, an appropriate performance measure

is the distortion outage probability introduced in Section 3.1. We are interested in finding the optimal power allocation that minimizes the outage probability subject to a long-term total power constraint. We call this power allocation scheme the long-term optimal power allocation (LT-OPA). The problem is given as

$$\begin{aligned} \min \quad & \Pr\left(D\left(\mathbf{P}(\mathbf{h}),\mathbf{h}\right) > D_{max}\right) \\ \text{s.t.} \quad & E\left[\sum_{i=1}^{N} P_i(\mathbf{h})\right] \leq \mathcal{P}_{tot},\ P_i(\mathbf{h}) \geq 0 \quad \forall i. \end{aligned} \tag{3.5}$$

Problem (3.5) can be solved in a similar way to [27]. First consider the following minimization problem given as

$$\begin{aligned} \min \quad & \langle \mathbf{P}\left(\mathbf{h}\right) \rangle \\ \text{s.t.} \quad & D(\mathbf{P}(\mathbf{h}),\mathbf{h}) \leq D_{max},\ P_i(\mathbf{h}) \geq 0 \quad \forall i. \end{aligned} \tag{3.6}$$

We have the following lemma:

Lemma 3.2.1. *With the knowledge of* $\mathbf{h}$*, the solution of problem (3.6) is given as*

$$P_i^*(\mathbf{h}) = P_{tot}(\mathbf{h}) c_i(h_i) \left(\sum_{j=1}^{N} c_j(h_j)\right)^{-1}, \quad i = 1, \ldots, N \tag{3.7}$$

where $c_i(h_i) = C_i h_i / \left(C_i + P_{tot}(\mathbf{h}) h_i \sigma_i^2/\sigma_c^2\right)^2$ *and* $P_{tot}(\mathbf{h})$ *is the solution of*

$$\gamma_{th} = \sum_{i=1}^{N} \frac{h_i}{\left(\frac{\sigma_c^2 C_i}{P_{tot}(\mathbf{h})} + \sigma_i^2 h_i\right)} \tag{3.8}$$

where $\gamma_{th} = 1/D_{max} - 1/\sigma_\theta^2$.

The proof of this lemma can be found in [23] and is hence omitted. One also has the following Lemma which is necessary to find the optimal solution of problem (3.5):

Lemma 3.2.2. *The long-term optimal power* $\mathbf{P}^*(\mathbf{h}) = [P_1^*(\mathbf{h}), \ldots, P_N^*(\mathbf{h})]^T$ *as given in (3.7), is a continuous function of* $\mathbf{h}$*. Furthermore,* $\langle \mathbf{P}^*(\mathbf{h}) \rangle$ *is a non-increasing function of* h_i *for* $i = 1, \ldots, N$*.*

Proof. See Appendix. □

Before we give the solution to problem (3.5), we will also need the following definitions and notations, similar to those in [27]. We first define the regions $\mathcal{R}_T(t) = \left\{\mathbf{h} : \sum_{i=1}^{N} P_i(\mathbf{h}) < t\right\}$, $\overline{\mathcal{R}}_T(t) = \left\{\mathbf{h} : \sum_{i=1}^{N} P_i(\mathbf{h}) \leq t\right\}$ and $\mathcal{B}_T(t) = \left\{\mathbf{h} : \sum_{i=1}^{N} P_i(\mathbf{h}) = t\right\}$. We then define two power sum quantities as

$$P_T(t) = \int_{\mathcal{R}_T(t)} \sum_{i=1}^{N} P_i(\mathbf{h}) dF(\mathbf{h}) \tag{3.9}$$

$$\overline{P}_T(t) = \int_{\bar{\mathcal{R}}_T(t)} \sum_{i=1}^{N} P_i(\mathbf{h}) dF(\mathbf{h}) \tag{3.10}$$

where $F(\mathbf{h})$ denotes the joint c.d.f. of $\mathbf{h}$. Finally, the power sum threshold t^* and the weight u^* are given as $t^* = \sup\{t : P_T(t) < \mathcal{P}_{tot}\}$ and $u^* = \frac{\mathcal{P}_{tot} - P_T(t^*)}{\overline{P}_T(t^*) - P_T(t^*)}$.

With the above lemma and definitions we can now present the solution to problem (3.5).

Theorem 3.2.1. *The solution of problem (3.5) is given as*

$$\hat{\mathbf{P}}(\mathbf{h}) = \begin{cases} \mathbf{P}^*(\mathbf{h}), & \text{if } \mathbf{h} \in \mathcal{R}_T(t^*) \\ \mathbf{0}, & \text{if } \mathbf{h} \notin \overline{R}_T(t^*) \end{cases} \tag{3.11}$$

while if $\mathbf{h} \in \mathcal{B}_T(t^*)$, $\hat{\mathbf{P}}(\mathbf{h}) = \mathbf{P}^*(\mathbf{h})$ *with probability* u^* *and* $\hat{\mathbf{P}}(\mathbf{h}) = 0$ *with probability* $1 - u^*$, *where* $\mathbf{P}^*(\mathbf{h})$ *is given in (3.7).*

The proof follows using similar techniques as in [27] and is hence excluded.

The long-term optimal power allocation scheme that minimizes the outage probability subject to a long-term total power constraint says that if the vector of channel gains falls inside the region defined by $\mathcal{R}_T(t^*)$, where t^* is a quantity that is associated with $\mathcal{P}_{tot}$, then the sensors should transmit with powers given by (3.7) and achieve a distortion of exactly D_{max}. Otherwise, none should transmit to save power, and this is where outage occurs.

We can also obtain another condition that determines whether the sensors transmit or not (hence the condition for an outage event to occur). Note that in order to compute the optimal powers $P_i^*(\mathbf{h})$, we first need to compute $P_{tot}(\mathbf{h})$. From $P_{tot}(\mathbf{h})$ and the definition of t^*, the outage event only occurs if $P_{tot}(\mathbf{h}) > t^*$. Hence in every

transmission, the fusion center simply computes the quantity $P_{tot}(\mathbf{h})$ and compares it against t^*. If $P_{tot}(\mathbf{h}) > t^*$, then all sensors should be turned off to save power. Otherwise, the sensors should transmit with power given by (3.7). The value of t^* would depend on the value of $\mathcal{P}_{tot}$ and it can be predetermined numerically in off-line mode via Monte-Carlo simulation. A closed-form expression of a lower bound on t is given in Section 3.3.3 which allows one to quickly compute a lower bound of t^* given $\mathcal{P}_{tot}$.

3.3 Diversity orders of estimation outage

We are interested in seeing how the outage probability decays as the number of sensors increases. In this section we will obtain for large N asymptotic closed-form expressions of $\log P_{outage}$, for the different power allocation schemes given in Section 3.2. Such expressions characterize the *diversity order of distortion outage* introduced in [22], who showed that the outage probability decays exponentially with the number of sensors for N i.i.d. orthogonal MAC. For analytical tractability, in the following theoretical analysis, we will only consider a homogeneous wireless sensor network where all the measurement noise and fading distributions are i.i.d. As a consequence, we will denote $\sigma_i^2 = \sigma^2$ and $C_i = C = \sigma_\theta^2 + \sigma^2$, $\forall i$.

Notation: For two functions $f(\cdot)$ and $g(\cdot)$, we will use the standard asymptotic notation (see for example [58]) and say that $f \sim g$ as $t \to t_0$, if $\frac{f(t)}{g(t)} \to 1$ as $t \to t_0$.

3.3.1 Equal power allocation

Substituting $P_i = \mathcal{P}_{tot}/N$ into (3.2), after some algebraic manipulation we obtain

$$\frac{D}{\sigma_\theta^2} = \frac{\frac{\sum_{i=1}^N h_i}{N} + \frac{\sigma_c^2}{\sigma^2 \mathcal{P}_{tot}}}{\frac{\sum_{i=1}^N h_i}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}} + \frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h_i}}{N}\right)^2}. \tag{3.12}$$

Inspecting the RHS of (3.12), we note that $\frac{1}{N}\sum_{i=1}^N h_i$ and $\frac{1}{N}\sum_{i=1}^N \sqrt{h_i}$ converge to $E[h]$ and $E[\sqrt{h}]$ respectively by the strong law of large numbers as N gets large. However we find that $\mathrm{var}\left(\frac{1}{N}\sum_{i=1}^N h_i\right) = \frac{1}{N}\mathrm{var}[h]$ and $\mathrm{var}\left(\frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h_i}}{N}\right)^2\right) \approx$

$\frac{4\sigma_\theta^4 N}{\sigma^4}(E[\sqrt{h}])^2\text{var}[\sqrt{h}]$ (obtained using the Delta method [59]). We see that the variance of $\frac{1}{N}\sum_{i=1}^{N} h_i$ decreases like $1/N$, whereas the approximate variance of $\frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^{N}\sqrt{h_i}}{N}\right)^2$ increases with N. We therefore choose to replace $\frac{1}{N}\sum_{i=1}^{N} h_i$ by its mean $E[h]$, and retain $\frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{h_i}\right)^2$ for large N. This gives us the following result where the distortion converges (for large N) almost surely to a random variable expressed as:

$$D \overset{a.s.}{\to} \sigma_\theta^2 \eta \left(\eta + \frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^{N}\sqrt{h_i}}{N}\right)^2\right)^{-1} \tag{3.13}$$

where $\eta = E[h] + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}$.

The asymptotic distortion outage probability for large N can therefore be found as

$$P_{outage} = \Pr\left(D > D_{max}\right) \tag{3.14}$$

$$\to \Pr\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{h_i} < \sqrt{\frac{\eta\sigma^2\left(\sigma_\theta^2 - D_{max}\right)}{D_{max}\sigma_\theta^2 N}}\right) \quad \text{(substitute (3.13) and re-arrange)} \tag{3.15}$$

$$= \Pr\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{h_i} < \frac{a}{\sqrt{N}}\right) \tag{3.16}$$

where $a = \sqrt{\frac{\eta\sigma^2\left(\sigma_\theta^2 - D_{max}\right)}{D_{max}\sigma_\theta^2}}$.

By inspecting (3.16) we see that the asymptotic outage probability is expressed in terms of the empirical mean of i.i.d. random variables $\sqrt{h_i}$ being less than a threshold that is a function of N. This resembles a more general form of the typical large deviation problem where the threshold is a constant. In Theorem 3.3.1 we will provide a generalized version of Cramer's Theorem which can be applied to (3.16). Before we give the theorem we need the following definitions. The moment-generating function of the random variable X is defined as $M_X(t) \triangleq E\left[e^{tX}\right]$. The cumulant-generating function of the random variable X is defined as $\Lambda_X(t) \triangleq \log M_X(t)$. The rate function of the random variable X is defined as $I_X(c) = \sup_t \{ct - \Lambda_X(t)\}$. We also define

the following notations relating to the rate function as $I_X^+(c) = \sup_{t>0} \{ct - \Lambda_X(t)\}$ and $I_X^-(c) = \sup_{t<0} \{ct - \Lambda_X(t)\}$. Note here that I_X^+ and I_X^- have the same value as I_X; these two notations are introduced only to further restrict the domain of the supremum without affecting the result of I_X. Hence these notations may be used interchangeably depending on whether we have extra knowledge of the domain over which the supremum is achieved.

Theorem 3.3.1. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with mean $\mu_X > 0$, and suppose that their moment generating function $M_X(t) = E\left[e^{tX}\right]$ is finite in some neighborhood of the origin $t = 0$. Let $\tilde{Y}_{n,i}$ be the exponential change of distribution of $Y_i = -X_i + \mu_X$ defined as*

$$dF_{\tilde{Y}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)} dF_Y(y) \tag{3.17}$$

Suppose that $\Pr\left(\frac{1}{n}\sum_{i=1}^n \tilde{Y}_{n,i} > E\left[\tilde{Y}_{n,i}\right]\right)$ is bounded away from zero as $n \to \infty$. Let $a_n = \frac{a}{n^p}$, $p \geq 0$ and $\Pr(X < a_n) > 0$, $\forall n$. Then $I_X(a_n) > 0$ for sufficiently large n, and

$$\log \Pr\left(\frac{1}{n}\sum_{i=1}^n X_i \leq a_n\right) \sim -nI_X\left(a_n\right) \quad \text{as } n \to \infty. \tag{3.18}$$

Proof. See Appendix. □

In order to apply Theorem 3.3.1 to (3.16), we need to verify the assumption that $\Pr\left(\frac{1}{n}\sum_{i=1}^n \tilde{Y}_{n,i} > E\left[\tilde{Y}_{n,i}\right]\right)$ is bounded away from zero as $n \to \infty$. The following lemma verifies this condition in the case of Rayleigh fading.

Lemma 3.3.1. *Let $Y_i = -\sqrt{h_i} + E\left[\sqrt{h_i}\right]$, where $\sqrt{h_i}$ is Rayleigh distributed with parameter κ (i.e. $f_{\sqrt{h}}(x) = \frac{x}{\kappa^2} e^{-x^2/2\kappa^2}$). Denote $\tilde{Y}_{n,i}$ as the exponential change of distribution of Y_i as defined in (3.17). Then*

$$\Pr\left(\frac{1}{n}\sum_{i=1}^n \tilde{Y}_{n,i} > E\left[\tilde{Y}_{n,i}\right]\right) \to 0.5 \quad \text{as } n \to \infty \tag{3.19}$$

Proof. See Appendix. □

Applying Theorem 3.3.1 to (3.16) we have

$$\log P_{outage} \sim -NI^{-}_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) \quad \text{as } N \to \infty \tag{3.20}$$

where

$$I^{-}_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) = \sup_{\theta<0}\left(\frac{a}{\sqrt{N}}\theta - \log M_{\sqrt{h}}(\theta)\right). \tag{3.21}$$

Since $\sqrt{h}$ is Rayleigh distributed with parameter κ, its moment generating function is available in closed form as

$$M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right) = 1 - \sqrt{\pi}xe^{x^2}\operatorname{erfc}(x) \tag{3.22}$$

where we have used a substitution of variables $\theta = -\sqrt{2}x/\kappa$.

We need to find the value of θ that attains the supremum in the rate function $I^{-}_{\sqrt{h}}(a/\sqrt{N})$. This value of θ can be found by using the stationary condition (first derivative) given as

$$\frac{dI^{-}_{\sqrt{h}}\left(a/\sqrt{N}\right)}{d\theta} = 0, \quad \theta < 0 \tag{3.23}$$

$$\Rightarrow \frac{\sqrt{N}}{a} = \psi(\theta) \tag{3.24}$$

where

$$\psi(\theta) = \left(\Lambda'_{\sqrt{h}}(\theta)\right)^{-1} = M_{\sqrt{h}}(\theta)/M'_{\sqrt{h}}(\theta). \tag{3.25}$$

After substituting $\theta = -\sqrt{2}x/\kappa$ in (3.24) and some algebraic manipulation, it is possible to obtain

$$\frac{\sqrt{N}}{2} = \psi\left(-\frac{\sqrt{2}x}{\kappa}\right) \tag{3.26}$$

where

$$\psi\left(-\frac{\sqrt{2}x}{\kappa}\right) = \frac{\sqrt{2}}{\kappa}\frac{xM_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)}{1 - M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right) - 2x^2 M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)}. \tag{3.27}$$

Note that $\psi\left(-\frac{\sqrt{2}x}{\kappa}\right)$ is a continuous non-decreasing function of x since

$$\frac{d\psi\left(-\frac{\sqrt{2}x}{\kappa}\right)}{dx} = \frac{\sqrt{2}}{\kappa}\frac{\Lambda''_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)}{\left(\Lambda'_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)\right)^2} \geq 0 \tag{3.28}$$

where the inequality is due to the cumulant generating function being a convex function and hence its second derivative is non-negative. The continuity of $\psi\left(-\frac{\sqrt{2}x}{\kappa}\right)$ can be seen from (3.25); since $M_{\sqrt{h}}(\theta)$ is a positive continuous strictly-increasing convex function, this implies that $M'_{\sqrt{h}}(\theta) > 0$, and the change of variables from θ to x preserves the continuity of the function.

Hence from (3.26), large N corresponds to the case of large x. We now show that $\psi\left(-\sqrt{2}x/\kappa\right)$ in fact increases linearly in x for large x. We substitute the asymptotic expansion of the complementary error function (for large x) given as $\mathrm{erfc}(x) = \frac{e^{-x^2}}{x\sqrt{\pi}}\sum_{n=0}^{\infty}(-1)^n\frac{(2n)!}{n!(2x)^{2n}}$ into the moment generating function (3.22) and obtain

$$M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right) = \frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots. \tag{3.29}$$

We then substitute (3.29) into (3.26) to obtain the following

$$\frac{\sqrt{N}}{a} = \frac{\sqrt{2}}{\kappa}\frac{x\left(\frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots\right)}{1 - \left(\frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots\right) - 2x^2\left(\frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots\right)} \tag{3.30}$$

$$= \frac{\sqrt{2}}{\kappa}\frac{\frac{1}{2x} - \frac{3}{4x^3} + \frac{15}{8x^5} + \cdots}{\frac{1}{x^2} - \frac{3}{x^4} + \cdots} \sim \frac{\sqrt{2}}{\kappa}\frac{x}{2} \quad \text{for large } x \tag{3.31}$$

Hence for large N,

$$\theta \sim -\frac{2\sqrt{N}}{a}. \tag{3.32}$$

Substituting this asymptotic expression for θ back into the rate function gives

$$I_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) \sim -\frac{a}{\sqrt{N}}\frac{2\sqrt{N}}{a} - \log M_{\sqrt{h}}\left(-\frac{2}{a_N}\right) = -2 - \log M_{\sqrt{h}}\left(-\frac{2}{a_N}\right) \tag{3.33}$$

$$\sim -2 - \log\left(\frac{a^2}{2\kappa^2 N}\right) \tag{3.34}$$

$$= -2 - \log\left(\frac{a^2}{2\kappa^2}\right) + \log N. \tag{3.35}$$

Hence from (3.20) the outage probability for large N satisfies

$$\log P_{outage} \sim - N I_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) \tag{3.36}$$

$$\sim - N\left(-2 - \log\left(\frac{a^2}{2\kappa^2}\right) + \log N\right) \tag{3.37}$$

$$\sim - N \log N \tag{3.38}$$

which shows that the diversity order of distortion outage in i.i.d. coherent MAC with Rayleigh fading using EPA is $N \log N$ for large N. In [22], the authors obtained a diversity order of N for i.i.d. orthogonal MAC with Rayleigh fading using EPA. We thus see that the coherent MAC achieves a higher diversity order over the orthogonal MAC case by a factor of $\log N$ for i.i.d. Rayleigh-faded channels.

Remark: Note that if the total power scales linearly with the number of sensors, then a diversity order of $N \log N$ for orthogonal MAC can also be achieved [55]. In contrast, here we showed that for coherent MAC a diversity order of $N \log N$ can still be achieved when the total power is fixed.

3.3.2 Short-term optimal power allocation

We first give the expression of distortion using ST-OPA. Substituting (3.4) into (3.2) gives

$$D = \left(\frac{1}{\sigma_\theta^2} + \frac{\left(\sum_{i=1}^N \sqrt{h_i P_i^*}\right)^2}{\sigma^2 \sum_{i=1}^N h_i P_i^* + \sigma_C^2 C}\right)^{-1} = \frac{\sigma_\theta^2 \sigma^2}{\sigma^2 + \sigma_\theta^2 \sum_{i=1}^N Z_i} \tag{3.39}$$

where $Z_i = h_i/(h_i + \rho)$ with $\rho = C\sigma_c^2/\mathcal{P}_{tot}\sigma^2$, and the second equality follows after some algebraic manipulation. The distortion outage probability can therefore be written as

$$P_{outage} = \Pr(D > D_{max}) = \Pr\left(\frac{1}{N}\sum_{i=1}^N Z_i < g_N\right) \tag{3.40}$$

where $g_N = g/N$ and $g = \sigma^2(1/D_{max} - 1/\sigma_\theta^2)$.

Denote Z as the random variable distributed according to the common distribution of Z_i. We now apply Theorem 3.3.1 to (3.40). We have the following lemma needed for verifying one of the assumptions in Theorem 3.3.1 (similar to lemma

3.3.1).

Lemma 3.3.2. *Let $Y_i = -Z_i + E[Z_i]$, where $Z_i = h_i/(h_i + \rho)$, with h_i being exponentially distributed. Denote $\tilde{Y}_{n,i}$ as the exponential change of distribution of Y_i as defined in (3.17). Then*

$$\Pr\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{Y}_{n,i} > E\left[\tilde{Y}_{n,i}\right]\right) \to 0.5 \quad as\ n \to \infty \tag{3.41}$$

This lemma can be proved in a similar manner to Lemma 3.3.1 and is excluded to avoid repetition.

Applying Theorem 3.3.1 to (3.40) we have

$$\log P_{outage} \sim -NI_Z^-(g_N) \quad \text{as } N \to \infty \tag{3.42}$$

where $I_Z^-(g_N) = \sup_{\theta<0}(g_N\theta - \log M_Z(\theta))$.

In order to obtain $M_Z(\theta)$, we need the distribution of Z. The common distribution of i.i.d. random variables Z_i can be easily obtained since $Z_i = \left(1 + \frac{\rho}{h_i}\right)^{-1}$, where h_i are i.i.d. exponentially distributed random variables with parameter λ. Note that the domain of Z_i is $[0, 1)$. The c.d.f. and p.d.f. of Z are given by $F_Z(z) = 1 - e^{-\frac{\lambda\rho}{1/z-1}}$ and $f_Z(z) = \lambda\rho\frac{1}{(1-z)^2}e^{-\lambda\rho\frac{z}{1-z}}$ respectively. The mean of Z is given as $\mu_Z = 1 - \lambda\rho e^{\lambda\rho}E_1(\lambda\rho)$, where $E_1(x) = \int_x^\infty \frac{e^{-t}}{t}dt$ is the exponential integral. The moment generating function of Z is given as $M_Z(\theta) = E\left[e^{\theta Z}\right] = \lambda\rho\int_0^1 \frac{1}{(1-z)^2}e^{\theta z - \lambda\rho\frac{z}{1-z}}dz$.

We need to find the value of θ that attains the supremum in the rate function $I_Z^-(g_N)$. This value of θ can be found by using the stationary condition $\frac{dI_Z^-(g_N)}{d\theta} = 0$, $\theta < 0$. Taking the first derivative of the rate function gives

$$g_N - \frac{M_Z'(\theta)}{M_Z(\theta)} = 0 \Rightarrow g_N = \frac{\int_0^1 \frac{z}{(1-z)^2}e^{\theta z - \lambda\rho\frac{z}{1-z}}dz}{\int_0^1 \frac{1}{(1-z)^2}e^{\theta z - \lambda\rho\frac{z}{1-z}}dz} \tag{3.43}$$

$$\Rightarrow g_N = \frac{\int_0^1 zg(z,t)dz}{\int_0^1 g(z,t)dz} \tag{3.44}$$

where $t = -\theta$ and $g(z,t) = \frac{1}{(1-z)^2}e^{-tz-\lambda\rho\frac{z}{1-z}}$.

Note that as N increases, g_N decreases to zero. Also note that $g(z,t) > 0$. Let $\varphi(\theta) = M_Z'(\theta)/M_Z(\theta)$. Replacing θ by $-t$ and taking the derivative of $\varphi(-t)$ w.r.t. t yields $\frac{d\varphi(-t)}{dt} = -\Lambda_Z''(-t) \leq 0$, where the inequality arises due to the cumulant generating function being a convex function. Hence $\varphi(-t)$ is a continuous non-increasing function of t (the continuity of $\varphi(-t)$ is evident by inspecting the RHS of (3.44)). Hence large N corresponds to the case of large t in (3.44). Let $x = 1/(1-z)$. It can be easily shown that (3.44) can be written as

$$g_N = 1 - \frac{\int_1^\infty \frac{1}{x} e^{\frac{t}{x} - cx} dx}{\int_1^\infty e^{\frac{t}{x} - cx} dx} \tag{3.45}$$

where $c = \lambda\rho$.

Lemma 3.3.3.

$$g_N \sim \frac{1}{t} \text{ as } t \to \infty. \tag{3.46}$$

Proof. See Appendix. □

Hence for large N, we have

$$\theta \sim -\frac{1}{g_N}. \tag{3.47}$$

Substituting this asymptotic expression for θ back into $M_Z(\theta)$ gives

$$M_Z(\theta) = \lambda\rho e^{-t+c} \int_1^\infty e^{-tp(x)} q(x) dx \sim \lambda\rho e^{-t+c} \frac{e^{t-c}}{t} \sim \lambda\rho g_N \tag{3.48}$$

Substituting $\theta \sim -\frac{1}{g_N}$ and $M_Z(\theta) \sim \lambda\rho g_N$ back into the rate function gives

$$I_Z(a_N) \sim -g_N \frac{1}{g_N} - \log\left(\frac{\lambda\rho g}{N}\right) \quad \text{for large } N \tag{3.49}$$

$$= -1 - \log(\lambda\rho g) + \log N. \tag{3.50}$$

Hence from (3.42) the outage probability for large N is asymptotically

$$\log P_{outage} \sim -N I_Z(g_N) \tag{3.51}$$

$$\sim -N(-1 - \log(\lambda\rho g) + \log N) \tag{3.52}$$

$$\sim -N \log N. \tag{3.53}$$

Hence the diversity order of distortion outage for i.i.d. coherent MAC with Rayleigh fading using ST-OPA is $N \log N$, which interestingly achieves the same diversity order of distortion outage as EPA.

3.3.3 Long-term optimal power allocation

In this section we first show that it is possible to use LT-OPA in coherent MAC to achieve zero distortion outage with a finite amount of power, if the number of sensors $N > 1$. We will later show that this result implies that for a given power constraint it is possible to achieve zero distortion outage with finite N, i.e., there exists a finite number of sensors that will drive the distortion outage to zero. We will obtain an approximate expression for finding such N.

We first analyze the power required to achieve zero outage. For $N = 1$, the sum power expression in (3.8) can be re-arranged and expressed as $P_{tot}(h) = \frac{K_1}{h}$ where $K_1 = \frac{\gamma_{th}\sigma_c^2 C}{(1-\sigma^2\gamma_{th})}$. The region $\mathcal{R}_T(t)$ can be easily found directly from the definition as $\mathcal{R}_T(t) = \{h : P_{tot}(h) < t\} = \left\{h : h > \frac{K_1}{t}\right\}$. The average power sum, $P_T(t)$, becomes

$$P_T(t) = \int_{\mathcal{R}_T(t)} P_{tot}(h) dF(h) = \int_{\frac{K_1}{t}}^{\infty} \frac{K_1}{h} \lambda e^{-\lambda h} dh = \lambda K_1 \int_{\frac{\lambda K_1}{t}}^{\infty} \frac{e^{-u}}{u} du \tag{3.54}$$

$$= \lambda K_1 E_1\left(\frac{\lambda K_1}{t}\right) \tag{3.55}$$

where $u = \lambda h$ and $E_1(x) = \int_x^{\infty} \frac{e^{-t}}{t} dt$ is the exponential integral. To find the maximum total power that achieves zero-outage, we simply let $t \to \infty$. This is because the region $\mathcal{R}_T(t)$ defines the set of channel realizations where the sensor *does* transmit to meet the distortion constraint. Hence, the outage probability is also given by $P_{outage} = \Pr(h \notin \overline{\mathcal{R}}_T(t))$. When we let $t \to \infty$, we increase $\mathcal{R}_T(t)$ to be the whole channel space, implying that the outage region is reduced to null, and hence outage probability is reduced to zero. However, as $t \to \infty$, $P_T(t) \to \infty$, implying that we need an infinite amount of power to achieve zero outage for $N = 1$.

For $N > 1$ it is difficult to obtain closed form expressions of the maximum power required to achieve zero-distortion. Instead, we show that it is possible to achieve zero-outage with finite power for $N > 1$. Suppose we have a sub-optimal

power allocation scheme as follows. For every transmission, we select the sensor with the best channel gain and use only that sensor to transmit with just enough power to meet the distortion constraint. Denote the power as $\tilde{P}(h_{max})$ where $h_{max} = \max(h_1, \ldots, h_N)$. $\tilde{P}(h)$ can be obtained from the distortion constraint and it is given as $\tilde{P}(h_{max}) = \frac{\gamma_{th}\sigma_c^2 C}{(1-\sigma^2\gamma_{th})h_{max}}$. We can see that power is proportional to the inverse of the channel gain. This power allocation scheme is simply a channel inversion scheme. The c.d.f. and p.d.f. of choosing the maximum channel gain out of a set of i.i.d. exponential-distributed random variables $\{h_1, \ldots, h_N\}$ is given respectively as $F_{h_{max}}(t) = \left(1 - \lambda e^{-\lambda t}\right)^N$ and $f_{h_{\max}}(t) = N\lambda \left(1 - \lambda e^{-\lambda t}\right)^{N-1} e^{-\lambda t}$. The transmission power averaged over all possible values of the channel realization and over time is then given as

$$E\left[\tilde{P}(h_{max})\right] = \int_0^\infty \frac{\gamma_{th}\sigma_c^2 C}{(1-\sigma^2\gamma_{th})h} \cdot N\lambda \left(1 - \lambda e^{-\lambda h}\right)^{N-1} e^{-\lambda h} dh. \tag{3.56}$$

The integral above is well-known to be finite for $N > 1$. Since this suboptimal power allocation scheme can achieve zero-outage with finite power, the optimal power allocation scheme will also achieve zero-outage with finite power.

We now proceed to find an approximation for the maximum number of sensors N_{max} that still has non-zero outage for a given $\mathcal{P}_{tot}$ for LT-OPA. Then N_{max}+1 can be regarded as the minimum number of sensors that achieves zero outage. To do this, we first find a lower bound on the instantaneous power $P_{tot}(\mathbf{h})$. We begin with the equation we need to solve to obtain $P_{tot}(\mathbf{h})$, given as $\sigma^2\gamma_{th} = \sum_{i=1}^N \left(\frac{\sigma_c^2 C}{\sigma^2 P_{tot}(\mathbf{h})h_i} + 1\right)^{-1}$. Let $f(h_i) = \left(\frac{\sigma_c^2 C}{\sigma^2 P_{tot}(\mathbf{h})h_i} + 1\right)^{-1}$. It is straight forward to show that f is concave in h_i $\forall i$. Applying Jensen's inequality we have

$$\sigma^2\gamma_{th} = \frac{\sum_{i=1}^N f(h_i)}{N} \leq f\left(\frac{\sum_{i=1}^N h_i}{N}\right) \Rightarrow \frac{\sigma^2\gamma_{th}}{N} \leq \frac{1}{\frac{\sigma_c^2 C}{\sigma^2 P_{tot}(\mathbf{h})\frac{1}{N}\sum_{i=1}^N h_i} + 1} \tag{3.57}$$

$$\Rightarrow \frac{\sigma^2\gamma_{th}}{N} \frac{\sigma_c^2 C}{\sigma^2 P_{tot}(\mathbf{h})\frac{1}{N}\sum_{i=1}^N h_i} \leq 1 - \frac{\sigma^2\gamma_{th}}{N} \tag{3.58}$$

$$\Rightarrow P_{tot}(\mathbf{h}) \geq \frac{K_N}{\sum_{i=1}^N h_i} \tag{3.59}$$

where $K_N = \gamma_{th}\sigma_c^2 C / \left(1 - \frac{\sigma^2\gamma_{th}}{N}\right)$.

Let $\breve{P}_{tot}(\mathbf{h}) = K_N / \sum_{i=1}^{N} h_i$. Using the lower bound expression $\breve{P}_{tot}(\mathbf{h})$, we obtain the following modified definitions and expressions to the ones given in Section 3.3.3. The definition of $\mathcal{R}_T(\breve{t})$ becomes

$$\breve{\mathcal{R}}_T(\breve{t}) = \left\{\mathbf{h} : \breve{P}_{tot}(\mathbf{h}) < \breve{t}\right\} = \left\{\mathbf{h} : \sum_{i=1}^{N} h_i > \frac{K_N}{\breve{t}}\right\}. \tag{3.60}$$

The definition of $P_T(\breve{t})$ becomes

$$\breve{P}_T(t) = \int_{\mathcal{R}_T(t)} \breve{P}_{tot}(\mathbf{h}) dF(\mathbf{h}) = K_T \int_{\sum_{i=1}^{N} h_i > \frac{K_N}{\breve{t}}} \frac{1}{\sum_{i=1}^{N} h_i} e^{-\lambda \sum_{i=1}^{N} h_i} dh_1 \cdots dh_N. \tag{3.61}$$

Note that h_i is exponentially distributed with mean $1/\lambda$. Let $T = \sum_{i=1}^{N} h_i$. It is well known that T is Gamma distributed with parameters $k = N$,$\theta = \frac{1}{\lambda}$. Hence $\breve{P}_T(t)$ becomes

$$\breve{P}_T(t) = K_N \int_{\sum_{i=1}^{N} h_i > \frac{K_N}{\breve{t}}} \frac{1}{\sum_{i=1}^{N} h_i} e^{-\lambda \sum_{i=1}^{N} h_i} dh_1 \cdots dh_N = K_N \frac{1}{\Gamma(k)\theta^k} \int_{\frac{K_N}{\breve{t}}}^{\infty} T^{k-2} e^{-\frac{T}{\theta}} dT \tag{3.62}$$

$$= \frac{K_N}{\Gamma(N)\lambda^{-n}} \int_{\frac{K_N}{\breve{t}}}^{\infty} T^{N-2} e^{-\lambda T} dT = \frac{K_N \lambda}{N-1} \cdot \frac{\Gamma\left(N-1, \lambda K_N/\breve{t}\right)}{\Gamma(N-1)}. \tag{3.63}$$

The definition of $\breve{t}^*$ becomes $\breve{t}^* = \sup\left\{\breve{t} : \breve{P}_T(\breve{t}) < \mathcal{P}_{tot}\right\}$. We can solve for $\breve{t}^*$ by letting $\breve{P}_T(\breve{t}^*) = \mathcal{P}_{tot}$ and obtain

$$\frac{K_N \lambda}{N-1} \cdot \frac{\Gamma\left(N-1, \lambda K_N/\breve{t}^*\right)}{\Gamma(N-1)} = \mathcal{P}_{tot}. \tag{3.64}$$

The outage event becomes $\breve{P}_{outage} = \left\{\mathbf{h} : \breve{P}_{tot}(\mathbf{h}) > \breve{t}^*\right\} = \left\{\mathbf{h} : \frac{1}{N}\sum_{i=1}^{N} h_i < \frac{K_N}{N\breve{t}^*}\right\}$. If we let $\breve{t}^* \to \infty$ in (3.64) for a given finite N then $K_N/\breve{t}^* \to 0$, $\frac{\Gamma(N-1,\lambda K_N/\breve{t}^*)}{\Gamma(N-1)} \to 1$ and

$$\frac{K_N \lambda}{N-1} = \mathcal{P}_{tot} \tag{3.65}$$

Equation (3.65) allows us to solve for N, and it gives an approximation $\breve{N}_{max}$ to the maximum number of sensors that has non-zero outage probability for a given $\mathcal{P}_{tot}$.

The solution of (3.65) can be found in closed-form and is given as

$$\breve{N}_{max} = \left\lfloor \frac{(1+\sigma^2\gamma_{th})\mathcal{P}_{tot} + \gamma_{th}\sigma_c^2 C\lambda + \sqrt{[(1+\sigma^2\gamma_{th})\mathcal{P}_{tot} + \gamma_{th}\sigma_c^2 C\lambda]^2 - 4\mathcal{P}_{tot}^2\gamma_{th}\sigma^2}}{2\mathcal{P}_{tot}} \right\rfloor \tag{3.66}$$

where $\lfloor x \rfloor$ denotes the floor function of x.

3.4 Simulation results

The following results, if not computed directly from the equations, are obtained via Monte Carlo simulation over 1,000,000 channel realizations. We first present the diversity order of distortion outage for EPA. We simulated the case where $\mathcal{P}_{tot} = 10m$W and plotted the results in Fig. 3.2. The lines plotted in plus signs shown are plots of $\log P_{outage}$ obtained via Monte Carlo simulation, where log is the natural log. The lines plotted in triangles are the exact values of $-NI_{\sqrt{h}}(a/\sqrt{N})$ where the values of $I_{\sqrt{h}}(a/\sqrt{N})$ are obtained by solving (3.21) numerically. The squares are plots of (3.37). The figure shows that as N gets large, the asymptotic expression (3.37) converges to $-NI_{\sqrt{h}}(a/\sqrt{N})$. Note that the asymptotic results $I_{\sqrt{h}}(a/\sqrt{N})$ and (3.37) only give us the slope of the outage probability when plotted on a log scale; these two lines may not necessarily converge to $\log P_{outage}$ but their gradients should coincide for large N, as can be seen in Fig. 3.2.

We now look at ST-OPA. Fig. 3.3 shows the log of the outage probability using ST-OPA as a function of N in circles and $-NI_Z(g_N)$ obtained numerically in squares for $\mathcal{P}_{tot} = 10$mW. It shows that $-NI_Z(g_n)$ gives a similar gradient as $\log P_{outage}$. In Fig. 3.3 we also show the asymptotic expression of $-NI_Z(g_N)$ plotted in plus signs. We see that as N increases, the asymptotic expression gives very similar gradients as $-NI_Z(g_N)$.

With the long-term OPA, we first present the relation between t^* and $\mathcal{P}_{tot}$ for a fixed total power constraint shown in Fig. 3.4. The circles and squares are obtained via Monte Carlo simulation. The solid lines on the graph are obtained by solving (3.64) numerically. We see that the results match closely. Note that $\mathcal{P}_{tot}$ is a

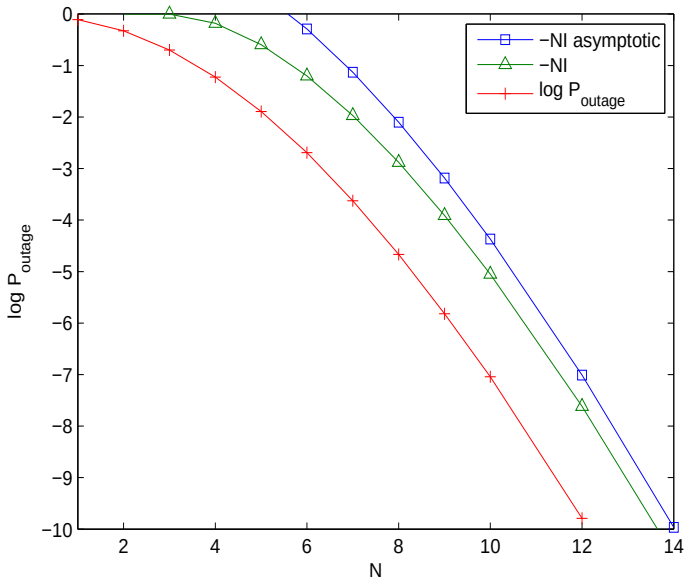

Figure 3.2: EPA with $\mathcal{P}_{tot} = 10m$W. Squares: (3.37) against N. Triangles: $-NI_{\sqrt{h}}(a/\sqrt{N})$ against N. Plus signs: $\log P_{outage}$ from Monte Carlo simulation. Simulation parameters: $\sigma = 0.0014$, $a = 0.003$, $\sigma_\theta^2 = 1$, $\sigma_i^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$.

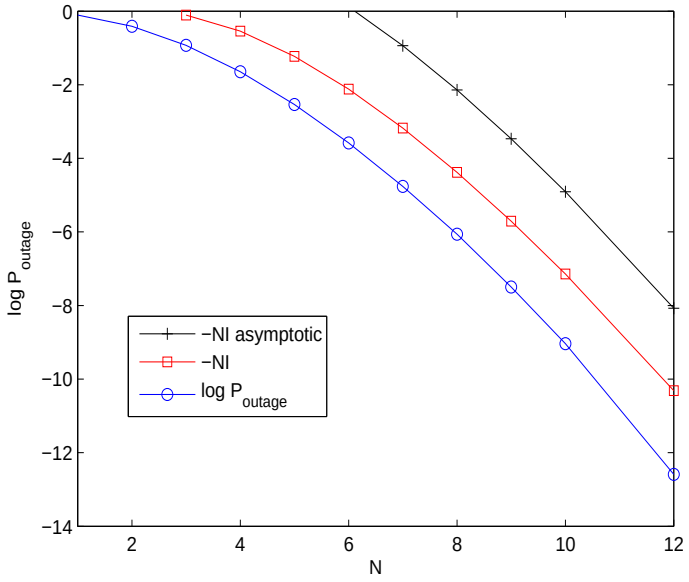

Figure 3.3: ST-OPA with $\mathcal{P}_{tot} = 10m$W. Plus signs: (3.52) against N. Squares: $-NI_Z(g/N)$ against N. Circles: $\log P_{outage}$ against N. Simulation parameters: $\lambda = 250,000$, $g = 0.09$, $\sigma_\theta^2 = 1$, $\sigma_i^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$.

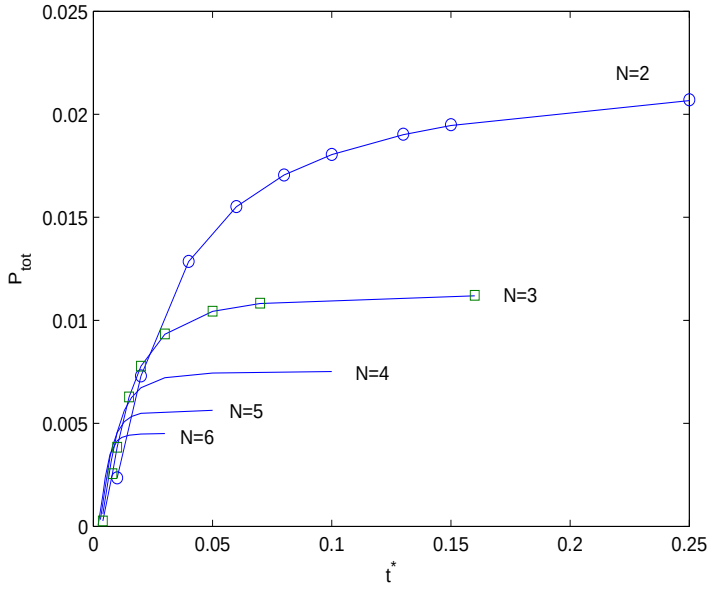

Figure 3.4: $\mathcal{P}_{tot}$ versus t^*. Circles and squares: from Monte Carlo simulation with 1,000,000 channel realizations. Solid lines: numerical solution of (3.64). Simulation parameters: $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$, $\lambda = 250,000$.

monotonically increasing function of t^* for any fixed N, and as t^* gets large, the value of $\mathcal{P}_{tot}$ saturates and approaches to some asymptotic value. The saturation behavior is due to $\frac{\Gamma(N-1,\lambda\zeta)}{\Gamma(N-1)} \to 1$ as $t \to \infty$ for fixed N, implying the existence of a finite $\mathcal{P}_{tot}$ achieving zero outage.

An approximate relationship between N_{max} and $\mathcal{P}_{tot}$ for LT-OPA has been obtained in (3.66). To see how good the approximation is, we plot (3.66) together with N_{max} obtained via Monte Carlo simulation, where we compute $E\left[\langle P_{tot}(\mathbf{h})\rangle\right]$ for a given N over 1,000,000 channel realizations. The results are shown in Fig. 3.5.

In Fig. 3.6 we compare the outage performance as a function of N for the three different power allocation schemes considered in this chapter, using $\mathcal{P}_{tot} = 1,600\mu$W. Note that for LT-OPA, due to the existence of N_{max}, the outage probability for $N > N_{max}$ is zero and hence we cannot show results for $N > N_{max}$ on the graph. In this example, $N_{max} = 15$. From this figure we can see that the gradients of EPA and ST-OPA are similar for large N, while the outage probability curve for LT-OPA approaches to a vertical asymptote located at N_{max}.

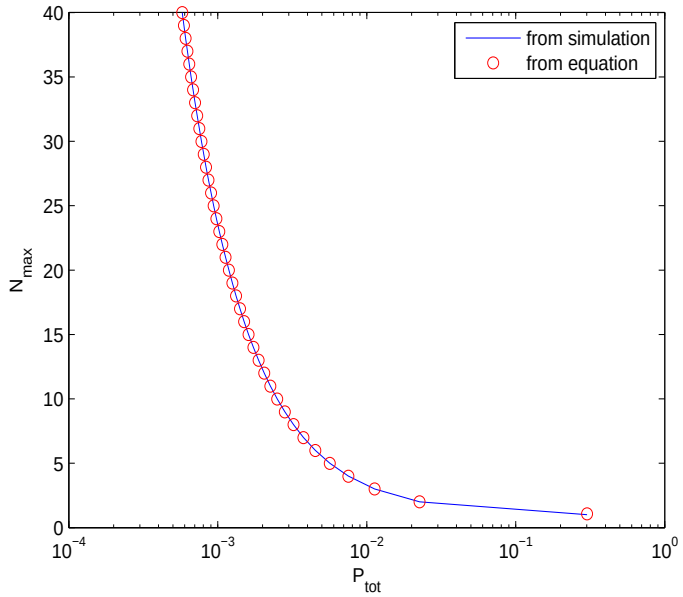

Figure 3.5: N_{max} versus $\mathcal{P}_{tot}$. Circles: (3.66) against $\mathcal{P}_{tot}$. Solid line: N_{max} from Monte Carlo simulation. Simulation parameters: $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$ and $\lambda = 250,000$.

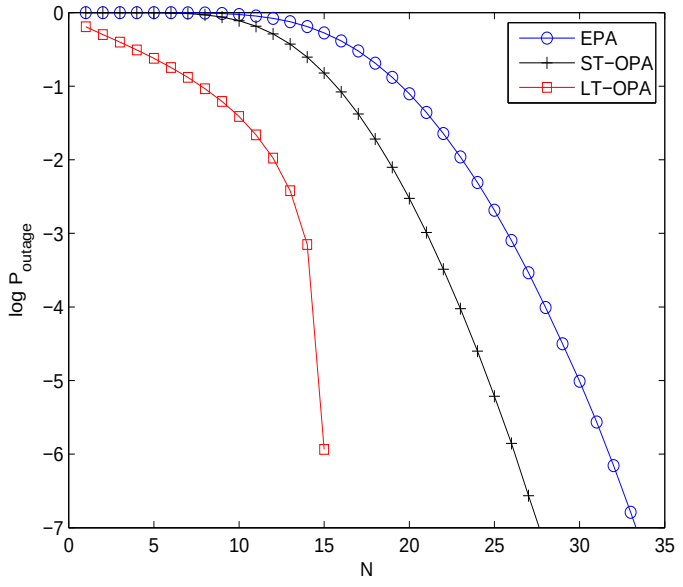

Figure 3.6: P_{outage} versus N. Simulation parameters: $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$, $\lambda = 250,000$ and $\mathcal{P}_{tot} = 1,600\mu$W.

3.5 Conclusions

In this chapter we have derived theoretical results on the diversity order of distortion outage in wireless sensor networks using different power allocation schemes. We presented three power allocation schemes - EPA, ST-OPA and LT-OPA. We then followed by presenting the theoretical results on the diversity order of distortion outage achieved by each of the power allocation schemes under Rayleigh fading. The equal power allocation asymptotically achieves a diversity order of $N \log N$, which is larger than the diversity order achieved by EPA in orthogonal MAC [22] by a factor of $\log N$. We have also shown that ST-OPA (minimizing distortion subject to a total power constraint) achieves the same diversity order of distortion outage as EPA. This suggests that in the case of a large number of sensors, the spatial diversity gain in EPA can overcome fading equally well as ST-OPA, which requires knowing CSIT. In the analysis of diversity order in LT-OPA, we found that the outage probability can be driven to zero with a finite amount of total power. We also obtained a closed form approximation to the minimum number of sensors that drives the outage probability to zero for a given total power constraint. Simulation results show that this approximation gives very close results to the true value.

Future extension of this work may include non-i.i.d. fading channels or different fading distributions. One may also extend this work to dynamical systems where the source is a time-varying Gauss Markov random process.

Chapter 4

Power allocation with correlated sensor data

4.1 Introduction

In the previous chapters we studied multiple-sensor estimation problems where the sensors observe a single Gaussian source and that their observations are spatially uncorrelated. However in many practical networks, especially for dense sensor networks, the sensor observations, although may differ slightly from one to another, can be spatially-correlated. This motivates us to study the power allocation that minimizes the distortion outage probability with correlated sensor data.

There are a number of existing works in the literature that considers correlation in wireless sensor networks. In [60] the authors exploit the sensor data correlation on the medium access control layer. They have developed a theoretical framework for transmission regulation of sensor nodes under a distortion constraint. They showed that a sensor node can act as a representative node for several other sensor nodes observing the correlated data, and the proposed medium access control can improve energy efficiency, packet drop rate and latency. The authors in [24] considered the case where the sensor noises are correlated and showed that the total energy consumption required for transmission in a sensor network can be minimized by considering jointly the number of quantization levels for each sensor and information about correlation of sensor observations. A more recent work in [52] considered optimal power allocation in linear estimation of correlated data in wireless sensor network using analog modulation and orthogonal multi-access transmission protocol. In this work the authors defined *averaged distortion measure* and solved optimization problems of minimizing total power subject to average distortion constraint. Problems involving correlation of data usually involve complex matrix algebra, and

closed-form solutions for problems with a general covariance matrix can be difficult to find. In [24, 52], they have both used a tridiagonal covariance matrix model for simulation due to the simplicity in calculating the inverse of tridiagonal matrices. Other practical models of the covariance matrix can be found in [60, 61].

In this chapter we consider the network model given in [52] which is summarized as follows. We assume a wireless sensor network where the sensor observations are jointly Gaussian, spatially correlated, and corrupted by i.i.d. AWGN. The sensors amplify and forward their signals to the FC by over orthogonal multi-access channels (e.g. FDMA). We assume block-fading channels and that the instantaneous CSI is available at the sensors (transmitters) and the FC (receiver), and that the sensor data correlation matrix is known and does not change in time. The distortion of the estimate computed at the FC using an MMSE estimator is a random quantity due to the random nature of the fading channels and hence may exceed the required distortion threshold with non-zero probability. In this chapter we will apply the techniques presented in Chapter 2 and find the optimal power allocation that minimizes the distortion outage probability (the probability that the average distortion exceeds a certain threshold) subject to a long-term average power constraint for the case of spatially-correlated sensor data. Simulation results show significant power savings can be achieved for highly-correlated sensor data.

The organization of this chapter is given as follows. In Section 4.2 we give the wireless sensor network model and the problem formulation. In Section 4.3 we present the power allocation schemes and solutions, followed by simulation results in Section 4.4. We conclude this chapter by giving some concluding remarks and discussions in Section 4.6.

4.2 Sensor network model and problem formulation

A schematic diagram of the wireless sensor network studied in this chapter is shown in Fig. 4.1. Assume that there are N sensors and each sensor observes a random

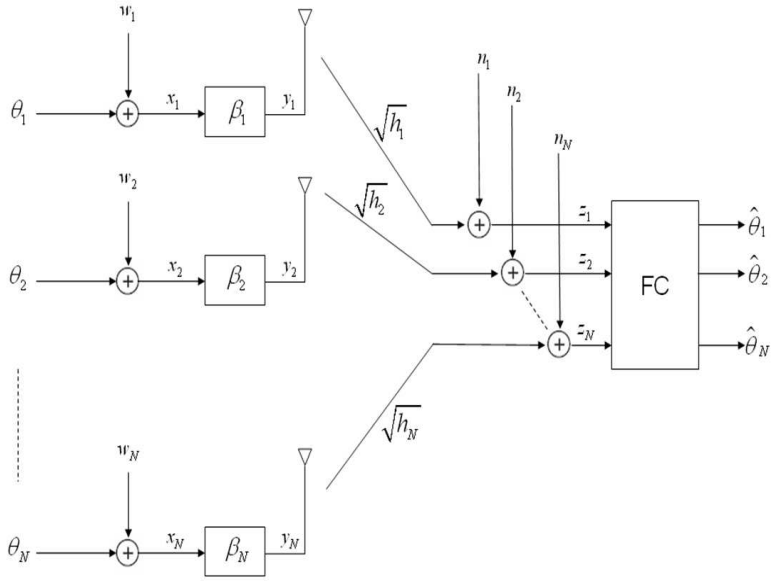

Figure 4.1: Schematic diagram of a wireless sensor network for distributed estimation

Gaussian process that is i.i.d in time. We assume that the sources may be spatially correlated, but the correlation pattern stays fixed in time. This model may be thought of as a wireless sensor network monitoring over a static spatially-correlated field.

Denote the realization of the ith source as θ_i and the respective measurement noise as w_i. The measurements of the sensors at discrete time $k = 0, 1, 2, \ldots$ are given as

$$\mathbf{x}[k] = \boldsymbol{\theta}[k] + \mathbf{w}[k]$$

where

$$\mathbf{x}[k] = [x_1[k], \ldots, x_N[k]]^T \tag{4.1}$$

$$\theta[k] = [\theta_1[k], \ldots, \theta_N[k]]^T \tag{4.2}$$

$$\mathbf{w}[k] = [w_1[k], \ldots, w_N[k]]^T, \tag{4.3}$$

where ^{T} denotes transposition. We assume that the observation noise $\mathbf{w}[k]$ is independent white Gaussian noise with zero mean and covariance matrix $\mathbf{C_w} = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$. We assume that the source $\boldsymbol{\theta}[k]$ is i.i.d Gaussian distributed random process with zero mean and autocorrelation matrix $\mathbf{C}_{\boldsymbol{\theta}} = E\left[\boldsymbol{\theta}\boldsymbol{\theta}^T\right]$. Note here

that since the source has zero mean, its autocorrelation matrix is identical to its covariance matrix. We assume that the sensors use amplify-and-forward scheme to transmit wirelessly their measurements to the fusion center through orthogonal block-fading channels. We denote the ith sensor's amplifier gain (voltage gain), the channel gain (voltage gain) and the channel noise of the ith channel as $\beta_i[k]$, $\sqrt{h_i[k]}$ and $n_i[k]$ respectively. Let $\mathbf{n}[k] = [n_1[k], \ldots, n_N[k]]^T$. We assume that the channel noise $\mathbf{n}[k]$ is independent white Gaussian noise with zero mean and covariance matrix $\mathbf{C_n} = \text{diag}(\xi_1^2, \ldots, \xi_N^2)$. We assume that the instantaneous CSI is known at both the transmitters (sensors) and the receiver (FC). The fusion center receives a vector of signals given as (for brevity we discard the time index k)

$$\mathbf{z} = \mathbf{H}\boldsymbol{\theta} + \mathbf{v} \tag{4.4}$$

where

$$\mathbf{z} = [z_1, \ldots, z_N]^T \tag{4.5}$$

$$\mathbf{H} = \text{diag}\left(\beta_1\sqrt{h_1}, \ldots, \beta_N\sqrt{h_N}\right) \tag{4.6}$$

$$\mathbf{v} = \left[\beta_1\sqrt{h_1}w_1 + n_1, \ldots, \beta_N\sqrt{h_N}w_N + n_N\right]^T. \tag{4.7}$$

Here $\mathbf{H}$ is a deterministic matrix, $\boldsymbol{\theta}$ is multivariate Gaussian with zero mean and covariance $\mathbf{C}_{\boldsymbol{\theta}}$ and $\mathbf{v}$ is white Gaussian noise with known covariance matrix $\mathbf{C_v} = \text{diag}\left(\beta_1^2 h_1 \sigma_1^2 + \xi_1^2, \ldots, \beta_N^2 h_N \sigma_N^2 + \xi_N^2\right)$. Since $\boldsymbol{\theta}$ and $\mathbf{v}$ are Gaussian and mutually uncorrelated, they are jointly Gaussian. Hence $\mathbf{z}$ and $\boldsymbol{\theta}$ are also jointly Gaussian, and we can therefore use MMSE estimator at the fusion center, which is the optimal estimator in this case, and reconstruct an estimate of the source vector $\hat{\boldsymbol{\theta}}$ given as

$$\hat{\boldsymbol{\theta}} = \mathbf{C}_{\boldsymbol{\theta}}\mathbf{H}^T\left(\mathbf{H}\mathbf{C}_{\boldsymbol{\theta}}\mathbf{H}^T\right)^{-1}\mathbf{z}$$

where $\hat{\boldsymbol{\theta}} = \left[\hat{\theta}_1, \ldots, \hat{\theta}_N\right]^T$. The estimation distortion is given by the error-covariance matrix associated with $\hat{\boldsymbol{\theta}}$ given as

$$\mathbf{D}_{\hat{\boldsymbol{\theta}}} = \left[\mathbf{H}^T\mathbf{C}_{\mathbf{v}}^{-1}\mathbf{H} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right]^{-1} \tag{4.8}$$

We define the average distortion measure $\bar{D}$ as

$$\begin{aligned}\bar{D} &= \frac{1}{N}E\left[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^T(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\right] \\ &= \frac{1}{N}\mathrm{tr}\left(E\left[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^T\right]\right) \\ &= \frac{1}{N}\mathrm{tr}\left(\mathbf{D}_{\hat{\boldsymbol{\theta}}}\right) \\ &= \frac{1}{N}\mathrm{tr}\left(\mathbf{H}^T\mathbf{C}_{\mathbf{v}}^{-1}\mathbf{H}+\mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1}\end{aligned}$$

Due to the randomly-varying nature of the wireless channels, the reconstructed sources at FC are also random quantities whose average distortion may be different upon each reception of the signals transmitted by the sensors. One commonly used performance metric is the distortion outage probability, which is defined as the probability that the average distortion $\bar{D}$ of the reconstructed source exceeds a given threshold $\bar{D}_{max}$. Note that sensors only have limited average power and the average distortion at the receiver may be very large when fading is severe in the wireless channels. The outage probability hence is a function of the channel gain distributions. In the following we formulate two problems - minimizing the short-term total transmit power subject to a distortion constraint and minimizing the distortion outage probability subject to a long-term average transmit power constraint. We assume that full CSI is available at both the sensors and FC, and we compute the sensors' transmit powers based on these instantaneous CSI.

4.2.1 Short-term power allocation

We would like to minimize the total transmit power of the sensors while satisfying an average distortion constraint at FC. We call this problem the *short-term* power allocation because we are primarily concerned with the sensor transmission powers for a single time instance k. The channel gain $\mathbf{h}$ here is invariant over each block due to block-fading. is We formulate this short-term power allocation as an optimization

problem given as

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} P_i \\ \text{s.t.} \quad & \tfrac{1}{N}\text{tr}\left(\mathbf{H}^T {\mathbf{C_v}}^{-1}\mathbf{H} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1} \leq \bar{D}_{max} \\ & P_i \geq 0 \quad \forall i \end{aligned} \tag{4.9}$$

where $P_i = W_i^2\beta_i^2$ and $W_i^2 = [\mathbf{C}_{\boldsymbol{\theta}}]_{i,i} + \sigma_i^2$. The notation $[\mathbf{A}]_{i,j}$ denotes the ith row and the jth column element of matrix $\mathbf{A}$. Note that this power allocation needs to be recomputed every time the sensors transmit their measurements to FC, since the fading gains would have changed in the next transmission.

Problem (4.9) has been solved in [52]. Here we simply give the solution of problem (4.9) in this chapter as it will be useful later. We first transform problem (4.9) into a convex optimization problem [52] given as

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} \frac{\alpha_i^2}{h_i}\left(\frac{r_i}{1-r_i\sigma_i^2}\right) \\ \text{s.t.} \quad & \tfrac{1}{N}\text{tr}\left(\mathbf{R} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1} \leq \bar{D}_{max} \\ & 0 \leq r_i \leq \tfrac{1}{\sigma_i^2} \quad \forall i \end{aligned} \tag{4.10}$$

where $\mathbf{R} = \text{diag}(r_1, \ldots, r_N)$, $r_i = \frac{P_i h_i}{P_i h_i \sigma_i^2 + \alpha_i^2}$ and $\alpha_i^2 = W_i^2\xi_i^2$. The solution can be obtained by solving the KKT conditions given as

$$\begin{aligned} & \frac{\alpha_i^2}{h_i}\left(\frac{1}{1-r_i\sigma_i^2}\right)^2 - \lambda_0\frac{1}{N}\left(\boldsymbol{e}_i^T\left(\mathbf{R} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-2}\boldsymbol{e}_i\right) - \gamma_i = 0 \\ & \frac{1}{N}\text{tr}\left(\mathbf{R} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1} = \bar{D}_{max} \\ & \gamma_i r_i = 0 \\ & \gamma_i \geq 0 \\ & \lambda_0 \geq 0 \end{aligned} \tag{4.11}$$

where γ_i and λ_0 are the Lagrangian multipliers, $\boldsymbol{e}_i$ is the $N \times 1$ vector which has its ith component equal to 1 and all of its other components equal to 0. For spatial correlated sources (when $\mathbf{C}_{\boldsymbol{\theta}}$ is not a diagonal matrix), the solution does not have a closed-form. Numerical methods may be used to solve for the unknowns. For

spatially uncorrelated sources, we may obtain the optimal power $\mathbf{P}^* = (P_1^*, \ldots, P_N^*)$ as follows [52]. With out loss of generality, order the channel gain vector $\mathbf{h}$ such that

$$\frac{\alpha_1}{[\mathbf{C}_{\boldsymbol{\theta}}]_{1,1}\sqrt{h_1}} \leq \frac{\alpha_2}{[\mathbf{C}_{\boldsymbol{\theta}}]_{2,2}\sqrt{h_2}} \leq \ldots \leq \frac{\alpha_N}{[\mathbf{C}_{\boldsymbol{\theta}}]_{N,N}\sqrt{h_N}} \tag{4.12}$$

The optimal power of the ith channel is given as

$$P_i^*(\mathbf{h}) = \frac{G_i}{\eta_i}\left[\frac{\sqrt{\eta_i}}{\rho_0(\mathbf{h}, N_1)} - 1\right]^+ \tag{4.13}$$

where $G_i = [\mathbf{C}_{\boldsymbol{\theta}}]_{i,i} / \left([\mathbf{C}_{\boldsymbol{\theta}}]_{i,i} + \sigma_i^2\right)$, $\eta_i = [\mathbf{C}_{\boldsymbol{\theta}}]_{i,i} h_i/\alpha_i^2$, $\rho_0(\mathbf{h}, N_1) = B(N_1)/A(\mathbf{h}, N_1)$, $B(N_1) = \sum_{j=1}^{N_1} G_j + N\bar{D}_{max} - \sum_{i=1}^{N} [\mathbf{C}_{\boldsymbol{\theta}}]_{i,i}$ and $A(\mathbf{h}, N_1) = \sum_{j=1}^{N_1} G_j/\sqrt{\eta_j}$. $[x]^+$ denotes $max(0, x)$. $N_1 \in \{1, \ldots, N\}$ can be obtained by solving $f(N_1) < 1$ and $f(N_1 + 1) \geq 1$, where $f(i) = \alpha_i B(i)\sqrt{N} / [\mathbf{C}_{\boldsymbol{\theta}}]_{i,i} A(\mathbf{h}, i)\sqrt{h_i}$.

4.2.2 Minimizing distortion outage probability

We denote the distortion outage probability as $\Pr(D(\mathbf{P}(\mathbf{h}), \mathbf{h}) > \bar{D}_{max})$, where $\Pr(x)$ denotes the probability that the event x occurs. We also define the long-term average transmit power as the power averaged over the total number of the sensors N and over time, denoted as $E[\langle \mathbf{P} \rangle]$ where $\mathbf{P} = [P_1, \ldots, P_N]^T$, $\langle \mathbf{x} \rangle$ denotes the arithmetic mean of the vector $\mathbf{x}$ of length M defined as $(1/M)\sum_{i=1}^{M} x_i$ and $E[x]$ denotes the expectation of x. The optimization problem that minimizes the distortion outage probability subject to a long-term average power constraint is given as

$$\begin{aligned} \min \quad & \Pr\left(D(\mathbf{P}(\mathbf{h}), \mathbf{h}) > \bar{D}_{max}\right) \\ \text{s.t.} \quad & E[\langle \mathbf{P} \rangle] \leq P_{av} \\ & \mathbf{P} \succeq 0 \end{aligned} \tag{4.14}$$

where P_{av} is the long-term average power threshold and $\succeq$ denotes component-wise inequality.

4.3 Power allocation schemes and solutions

Problem (4.14) can be solved in the same way as in [27]. First consider the following minimization problem given as

$$\begin{aligned} \min \quad & \langle \mathbf{P}(\mathbf{h}) \rangle \\ \text{s.t.} \quad & D(\mathbf{P}(\mathbf{h}), \mathbf{h}) \leq \bar{D}_{max} \\ & \mathbf{P}(\mathbf{h}) \succeq 0 \end{aligned} \tag{4.15}$$

We have the following lemma:

Lemma 4.3.1. *With the knowledge of* $\mathbf{h}$*, the solutions of (4.15) for spatially uncorrelated sources and spatially correlated sources are given by (4.13) and (4.11) respectively.*

One can also obtain the following Lemma which is necessary to find the optimal solution to problem (4.14).

Lemma 4.3.2. *The optimal power,* $\mathbf{P}^*(\mathbf{h}) = [P_1^*(\mathbf{h}), \ldots, P_N^*(\mathbf{h})]^T$*, is a continuous function of* $\mathbf{h}$*. Furthermore,* $\langle \mathbf{P}^*(\mathbf{h}) \rangle$ *is a non-increasing function of* h_i *for* $n = 1, \ldots, N$*. In fact, we can show that for all* $P_n^*(\mathbf{h})$*,* $n = 1, \ldots, N$*, the following is true:*

$$\frac{\partial \langle \mathbf{P}^*(\mathbf{h}) \rangle}{\partial h_n} = -\frac{1}{N} \frac{P_n^*}{h_n} \tag{4.16}$$

The proof of this lemma can be found in the appendix. We define two regions, $\mathcal{R}(s)$ and $\bar{\mathcal{R}}(s)$ and the boundary surface $\mathcal{B}(s)$ for some non-negative s as in [27]:

$$\begin{aligned} \mathcal{R}(s) &= \{\mathbf{h} \in \mathbb{R}_+^N : \langle \mathbf{P}(\mathbf{h}) \rangle < s\} \\ \overline{\mathcal{R}}(s) &= \{\mathbf{h} \in \mathbb{R}_+^N : \langle \mathbf{P}(\mathbf{h}) \rangle \leq s\} \\ \mathcal{B}(s) &= \{\mathbf{h} \in \mathbb{R}_+^N : \langle \mathbf{P}(\mathbf{h}) \rangle = s\} \end{aligned} \tag{4.17}$$

We then define two average power sums as

$$\begin{aligned} P(s) &= \int_{\mathcal{R}(s)} \langle \mathbf{P}(\mathbf{h}) \rangle dF(\mathbf{h}) \\ \overline{P}(s) &= \int_{\overline{\mathcal{R}}(s)} \langle \mathbf{P}(\mathbf{h}) \rangle dF(\mathbf{h}) \end{aligned} \tag{4.18}$$

where $F(\mathbf{h})$ denotes the c.d.f. of $\mathbf{h}$. Finally, the power sum threshold s^* and the weight w^* are given as

$$\begin{aligned} s^* &= \sup\{s : P(s) < P_{av}\} \\ w^* &= \frac{P_{av} - P(s^*)}{\overline{P}(s^*) - P(s^*)} \end{aligned} \tag{4.19}$$

With the above lemma and definitions we now present the solution to problem (4.14). The proof follows using similar techniques as in [27] and hence excluded.

Theorem 4.3.1. *The solution of problem (4.14) is given as*

$$\hat{\mathbf{P}}(\mathbf{h}) = \begin{cases} \mathbf{P}^*(\mathbf{h}), & \text{if } \mathbf{h} \in \mathcal{R}(s^*) \\ \mathbf{0}, & \text{if } \mathbf{h} \notin \overline{\mathcal{R}}(s^*) \end{cases} \tag{4.20}$$

while if $\mathbf{h} \in \mathcal{B}(s^*)$, $\hat{\mathbf{P}}(\mathbf{h}) = \mathbf{P}^*(\mathbf{h})$ with probability s^* and $\hat{\mathbf{P}}(\mathbf{h}) = 0$ with probability $1 - w^*$. The optimal power allocation scheme that minimizes the outage probability subject to a long-term average power constraint says that if the vector of channel gains falls inside the region defined by $\mathcal{R}(s^*)$, where s^* is a quantity that is associated with P_{av}, then the sensors should transmit with power given by (4.11) or (4.13) depending on whether the sources are spatially correlated. Otherwise, none should transmit to save power.

4.4 Simulation results

We modeled two wireless sensor networks, one with 2 sensors and the other with 4 sensors ($N = 2, 4$). For both sensor networks, half of the sensors are placed at a distance 500 meters from FC, and the other half 600 meters. The measurement noises are i.i.d. Gaussian of zero mean and variance 10^{-3} for all sensors. The AWGN channel noises are i.i.d. Gaussian of zero mean and variance 10^{-10} for all channels. We assume the wireless channels are Rayleigh-faded. The signal power gains are hence independently exponentially-distributed and we assume that the mean channel gains are equal to inverse of the transmission distances squared. We use a source correlation model taken from [52] where we define a single correlation

coefficient, ρ, and the autocorrelation matrix $\mathbf{C}_{\boldsymbol{\theta}}$ is given as

$$[\mathbf{C}_{\boldsymbol{\theta}}] = \rho^{|j-i|}, \quad \rho < 1 \tag{4.21}$$

We set $\rho = 0.3, 0.6$ and 0.9.

The minimum achievable average distortion at FC can be obtained by setting $P_{av} \to \infty$. This in effect eliminates the uncertainties caused by the channel noise. The minimum achievable average distortion does not become zero due to the existence of measurement noise. The minimum achievable average distortion $\bar{D}_{min}$ is given as

$$\bar{D}_{min} = \frac{1}{N}\text{tr}\left(\sigma^2\mathbf{I} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1} \tag{4.22}$$

where $\sigma^2 = 10^{-3}$ (measurement noise variance) and $\mathbf{I}$ is the identity matrix of size $N \times N$. Given the above parameters, $\bar{D}_{min} \approx 0.001$. Based on this fact, we set the average distortion constraint $\bar{D}_{max} = 0.05$. The following outage probability results are obtained using Monte Carlo simulation over 1,000,000 channel realizations. Closed form solutions of (4.10) for general correlation matrices do not exist and we use numerical tools in MATLAB in obtaining the optimal power allocations.

Fig. 4.2 shows the distortion outage probability against long-term average power for different correlation coefficients. The top curve indicate the outage performance achieved when the source data are uncorrelated. Not surprisingly, as the correlation coefficient increases, less long-term average power is required to achieve the same outage probability. Large power savings can be achieved when the correlation is high e.g. $\rho = 0.9$, while performances of uncorrelated and $\rho = 0.3, 0.6$ have very similar performance.

Fig. 4.3 shows the outage performance of 4 sensors measuring four correlated sources. The improvement in the power savings can be easily seen even for some correlation $\rho = 0.3, 0.6$, and a large power gain is obtained for $\rho = 0.9$. This is as a result of exploiting the correlation between the sources.

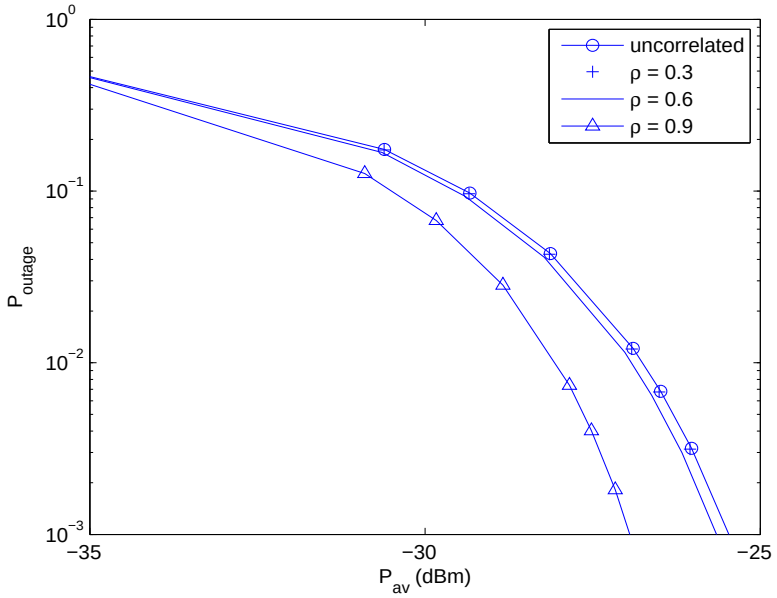

Figure 4.2: Outage performance of $N = 2$, $\bar{D}_{max} = 0.05$

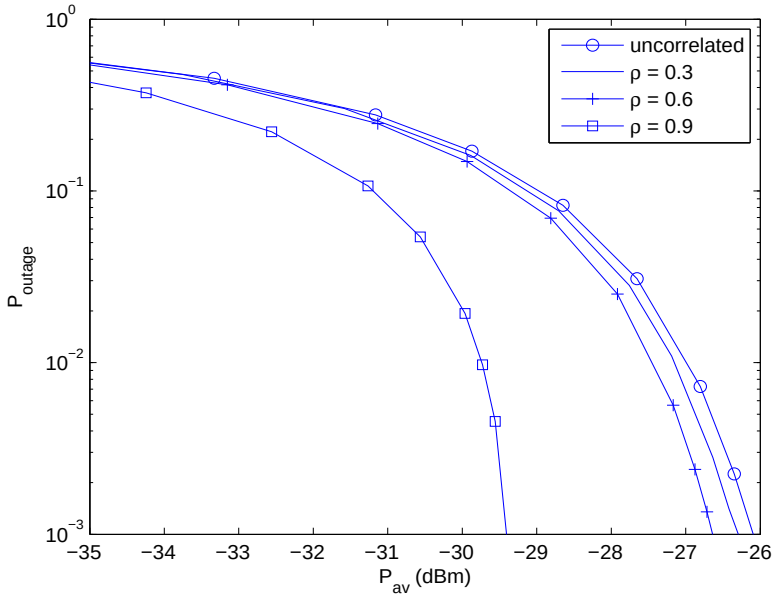

Figure 4.3: Outage performance of $N = 4$, $\bar{D}_{max} = 0.05$

4.5 Analysis of asymptotic average distortion for equal power allocation and some ideas for future work

The difficulty in finding the closed-form expressions of the optimal power allocations lies in the expression of the average distortion $\frac{1}{N}\text{tr}\left(\mathbf{R}+\mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1}$. Here we provide a rough idea for evaluating the approximate asymptotic average distortion for large number of sensors transmitting with constant equal powers for correlation matrices of the form given in 4.21. The inverse of the autocorrelation matrix is given as

$$\mathbf{C}_{\boldsymbol{\theta}}^{-1}=\begin{bmatrix} a_1 & b & 0 & \cdots & 0 \\ b & a_2 & b & \ddots & \vdots \\ 0 & b & \ddots & b & 0 \\ \vdots & \ddots & b & a_{N-1} & b \\ 0 & \cdots & 0 & b & a_N \end{bmatrix} \tag{4.23}$$

where

$$b=\frac{-\rho}{1-\rho^2},\ a_i=\begin{cases} \frac{1}{1-\rho^2}, & i=1,N \\ \frac{1+\rho^2}{1-\rho^2}, & i=2,\ldots,N-1 \end{cases} \tag{4.24}$$

Note that in the diagonal elements of the above matrix only the first and the last elements have different expressions from the rest of the elements in a_i. For large N, we assume that the effect of a_1 and a_N is diminished by the large numbers of a_i in between where $i=2,\ldots,N-1$ and we replace a_1 and a_N by $\frac{1+\rho^2}{1-\rho^2}$. This leads us to study the quantity $\mathbf{X}=\mathbf{X}_1+\mathbf{X}_2$ where

$$\mathbf{X}_1=\begin{bmatrix} \frac{1+\rho^2}{1-\rho^2} & \frac{-\rho}{1-\rho^2} & 0 & \cdots & 0 \\ \frac{-\rho}{1-\rho^2} & \frac{1+\rho^2}{1-\rho^2} & \frac{-\rho}{1-\rho^2} & \ddots & \vdots \\ 0 & \frac{-\rho}{1-\rho^2} & \ddots & \frac{-\rho}{1-\rho^2} & 0 \\ \vdots & \ddots & \frac{-\rho}{1-\rho^2} & \frac{1+\rho^2}{1-\rho^2} & \frac{-\rho}{1-\rho^2} \\ 0 & \cdots & 0 & \frac{-\rho}{1-\rho^2} & \frac{1+\rho^2}{1-\rho^2} \end{bmatrix} \tag{4.25}$$

and $\mathbf{X}_2 = diag(r_1, \ldots, r_N)$, where the r_i's are i.i.d.. Note that $\mathbf{X}_1$ is a tridiagonal Toeplitz matrix. We are interested in

$$\frac{1}{N}\text{tr}\left(\mathbf{X}^{-1}\right) \tag{4.26}$$

for large N.

Consider the Stieltjes transform of $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$,

$$G_X^N(z) = \frac{1}{N}\text{tr}\left(-z\mathbf{I}_N + \mathbf{X}\right)^{-1} \tag{4.27}$$

where $\mathbf{I}_N$ is the identity matrix of dimension N by N. Notice that

$$G_X^N(0) = \frac{1}{N}\text{tr}\left(\mathbf{X}\right)^{-1} \tag{4.28}$$

is the quantity that we are intersected in.

Under certain asymptotic freeness assumptions [62,63], $G_X(z) = \lim_{N\to\infty} G_X^N(z)$ can be determined from the asymptotic eigenvalue distributions of $\mathbf{X}_1$ and $\mathbf{X}_2$. More specifically, we can determine $G_X(0)$ as follows. From equations (23)-(24) of [64], $G_X(0)$ can be found by solving the following set of simultaneous equations in the variables $\{G_X(0), \rho_1, \rho_2\}$

$$G_X(0) = \frac{1}{1/\rho_1 + 1/\rho_2} \tag{4.29}$$

$$G_X(0) = E\left[\frac{1}{\mathbf{X}_1 + 1/\rho_1}\right] \tag{4.30}$$

$$G_X(0) = E\left[\frac{1}{\mathbf{X}_2 + 1/\rho_2}\right] \tag{4.31}$$

where $\mathbf{X}_i$ is a random variable distributed according to the asymptotic eigenvalue distribution of $\mathbf{X}_i$.

For the matrix $\mathbf{X}_1$ given by (4.25), it is well known (e.g. pp.34-35 of [65]) that the eigenvalues have the form

$$\lambda_j(\mathbf{X}_1) = \frac{1+\rho^2}{1-\rho^2} + \frac{2\rho}{1-\rho^2}\cos\left(\frac{\pi j}{N+1}\right),\ j = 1, \ldots, N. \tag{4.32}$$

After some manipulations, we can the find that the asymptotic eigenvalue distribution of $\mathbf{X}_1$ is given as

$$f_{X_1}(x_1) = \frac{1-\rho^2}{\pi} \frac{1}{\sqrt{4\rho^2 - \left(x_1\left(1-\rho^2\right) - \left(1+\rho^2\right)^2\right)^2}},\ \frac{1-\rho}{1+\rho} < x_1 < \frac{1+\rho}{1-\rho} \tag{4.33}$$

For the matrix $\mathbf{X}_2 = diag(r_1, \ldots, r_N)$, if we assume that the transmission powers P are equal and constant across all the sensors and over time, and that the channel gains, measurement noise and channel noise are i.i.d., then r_i's are i.i.d. with

$$r_i = \frac{Ph_i}{P\sigma^2 h_i + \alpha^2} \tag{4.34}$$

where $\sigma^2 = \sigma_i^2$, $\forall i$, denotes the common variance of the measurement noise and $\alpha^2 = W^2\xi^2$ where $W^2 = [\mathbf{C}_{\boldsymbol{\theta}}]_{i,i} + \sigma^2$ and $\xi^2 = \xi_i^2$, $\forall i$, is the common variance of the channel noise. For i.i.d. Rayleigh fading channels, h_i are exponentially distributed with parameter λ. We can find that

$$f_{X_2}(x_2) = \frac{P\alpha\lambda}{\left(P - P\sigma^2 x_2\right)^2} \exp\left(\frac{-\alpha\lambda x_2}{P - P\sigma^2 x_2}\right),\ 0 \le x_2 < \frac{1}{\sigma^2} \tag{4.35}$$

Hence solving the simultaneous equations (4.30) is equivalent to solving for $\{\rho_1, \rho_2\}$ the simultaneous equations

$$\frac{1}{1/\rho_1 + 1/\rho_2} = E\left[\frac{1}{X_1 + 1/\rho_1}\right] \tag{4.36}$$

$$\frac{1}{1/\rho_1 + 1/\rho_2} = E\left[\frac{1}{X_2 + 1/\rho_2}\right] \tag{4.37}$$

with the distributions of X_1 and X_2 given by (4.33) and (4.35) respectively. This can be done numerically. The quantity

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{tr}\left[\mathbf{X}^{-1}\right] = G_X(0) = \frac{1}{1/\rho_1 + 1/\rho_2} \tag{4.38}$$

can then be determined.

Fig. 4.4 shows how the average distortion behaves as a function of N for one realization of the channel gains for different values of ρ. We can observe that the av-

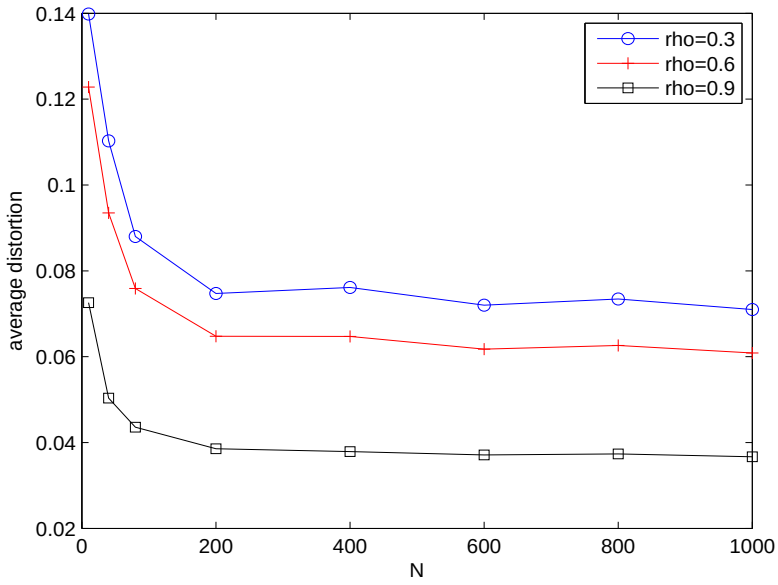

Figure 4.4: Average distortion against N for one realization of channel gains. Simulation parameters: $\sigma^2 = 10^{-3}$, $\xi^2 = 10^{-10}$, $P = 1m\text{W}$ and $\lambda = 250,000$.

erage distortions converges to some asymptotic values. Comparing these simulation results against the theoretical asymptotic average distortions computed by (4.38), given as 0.0775, 0.0656 and 0.0382 for $\rho = 0.3$, 0.6 and 0.9 respectively, shows they are closely matched. Similar result can be observed in Fig. 4.5 where the average distortion is plotted against N for different sensor power P. The asymptotic average distortions obtained by (4.38) are 0.1262, 0.0382 and 0.0091 for $P = 0.1m\text{W}$, $1m\text{W}$ and $10m\text{W}$ respectively.

Here we provide another idea for dealing with the inversion of symmetric tridiagonal matrices in the expression $\mathbf{tr}\left(\mathbf{Y}^{-1}\right)$ where

$$\mathbf{Y} = \begin{bmatrix} r_1 + a_1 & b & 0 & \cdots & 0 \\ b & r_2 + a_2 & b & \ddots & \vdots \\ 0 & b & \ddots & b & 0 \\ \vdots & \ddots & b & r_{N-1} + a_{N-1} & b \\ 0 & \cdots & 0 & b & r_N + a_N \end{bmatrix} \tag{4.39}$$

The study of analytic inversion of symmetric tridiagonal matrices can be found in [66, 67] where both references provided closed form expressions of the elements

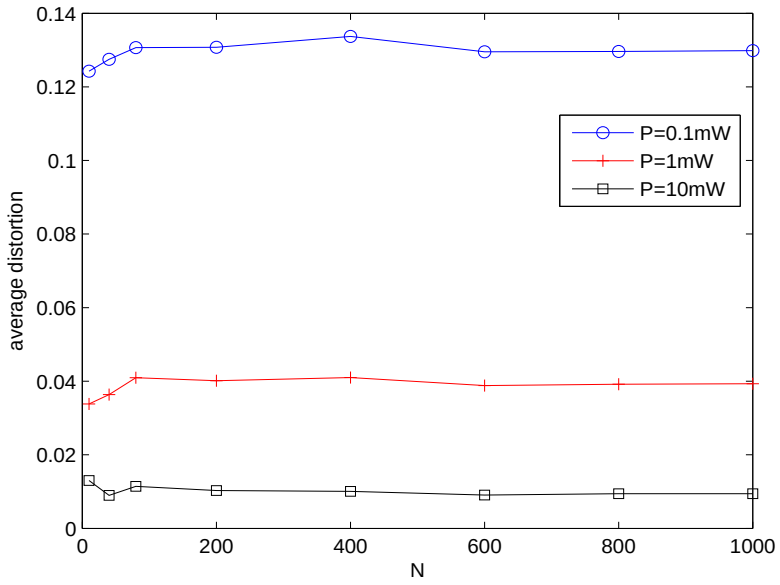

Figure 4.5: Average distortion against N for one realization of channel gains. Simulation parameters: $\sigma^2 = 10^{-3}$, $\xi^2 = 10^{-10}$, $\rho = 0.9$ and $\lambda = 250,000$.

of the inverse of symmetric tridiagonal matrices. Using the results in [66], we can express $\mathbf{tr}\left(\mathbf{Y}^{-1}\right)$ as

$$\mathbf{tr}\left(\mathbf{Y}^{-1}\right) = \sum_{i=1}^{N} \left[\mathbf{Y}^{-1}\right]_{i,i} = \sum_{i=1}^{N} \frac{d_{i+1} \cdots d_N}{\delta_i \cdots \delta_N} \tag{4.40}$$

where

$$\delta_1 = r_1 + a_1,\ \delta_i = r_i + a_i - \frac{b^2}{\delta_{i-1}},\ i = 2, \ldots, N \tag{4.41}$$

and

$$d_N = r_n + a_N,\ d_i = r_i + a_i - \frac{b^2}{d_{i+1}},\ i = N-1, \ldots, 1. \tag{4.42}$$

We observe that we can also represent the δ_i's and d_i's in terms of determinants of sub-matrices of $\mathbf{Y}$ as

$$\delta_i = \frac{\left|\mathbf{Y}^{(i)}\right|}{\left|\mathbf{Y}^{(i-1)}\right|},\ i = 1, \ldots, N \tag{4.43}$$

and

$$d_i = \frac{\left|\mathbf{Y}_{(N-i+1)}\right|}{\left|\mathbf{Y}_{(N-i)}\right|},\ i = 1, \ldots, N \tag{4.44}$$

where $\left|\mathbf{Y}^{(i)}\right| = |\mathbf{Y}(1:i, 1:i)|$ and $\left|\mathbf{Y}_{(i)}\right| = |\mathbf{Y}(N-i+1:N, N-i+1:N)|$ (using

the matrix notations in MATLAB). Note that by default $\left|\mathbf{Y}^{(0)}\right| = \left|\mathbf{Y}_{(0)}\right| = 1$. This analytic expressions of the inverse of the symmetric tridiagonal matrix may be another idea to consider for finding the analytic solution of the problems studied in this chapter.

4.6 Conclusions

In this chapter we investigated the power allocation scheme that minimizes the distortion outage probability for a wireless sensor network used in monitoring sources that are jointly Gaussian and spatially-correlated. We derived the necessary and sufficient conditions for solving such power allocation scheme subject to a long-term average power constraint. These conditions can be applied for any arbitrary fading distributions and correlation matrices, however closed form expressions could not be found. We present numerical simulation results based on a particular autocorrelation matrix and show that significant power savings can be achieved by exploiting the correlation structure of the sensor data, and that the higher the data are correlated, the more power can be saved.

For future work one may try to find the closed-form power allocation for general correlation matrices for arbitrary number of sensors in the network. One may also investigate asymptotic behavior of distortion outage probability and the diversity gain.

Chapter 5

Conclusions

This thesis studies power allocation problems in wireless sensor networks used in distributed estimation. Based on the amplify-and-forward framework, we investigated a number of energy-efficient power allocation algorithms that minimize distortion outage probability subject to long-term power constraint for different transmission protocols (coherent and orthogonal MAC), fading channels (Rayleigh fading, Nakagami-m fading and general fading distributions), limited feedback and correlation amongst sensor data. Below we shall give a summary of our work and some possible future research topics relating to our work in this thesis.

5.1 Summary

In Chapter 2 we studied outage-minimizing power allocation algorithms of a clustered wireless sensor network where the clusterhead transmitters send the observation signals from which they received from the sensors within clusters to the fusion center via an orthogonal MAC. We derived the optimal power allocation for fading channels under the assumption that the instantaneous CSI is available at both the cluster heads and the fusion center. We also studied some heuristic power allocation schemes based on the channel statistics. However simulation results showed that the performance of these schemes is significantly worse than the full CSI case. We then investigated the power allocation algorithm with limited feedback (partial CSI at the cluster heads) for Nakagami-m fading channels. We showed that it is difficult to find the optimal power allocation due to the highly irregularity in the structure of the quantization regions. Nonetheless we proposed a sub-optimal power allocation algorithm that have low computational complexity and simulation results showed

significant power gains can be achieved by using just a few number of feedback bits.

In Chapter 3 we assumed that the sensors directly send their observations to the fusion center through a coherent MAC. We proposed three power allocation algorithms based on full CSI and gave theoretical analysis on the diversity order of estimation outage for each of the power allocation schemes. We found that EPA scheme and the power allocation scheme that minimizes the distortion subject to a total power constraint (ST-OPA) achieve the same diversity order, showing that spatial diversity in EPA can overcome fading equally well as having knowledge of full CSI in ST-OPA for large number of sensors. On the other hand, the analysis for LT-OPA (minimizing outage probability subject to a long-term total power constraint) showed that we can drive the outage probability to zero using finite amount of long-term total power for networks with two or more number of sensors. We obtained an approximate expression for finding the required amount of long-term power for a given number of sensors and vice versa.

In Chapter 4 we exploited correlation amongst the sensor data and used this knowledge to further reduce the outage probability for a given average power. We studied a wireless sensor network where the sensors transmit their signals to the fusion center via orthogonal MAC. We derived the necessary and sufficient conditions for which the optimal power allocation must satisfy (which can be solved numerically) for any arbitrary fading distributions; however we did not find the closed-form expressions of the power allocation scheme. We present numerical results based on a particular correlation matrix and show that significant power savings can be achieved by taking into account the correlation structure of the sensor data.

5.2 Future research

For the work on optimal power allocation for distortion outage minimization, one may investigate how much additional energy is required to implement the algorithm, such as the energy used in the computation of transmission power and CSI acquisition, and see if there is an advantage in adopting such algorithm and by how much. The wireless sensor network considered in this problem only considers a single

Gaussian source. One may extend the problem to estimate a field of some physical phenomenon, which may have more practical applications. In the clustered wireless sensor network we have assumed that sensors within a cluster transmit with the same power and studied the effect of the sensor powers via simulation. One may investigate how the sensor powers should be adapted or allocated to transmit their signals to the cluster heads. Some other research areas include finding the optimal size of the clusters, the number of clusters and cross-layer optimization.

For the work on power allocation with limited feedback, due to the difficulties in evaluating the average power over the multi-dimensional channel state space, we have used a number of approximations and imposed extra constraints to simplify the problem. Although it provides a simple algorithm to implement, the performance that it can achieve is still sub-optimal. For example, in the case of 4-bit feedback, the outage performance achieved by SPSA (high computation complexity) is better than our proposed algorithm. One can investigate power allocation schemes that can perform better than our proposed algorithm and power allocation algorithms that can be applied to other fading channels (other than Nakagami-m). Optimal distortion-minimizing power allocation with limited feedback still remains as an open problem.

For the work on power allocation schemes for coherent MAC, we mainly focused our work on analyzing the diversity order of estimation outage based on the assumption that full CSI is available. In real networks having knowledge of CSI at the transmitters can be hard or expensive to implement as it requires a perfect feedback channel (zero-delay, error-free and infinite bandwidth). Therefore one may investigate the problem of energy-efficient power allocation schemes with limited feedback for sensor networks using coherent MAC protocol. In our work we have also assumed i.i.d. fading channels. Future extension of this work may include non-i.i.d. fading channels or different fading distributions. One may also extend this work to dynamical systems where the source is a time-varying Gauss Markov random process.

For the work on correlated sensor data, the closed-form optimal power allocation for a general correlation matrix is still unknown. One may investigate the asymptotic

behavior of average distortion outage probability and diversity gain. Other research directions include extending the problem to limited feedback and/or channel state estimation.

Another future work could look at energy minimization problem for different network lifetime definitions. The work in this thesis looks at the performance of the wireless sensor network with long-term average power constraint assuming all sensor nodes have power to perform distributed estimation. However as one sensor dies the network would have to continue with the remaining nodes provided that the distortion constraint can still be met. One may investigate optimizing the network lifetime where the lifetime could be defined as the first instant of time after which the distortion constraint can no longer be met.

Chapter 6

Appendix

6.1 Proof of Lemma 2.3.2

Proof. The first statement can be proved in a similar way to the one given in [27] and is omitted. The second statement can be proved by differentiating $\langle \mathbf{P}^*(\mathbf{h}) \rangle$ with respect to h_n $\forall n$ and show that it is non-positive. We begin with

$$\frac{\partial \langle \mathbf{P}^*(\mathbf{h}) \rangle}{\partial h_n} = \frac{1}{N} \frac{\partial}{\partial h_n} \sum_{i=1}^{N} P_i^*(\mathbf{h}) \tag{6.1}$$

$$= \frac{1}{N} \left(\sum_{i=1, i \neq n}^{N} \frac{\partial P_i^*(\mathbf{h})}{\partial h_n} + \frac{\partial P_n^*(\mathbf{h})}{\partial h_n} \right) \tag{6.2}$$

We will obtain the two derivatives inside the bracket of (6.2) separately. It is straight forward to show that the first term in the bracket is equal to

$$\frac{\partial P_i^*(\mathbf{h})}{\partial h_n}_{i \neq n} = -\frac{G_i G_n}{2D(N_1)\sqrt{\bar{\eta}_i} h_n \sqrt{\bar{\eta}_n}} \tag{6.3}$$

and the second term in the bracket is

$$\frac{\partial P_n^*(\mathbf{h})}{\partial h_n} = \frac{C_n G_n}{h_n \bar{H}_n} \left(1 - \frac{\sqrt{\bar{\eta}_n}}{2\bar{\rho}_0} - \frac{G_n}{2D(N_1)} \right). \tag{6.4}$$

For $n > N_1$,

$$\frac{\partial \langle \mathbf{P}^*(\mathbf{h}) \rangle}{\partial h_n} = \frac{1}{N} \left(\sum_{i=1, i \neq n}^{N_1} \frac{\partial P_i^*(\mathbf{h})}{\partial h_n} + \frac{\partial P_n^*(\mathbf{h})}{\partial h_n} \right) = \frac{1}{N} \left(\sum_{i=1, i \neq n}^{N_1} \frac{\partial P_i^*(\mathbf{h})}{\partial h_n} \right) \tag{6.5}$$

$$= -\frac{1}{N} \sum_{i=1, i \neq n}^{N_1} \frac{G_i G_n}{2D(N_1)\sqrt{\bar{\eta}_i} h_n \sqrt{\bar{\eta}_n}} \leq 0. \tag{6.6}$$

For $n \leq N_1$,

$$\frac{\partial \langle \mathbf{P}^*(\mathbf{h}) \rangle}{\partial h_n} = \frac{1}{N}\left(\sum_{i=1, i\neq n}^{N} \frac{\partial P_i^*(\mathbf{h})}{\partial h_n} + \frac{\partial P_n^*(\mathbf{h})}{\partial h_n} \right) \tag{6.7}$$

$$= \frac{1}{N}\left(-\sum_{i=1, i\neq n}^{N} \frac{G_i G_n}{2D(N_1)\sqrt{\bar{\eta}_i} h_n \sqrt{\bar{\eta}_n}} + \frac{C_n G_n}{h_n \bar{H}_n}\left(1 - \frac{\sqrt{\bar{\eta}_n}}{2\bar{\rho}_0} - \frac{G_n}{2D(N_1)}\right)\right) \tag{6.8}$$

$$= -\frac{1}{N}\left(\sum_{i=1}^{N} \frac{G_i G_n}{2D(N_1)\sqrt{\bar{\eta}_i} h_n \sqrt{\bar{\eta}_n}} + \frac{C_n G_n}{h_n \bar{H}_n}\left(\frac{\sqrt{\bar{\eta}_n}}{2\bar{\rho}_0} - 1\right)\right) \tag{6.9}$$

$$= -\frac{G_n}{2D(N_1) N h_n \sqrt{\bar{\eta}_n}}\left(2C(N_1) - \frac{2D(N_1)}{\sqrt{\bar{\eta}_n}}\right) \tag{6.10}$$

$$= -\frac{G_n}{N h_n \bar{\eta}_n}\left[\frac{\sqrt{\bar{\eta}_n}}{\bar{\rho}_0} - 1\right]^+ \leq 0. \tag{6.11}$$

Hence $\frac{\partial \langle \mathbf{P}^*(\mathbf{h}) \rangle}{\partial h_n} \leq 0 \ \forall n$ and this completes the proof of the second statement. □

6.2 Proof of Lemma 2.4.1

Proof. Recall the cdf expressed in the 'infinite-sum-series' form given in (2.33). Note that when $j = 1$, the expression corresponds to the outage probability, i.e., $\bar{F}_N(\mathbf{s}_1^{(N)}) \equiv P_{outage}$. As $P_{av} \to \infty$, $s_{i,j} \to 0$, and (2.33) can be simplified as,

$$\bar{F}_N(\mathbf{s}_j^{(N)}) = \frac{\prod_{i=1}^{N} (m_i \lambda_i s_{i,j})^{m_i}}{\Gamma\left(1 + \sum_{i=1}^{N} m_i\right)} \times \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \frac{\left[\prod_{i=1}^{N} (m_i)_{n_i} \left(-m_i \lambda_i s_{i,j}\right)^{n_i} \frac{1}{n_i!}\right]}{\left(1 + \sum_{i=1}^{N} m_i\right)_{n_T}} \tag{6.12}$$

$$\approx \frac{\prod_{i=1}^{N} (m_i \lambda_i s_{i,j})^{m_i}}{\Gamma\left(1 + \sum_{i=1}^{N} m_i\right)} \triangleq \tilde{F}_N(\mathbf{s}_j^{(N)}) \tag{6.13}$$

The partial derivative of $\tilde{F}_N(\mathbf{s}_j^{(N)})$ w.r.t. $s_{i,j}$ is given as

$$\frac{\partial \tilde{F}_N(\mathbf{s}_j^{(N)})}{\partial s_{i,j}} = \tilde{F}_N(\mathbf{s}_j^{(N)}) \cdot \frac{m_i}{s_{i,j}} \tag{6.14}$$

Substituting (6.14) into the KKT conditions (2.32) gives

$$\begin{aligned}
&\frac{\partial \Lambda_j}{\partial s_{i,j}} \Bigg/ \frac{\partial \tilde{F}_N(\mathbf{s}_j)}{\partial s_{i,j}} = \frac{\partial \Lambda_j}{\partial s_{k,j}} \Bigg/ \frac{\partial \tilde{F}_N(\mathbf{s}_j)}{\partial s_{k,j}} \\
&\Rightarrow -\frac{\phi_i}{s_{i,j}^2} \Bigg/ \frac{\tilde{F}_N(\mathbf{s}_j) m_i}{s_{i,j}} = -\frac{\phi_k}{s_{k,j}^2} \Bigg/ \frac{\tilde{F}_N(\mathbf{s}_j) m_k}{s_{k,j}} \\
&\Rightarrow P_{k,j} = \frac{m_k}{m_i} P_{i,j} \quad \forall i,k \in \{1,\ldots,N\}, j \in \{1,\ldots,L\}
\end{aligned} \tag{6.15}$$

If all $m_i = m_k \forall i,k$, then $P_{i,j} = P_{k,j}$, $\forall i,j,k$. This completes the proof. □

6.3 Proof of Lemma 2.4.2

Proof. In *Lemma 2.4.1* we obtained $\tilde{F}_N(\mathbf{s}_j^{(N)})$ and showed that as $P_{av} \to \infty$ it is asymptotically optimal to transmit with $P_{i,j} = \frac{m_i}{m_k} P_{k,j}$. We can use this result to express any $P_{i,j}$ in terms of $P_{1,j}$ $\forall i,j$. Hence instead of dealing with a vector space, we can reduce the problem down to a scalar problem by expressing (6.13) as a function of $P_{1,j}$ given as

$$\tilde{F}_N(\mathbf{P}_j) = \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} \cdot \frac{m_1^Q}{P_{1,j}^Q} \tag{6.16}$$

where $Q = \sum_{i=1}^{N} m_i$. The quantity $\Lambda_j = \sum_{i=1}^{N} P_{i,j}$ can also be written as a function of $P_{1,j}$ given as

$$\Lambda_j = \frac{Q}{m_1} P_{1,j} \tag{6.17}$$

As $P_{av} \to \infty$ the channel thresholds become small, $s_{i,j} \to 0$ $\forall i,j$. Hence the long-term average power in each quantized region can be approximated to be the same (EPPR), as shown in [31]. Applying (6.16) and (6.17) to the constraints in problem (2.35), we can derive the outage probability as a function of P_{av}, N and L. This expression is also used to obtain the diversity gain of the network. Starting from the last equation in (2.35), as $P_{av} \to \infty$, $s_{i,j} \to 0$,

$$\Lambda_L \left(1 - \tilde{F}_N(\mathbf{P}_L)\right) \approx \frac{N P_{av}}{L}$$

$$\Rightarrow \frac{Q\phi_1}{m_1 s_{1,L}} \left(1 - \left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,L}^Q \right) \approx \frac{NP_{av}}{L}$$

$$\Rightarrow 1 - \left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,L}^Q \approx \frac{m_1 NP_{av}}{\phi_1 QL} s_{1,L} \tag{6.18}$$

Since $s_{1,L}$ is small, $s_{1,L}^Q << s_{1,L}$ and we can discard the term with $s_{1,L}^Q$ in (6.18). After rearranging we obtain an expression of $s_{1,L}$ given as

$$s_{1,L} \approx \frac{\phi_1 QL}{m_1 NP_{av}} \tag{6.19}$$

Applying (6.16) and (6.17) to the constraint with $j = L-1$ in problem (2.35) gives

$$\frac{\phi_1 Q}{m_1 s_{1,L-1}} \left[\left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,L}^Q - \left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,L-1}^Q \right] = \frac{NP_{av}}{L}$$

$$\Rightarrow \left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,L}^Q = \left(\frac{m_1 NP_{av}}{\phi_1 LQ} + \left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,L-1}^{Q-1} \right) s_{1,L-1}$$

$$\Rightarrow \left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,L}^Q \approx \frac{m_1 NP_{av}}{\phi_1 LQ} s_{1,L-1}$$

$$\Rightarrow s_{1,L-1} \approx \left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\lambda_i \phi_i)^{m_i}}{\Gamma(1+Q)} \left(\frac{\phi_1 LQ}{m_1 NP_{av}} \right)^{Q+1} \tag{6.20}$$

where the last line is obtained after substituting (6.19). Repeating the above steps for the remaining constraints in (2.35), we can obtain

$$s_{1,1} \approx \left(\left(\frac{m_1}{\phi_1} \right)^Q \frac{\prod_{i=1}^{N} (\phi_i \lambda_i)^{m_1}}{\Gamma(1+Q)} \right)^{Q^{L-2}+\cdots+Q+1} \times \left(\frac{\phi_1 LQ}{m_1 NP_{av}} \right)^{Q^{L-1}+\cdots+Q+1} \tag{6.21}$$

and the outage probability is

$$\begin{aligned}\tilde{F}_N(s_{1,1}) &\approx \left(\frac{m_1}{\phi_1}\right)^Q \frac{\prod_{i=1}^{N}(\lambda_i\phi_i)^{m_i}}{\Gamma(1+Q)} s_{1,1}^Q \\ &\approx \left(\frac{\prod_{i=1}^{N}(\lambda_i\phi_i)^{m_i}}{\Gamma(1+Q)}\right)^{Q^{L-1}+\cdots+Q+1} \times \left(\frac{LQ}{NP_{av}}\right)^{Q^L+\cdots+Q} \end{aligned} \tag{6.22}$$

□

6.4 Proof of Theorem 2.4.1

Proof. Let $J(Q) = Q^L + \cdots + Q^2 + Q$. The diversity gain for the limited-feedback system can be obtained by substituting (6.22) to (2.41) and is given as

$$\begin{aligned} d \approx & -\lim_{P_{av}\to\infty} \frac{\log\left(\left(\frac{\prod_{i=1}^{N}(\lambda_i\phi_i)^{m_i}}{\Gamma(1+Q)}\right)^{J(Q)/Q}\left(\frac{LQ}{NP_{av}}\right)^{J(Q)}\right)}{\log P_{av}} \\ = & -\lim_{P_{av}\to\infty}\left(\frac{\frac{J(Q)}{Q}\left[\log\prod_{i=1}^{N}(\lambda_i\phi_i)^{m_i} - \log\Gamma(1+Q)\right]}{\log P_{av}} + \frac{J(Q)\left[\log L + \log Q - \log N - \log P_{av}\right]}{\log P_{av}}\right) \\ = & J(Q) \end{aligned} \tag{6.23}$$

□

6.5 Proof of Lemma 3.2.2

Proof. In the first statement it is immediate to see that $\mathbf{P}^*(\mathbf{h})$ is a continuous function of $\mathbf{h}$. In the second statement we need to show that $\langle\mathbf{P}(\mathbf{h})\rangle$ is a non-increasing function of h_i, $i = 1, \ldots, N$. We begin with the partial derivative of the short-term

average power given as

$$\frac{\partial \langle \mathbf{P}(\mathbf{h}) \rangle}{\partial h_i} = \frac{\partial}{\partial h_i} \frac{P_{tot}(\mathbf{h})}{N} = \frac{\sigma_c^2}{N} \frac{\partial \nu}{\partial h_i} \tag{6.24}$$

where $\nu = \frac{\mathcal{P}_{tot}}{\sigma_c^2}$ is the Lagrangian multiplier in one of the KKT conditions (see [23]). Also from the KKT conditions [23] we have

$$\sum_{i=1}^{N} \frac{\nu h_i}{C + \nu h_i \sigma^2} = \gamma_{th}. \tag{6.25}$$

Taking the partial derivative w.r.t. h_i on both sides of (6.25) gives

$$\begin{aligned}
&\frac{\partial}{\partial h_i} \sum_{j=1}^{N} \frac{\nu h_j}{C + \nu h_j \sigma^2} = 0 \Rightarrow \frac{\partial}{\partial h_i} \frac{\nu h_i}{C + \nu h_i \sigma^2} + \sum_{j=1}^{N} \frac{\partial}{\partial \nu} \frac{\nu h_j}{C + \nu h_j \sigma^2} \frac{\partial \nu}{\partial h_i} = 0 \\
\Rightarrow &\frac{\nu C}{\left(C + \nu h_i \sigma^2\right)^2} + \sum_{j=1}^{N} \frac{C h_j}{\left(C + \nu h_j \sigma^2\right)^2} \frac{\partial \nu}{\partial h_i} = 0 \\
\Rightarrow &\frac{\partial \nu}{\partial h_i} = -\frac{\nu C}{\left(C + \nu h_i \sigma^2\right)^2} \left(\sum_{j=1}^{N} \frac{C h_j}{\left(C + \nu h_j \sigma^2\right)^2} \right)^{-1} < 0 \Rightarrow \frac{\partial \langle \mathbf{P}(\mathbf{h}) \rangle}{\partial h_i} = \frac{\sigma_c^2}{N} \frac{\partial \nu}{\partial h_i} < 0
\end{aligned}$$

which completes the proof. □

6.6 Proof of Theorem 3.3.1

Proof. We prove the theorem by obtaining upper and lower bounds of $\log \Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \leq a_n\right)$, which asymptotically are equivalent for large n. The proof uses similar techniques to those provided in the proof of Theorem 5.11.4 in [68].

Upper bound. Assume that $X_1, X_2, \ldots$ are i.i.d. distributed random variables with a common c.d.f. and p.d.f. denoted as $F_X(x)$ and $f_X(x)$ respectively. Denote μ_X as the mean of X_i. Let $Y_i = -X_i + \mu_X$, hence $E[Y_i] = \mu_Y = 0$. The transformation allows us to obtain the following relationships $M_Y(t) = e^{\mu_X t} M_X(-t)$, $\Lambda_Y(t) = \mu_X t + \Lambda_X(-t)$ and

$$I_Y(c_n) = \sup_{-t} \left\{ (\mu_X - c_n)\, t - \Lambda_X(t) \right\}. \tag{6.26}$$

Note that $c_n = \mu_X - a_n$.

We prove first that $I_Y(c_n) > 0$ under the assumptions of the theorem. We note that $c_n t - \Lambda(t) = \log\left(\frac{e^{c_n t}}{M_Y(t)}\right) = \log\left(\frac{1+c_n t+o(t)}{1+\frac{1}{2}\sigma_Y^2 t^2+o(t^2)}\right)$ for small positive t, where $\sigma_Y^2 = \mathrm{var}(Y)$; we have used here the assumption that $M_Y(t) < \infty$ near the origin. For sufficiently small positive t, $1+c_n t+o(t) > 1+\frac{1}{2}\sigma_Y^2 t^2+o(t^2)$, whence $I_Y(c_n) > 0$ by the definition of the rate function.

We make two notes for future use. First, since $\Lambda_Y(t)$ is convex with $\Lambda_Y'(0) = E[Y] = 0$, and since $c_n > \mu_Y = 0$ for $n \geq \mathscr{N}$ (the value of $\mathscr{N}$ can be found by solving for the smallest integer n such that $c_n > 0$), the supremum of $c_n t - \Lambda_Y(t)$ over $t \in \mathbf{R}$ is unchanged by the restriction $t > 0$, which is to say that

$$I_Y(c_n) = \sup_{t>0} \left\{c_n t - \Lambda_Y(t)\right\}, \quad c_n > 0 \text{ for } n \geq \mathscr{N}. \tag{6.27}$$

Secondly, $\Lambda_Y(t)$ is strictly convex wherever the second derivative $\Lambda_Y''(t)$ exists. To see this, note that $\mathrm{var}(Y) > 0$ under the hypothesis of the theorem and

$$\Lambda_Y''(t) = \frac{M_Y(t)M_Y''(t) - M_Y'(t)^2}{M_Y(t)^2} = \frac{E\left[e^{tY}\right] E\left[Y^2 e^{tY}\right] - E\left[Y e^{tY}\right]^2}{M_Y(t)^2} > 0 \tag{6.28}$$

where the inequality is due to the Cauchy-Schwartz inequality applied to the random variables $Y e^{\frac{1}{2}tY}$ and $e^{\frac{1}{2}tY}$.

We have the following

$$\Pr\left(\frac{1}{n}\sum_{i=1}^n X_i \leq a_n\right) = \Pr\left(\sum_{i=1}^n Y_i \geq n c_n\right) = \Pr\left(e^{t\sum_{i=1}^n Y_i} \geq e^{n c_n t}\right) \quad \text{for } t > 0 \tag{6.29}$$

$$\leq \frac{E\left[\exp\left(t\sum_{i=1}^n Y_i\right)\right]}{e^{n c_n t}} = e^{-n c_n t} M_Y(t)^n = e^{-n(c_n t - \Lambda_Y(t))} \tag{6.30}$$

where the inequality is due to Markov's inequality. Taking log on both sides gives

$$\log \Pr\left(\frac{1}{n}\sum_{i=1}^n X_i \leq a_n\right) \leq -n\left(c_n t - \Lambda_Y(t)\right) \quad \forall t > 0 \tag{6.31}$$

Since the upper bound in (6.31) is true for all $t > 0$ and we are looking for the

tightest bound, we can further bound the LHS by taking the infimum on the RHS

$$\log \Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \leq a_n\right) \leq \inf_{t>0}\left\{-n\left(c_n t - \Lambda_Y(t)\right)\right\} = -n\sup_{t>0}\left\{c_n t - \Lambda_Y(t)\right\} \quad (6.32)$$

$$= -nI_Y^+(c_n) = -nI_X^-(a_n) \quad \text{from (6.26)} \quad (6.33)$$

Lower bound. We first show that the problem falls under the regular case, i.e., that the supremum of the rate function $I_Y(c_n)$, $n \geq \mathscr{N}$ is attained at some point $\tau \in (0, \infty)$. Denote $F_Y(y)$ and $f_Y(y)$ the common c.d.f. and p.d.f. of $Y_1, Y_2, \ldots$ respectively. Since $\Pr(Y_i > c_n) > 0$ for $n \geq \mathscr{N}$, there exists $b_n \in (c_n, \infty)$ such that $\Pr(Y_i > b_n) > 0$.

It follows that for $t > 0$,

$$c_n t - \Lambda_Y(t) = c_n t - \log E\left[e^{tY}\right] = c_n t - \log \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy \quad (6.34)$$

$$= c_n t - \log\left[\int_{-\infty}^{b_n} e^{ty} f_Y(y) dy + \int_{b_n}^{\infty} e^{ty} f_Y(y) dy\right] \quad (6.35)$$

$$\leq c_n t - \log \int_{b_n}^{\infty} e^{ty} f_Y(y) dy \leq c_n t - \log\left\{e^{tb_n} \int_{b_n}^{\infty} f_Y(y) dy\right\} \quad (6.36)$$

$$= c_n t - \log\left\{e^{tb_n} \Pr(Y_i > b_n)\right\} \quad (6.37)$$

$$= -(b_n - c_n)t - \log \Pr(Y_i > b_n) \to -\infty \text{ as } t \to \infty \quad (6.38)$$

since $b_n - c_n > 0$ for finite and fixed n. We deduce that the supremum of $c_n t - \Lambda_Y(t)$ over values $t > 0$ is attained at some point $\tau_n \in (0, \infty)$. The random sequence $Y_1, Y_2, \ldots$ is therefore a regular case of the large deviation problem.

We now introduce an ancillary random variable (as a function of n) $\tilde{Y}_n$ with distribution function $F_{\tilde{Y}_n}(y)$, sometimes called an 'exponential change of distribution' or a 'tilted distribution', by

$$dF_{\tilde{Y}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)} dF_Y(y) \quad (6.39)$$

which can also be interpreted as $F_{\tilde{Y}_n}(y) = \frac{1}{M_Y(\tau_n)} \int_{-\infty}^{y} e^{\tau_n u} dF_Y(u)$. Let $\tilde{Y}_{n,1}, \tilde{Y}_{n,2}, \ldots$ be i.i.d. distributed with c.d.f. $F_{\tilde{Y}_n}$. We note the following properties of $\tilde{Y}_{n,i}$. The

moment generating function of $\tilde{Y}_{n,i}$ is

$$M_{\tilde{Y}_n}(t) = \int_{-\infty}^{\infty} e^{tu} dF_{\tilde{Y}_n}(u) = \int_{-\infty}^{\infty} \frac{e^{(t+\tau_n)u}}{M_Y(\tau_n)} dF_Y(u) = \frac{M_Y(t+\tau_n)}{M_Y(\tau_n)} \tag{6.40}$$

The first two moments of $\tilde{Y}_{n,i}$ satisfy

$$E\left[\tilde{Y}_{n,i}\right] = M'_{\tilde{Y}_n}(0) = \frac{M'_Y(\tau_n)}{M_Y(\tau_n)} = \Lambda'_Y(\tau_n) = c_n, \tag{6.41}$$

$$\text{var}\left(\tilde{Y}_{n,i}\right) = E\left[\left(\tilde{Y}_{n,i}\right)^2\right] - \left(E\left[\tilde{Y}_{n,i}\right]\right)^2 = M''_{\tilde{Y}_n}(0) - M'_{\tilde{Y}_n}(0)^2 \tag{6.42}$$

$$= \Lambda''_Y(\tau_n) \in (0, \infty). \tag{6.43}$$

Denote $\tilde{S}_n = \sum_{i=1}^n \tilde{Y}_{n,i}$. Since $\tilde{S}_n$ is the sum of n i.i.d. random variables, it has moment generating function

$$M_{\tilde{S}_n}(t) = \left(\frac{M_Y(t+\tau_n)}{M_Y(\tau_n)}\right)^n = \frac{E\left[e^{(t+\tau_n)\tilde{S}_n}\right]}{M_Y(\tau_n)^n} = \frac{1}{M_Y(\tau_n)^n} \int_{-\infty}^{\infty} e^{(t+\tau_n)u} dF_{S_n}(u) \tag{6.44}$$

where F_{S_n} is the c.d.f. of $S_n = \sum_{i=1}^n Y_i$. Therefore, the cumulative distribution function of $\tilde{S}_n$, denoted as $F_{\tilde{S}_n}$, satisfies

$$dF_{\tilde{S}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)^n} dF_{S_n}(y). \tag{6.45}$$

Let $d > 0$. We have

$$\Pr\left(\frac{1}{n}\sum_{i=1}^n X_i \le a_n\right) = \Pr\left(\sum_{i=1}^n Y_i \ge nc_n\right) = \int_{nc_n}^{\infty} dF_{S_n}(u) = \int_{nc_n}^{\infty} M_Y(\tau_n)^n e^{-\tau_n u} dF_{\tilde{S}_n}(u)$$

$$\ge M_Y(\tau_n)^n \int_{nc_n}^{n(c_n+d)} e^{-\tau_n u} dF_{\tilde{S}_n}(u) \ge M_Y(\tau_n)^n e^{-n(c_n+d)\tau_n} \int_{nc_n}^{n(c_n+d)} dF_{\tilde{S}_n}(u)$$

$$= e^{-n(\tau_n(c_n+d)-\Lambda_Y(\tau_n))} \Pr\left(nc_n < \tilde{S}_n < n(c_n+d)\right)$$

$$= e^{-n(\tau_n(c_n+d)-\Lambda_Y(\tau_n))} \Pr\left(c_n < \frac{1}{n}\tilde{S}_n < c_n + d\right)$$

Since $E\left[\tilde{Y}_{n,i}\right] = c_n$ and $\text{var}\left(\tilde{Y}_{n,i}\right) > 0$, we have from the assumption of the theorem that $\Pr\left(\frac{1}{n}\tilde{S}_n > c_n\right)$ is bounded away from zero as $n \to \infty$. We also have

$\Pr\left(\frac{1}{n}\tilde{S}_n < c_n + d\right) \to 1$ as $n \to \infty$, which can be shown using a strong law of large numbers for triangular arrays [69]. Therefore,

$$\log \Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \leq a_n\right) \geq -n\left(\tau_n\left(c_n + d\right) - \Lambda_Y(\tau_n)\right) + \log \Pr\left(c_n < \frac{1}{n}\tilde{S}_n < c_n + d\right) \tag{6.46}$$

$$\sim -n\left(\tau_n\left(c_n + d\right) - \Lambda_Y(\tau_n)\right) \quad \text{as } n \to \infty \tag{6.47}$$

$$\sim -n\left(\tau_n c_n - \Lambda_Y(\tau_n)\right) \quad \text{as } d \to 0 \tag{6.48}$$

$$= -nI_Y^+\left(c_n\right) = -nI_X^-\left(a_n\right) \tag{6.49}$$

□

6.7 Proof of Lemma 3.3.1

Proof. Here we want to show that

$$\Pr\left(\frac{1}{n}\tilde{S}_n > c_n\right) \to 0.5 \tag{6.50}$$

as $n \to \infty$. We note that the L.H.S. of (6.50) involves a sum of random variables $\sum_{i=1}^{n} \tilde{Y}_{n,i}$ that are i.i.d. across i for a given n. We will show that the central limit theorem (CLT) applies in this case by showing that Lindeberg's condition holds. Before we state Lindeberg's condition, we first introduce a change of variable to simplify the problem in the later stage. Denote $\tilde{Y}_n$ the common distribution of $\tilde{Y}_{n,i}$, $\forall i$. Let $\tilde{Z}_n = \tilde{Y}_n - E\left[\tilde{Y}_n\right]$. Hence $E\left[\tilde{Z}_n\right] = 0$ and $\text{var}\left(\tilde{Z}_n\right) = \text{var}\left(\tilde{Y}_n\right)$. Note also that $E\left[\tilde{Y}_n\right] = c_n$ and $\text{var}\left(\tilde{Y}_n\right) = \Lambda_Y''\left(\tau_n\right)$. Lindeberg's condition is hence given as

$$\frac{1}{\sigma^2_{\tilde{Z}_n}} \int_{\left\{|\tilde{Z}_n| > \epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}\right\}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} \to 0 \text{ as } n \to \infty \tag{6.51}$$

for every $\epsilon > 0$. Proving that this condition is true for any general distribution is hard because we do not have the closed-form expression of τ_n. Instead we will here verify Lindeberg's condition for $\sqrt{h_i}$, where $\sqrt{h_i}$ is Rayleigh distributed.

We first give the asymptotic expression of var $\left(\tilde{Y}_n\right)$ as $n \to \infty$ for the Rayleigh distribution. We have the following results:

$$\frac{d\Lambda_Y(\theta)}{d\theta} = \mu_X + \frac{1}{M_{\sqrt{h}}(-\theta)} \frac{dM_{\sqrt{h}}(-\theta)}{d\theta}$$

$$\frac{d^2\Lambda_Y(\theta)}{d\theta^2} = \frac{\frac{d^2M_{\sqrt{h}}(-\theta)}{d\theta^2} M_{\sqrt{h}}(-\theta) - \left(\frac{dM_{\sqrt{h}}(-\theta)}{d\theta}\right)^2}{M_{\sqrt{h}}(-\theta)^2} \tag{6.52}$$

Note that

$$\frac{dM_{\sqrt{h}}(-\theta)}{d\theta} = \frac{\left(\kappa^2\theta^2 + 1\right) M_{\sqrt{h}}(-\theta) - 1}{\theta} \tag{6.53}$$

$$\frac{d^2M_{\sqrt{h}}(-\theta)}{d\theta^2} = \kappa^2 \left[\left(\kappa^2\theta^2 + 3\right) M_{\sqrt{h}}(-\theta) - 1\right] \tag{6.54}$$

Substituting (6.53) and (6.54) into (6.52) gives

$$\frac{d^2\Lambda_Y(\theta)}{d\theta^2} = \frac{\left(\kappa^2\theta^2 - 1\right) M_{\sqrt{h}}(-\theta)^2 + \left(\kappa^2\theta^2 + 2\right) M_{\sqrt{h}}(-\theta) - 1}{\theta^2 M_{\sqrt{h}}(-\theta)} \tag{6.55}$$

Using the asymptotic expansion of $M_{\sqrt{h}}(-\theta)$ (since $\theta \to \infty$ as $n \to \infty$)

$$M_{\sqrt{h}}(-\theta) = \frac{1}{\kappa^2\theta^2} - \frac{3}{\left(\kappa^2\theta^2\right)^2} + \frac{15}{\left(\kappa^2\theta^2\right)^3} - \cdots$$

we obtain $\frac{d^2\Lambda_Y(\theta)}{d\theta^2} \sim \frac{\frac{2}{\kappa^4\theta^4}}{\theta^2 \frac{1}{\kappa^4\theta^4}} = \frac{2}{\theta^2}$ and hence

$$\text{var}\left(\tilde{Y}_n\right) = \Lambda_Y''\left(\tau_n\right) \sim \frac{a^2}{2n} \tag{6.56}$$

The expression of $f_{\tilde{Z}_n}(\tilde{z})$ can be easily found and is given as

$$f_{\tilde{Z}_n}(\tilde{z}) = \frac{a_n - \tilde{z}}{\kappa^2 M_Y(\tau_n)} e^{-\frac{(a_n - \tilde{z})^2}{2\kappa^2} + \tau_n(\tilde{z} + c_n)} \tag{6.57}$$

Note that $\tilde{Z}_n \in (-\infty, a_n]$.

We are now ready to look at Lindeberg's condition (6.51) after obtaining the

expressions (6.56) and (6.57). We have

$$\frac{1}{\sigma^2_{\tilde{Z}_n}}\int_{\left\{|\tilde{Z}_n|>\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}\right\}}\tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z})d\tilde{z} = \frac{1}{\sigma^2_{\tilde{Z}_n}}\left(\int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}\tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z})d\tilde{z} + \int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{a_n}\tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z})d\tilde{z}\right) \tag{6.58}$$

$$\overset{(a)}{\sim}\frac{1}{\sigma^2_{\tilde{Z}_n}}\int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}\tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z})d\tilde{z} = \frac{1}{\kappa^2\sigma^2_{\tilde{Z}_n}M_Y(\tau_n)}\int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}(a_n-\tilde{z})\,\tilde{z}^2 e^{-\frac{(a_n-\tilde{z})^2}{2\kappa^2}+\tau_n(\tilde{z}+c_n)}d\tilde{z} \tag{6.59}$$

$$=\frac{e^{-\mu_{\sqrt{h}}\tau_n}}{\kappa^2\sigma^2_{\tilde{Z}_n}M_{\sqrt{h}}(-\tau_n)}\int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{\infty}(a_n+u)\,u^2 e^{-\frac{(a_n+u)^2}{2\kappa^2}+\tau_n(c_n-u)}du \tag{6.60}$$

$$=\frac{1}{\kappa^2\sigma^2_{\tilde{Z}_n}M_{\sqrt{h}}(-\tau_n)}\int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{\infty}(a_n+u)\,u^2 e^{-\frac{(a_n+u)^2}{2\kappa^2}-\tau_n u+\tau_n c_n-\tau_n\mu_{\sqrt{h}}}du \tag{6.61}$$

where $\mu_{\sqrt{h}} = E\left[\sqrt{h}\right]$, $u=-\tilde{z}$ and step (a) is due to the second integral vanishing to zero as the integration interval becomes null, since $a_n \to 0$ and $\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}} \to a\epsilon/\sqrt{2}$ as $n\to\infty$. Also note that we have the following asymptotic expressions (as $n\to\infty$)

$$a_n = a/\sqrt{n} \to 0 \tag{6.62}$$

$$c_n = \mu_{\sqrt{h}} - a_n \to \mu_{\sqrt{h}} \tag{6.63}$$

$$\tau_n \sim \frac{2\sqrt{n}}{a} \quad \text{(from (3.32) and } \tau=-\theta) \tag{6.64}$$

$$M_{\sqrt{h}}(-\tau_n) \sim \frac{a^2}{4\kappa^2 n} \tag{6.65}$$

$$\sigma^2_{\tilde{Z}_n} \sim \frac{a^2}{2n}. \tag{6.66}$$

We now show that (6.61) goes to zero as $n\to\infty$ by using an upper bound of (6.61) and show that the upper bound goes to zero as $n\to\infty$. We can obtain the following upper bounds by inspecting the exponential terms in (6.61): 1) $e^{-\frac{(a_n+u)^2}{2\kappa^2}} \le e^{-\frac{u^2}{2\kappa^2}}$ (since $a_n>0$), 2) $e^{\tau_n c_n-\tau_n\mu_{\sqrt{h}}} = e^{\tau_n\mu_{\sqrt{h}}-\tau_n a_n-\tau_n\mu_{\sqrt{h}}} = e^{-\tau_n a_n} = O(1)$ (from (6.62) and (6.64)) $\Rightarrow e^{\tau_n c_n-\tau_n\mu_{\sqrt{h}}} \le C$ (for sufficiently large n), 3) $e^{-\tau_n u} = e^{-\frac{2\sqrt{n}}{a}u(1+o(1))} \le e^{-\frac{\sqrt{n}}{a}u}$ (for sufficiently large n), where C is a constant.

Hence we substitute the upper bounds obtained above and the asymptotic ex-

pressions (6.62), (6.65) and (6.66) into (6.61) and obtain the following upper bound

$$\frac{1}{\kappa^2\sigma^2_{\tilde{Z}_n} M_{\sqrt{h}}(-\tau_n)} \int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{\infty} (a_n + u)\, u^2 e^{-\frac{(a_n+u)^2}{2\kappa^2} - \tau_n u + \tau_n c_n - \tau_n \mu_{\sqrt{h}}} du \tag{6.67}$$

$$\leq \frac{8Cn^2}{a^4} \int_{a\epsilon/\sqrt{2}}^{\infty} u^3 e^{-\frac{u^2}{2\kappa^2}} e^{-\frac{\sqrt{n}}{a}u} du\,(1 + o(1)) \tag{6.68}$$

We may use Laplace's method [70] to obtain an asymptotic approximation of

$$\mathcal{I}(\sqrt{n}) = \int_{a\epsilon/\sqrt{2}}^{\infty} u^3 e^{-\frac{u^2}{\kappa^2}} e^{-\sqrt{n}\frac{u}{a}} du \tag{6.69}$$

in (6.68). Let $h(u) = u/a$ and $\varphi(u) = u^3 e^{-\frac{u^2}{2\kappa^2}}$. Hence the integral becomes

$$\mathcal{I}(\sqrt{n}) = \int_{A}^{\infty} \varphi(u) e^{-\sqrt{n}h(u)} du \tag{6.70}$$

where $A = a\epsilon/\sqrt{2}$. It is straight forward to see that $h(u)$ and $\varphi(u)$ satisfy the four conditions necessary for using Laplace's method. The Taylor series for $h(u)$ and $\varphi(u)$ as $u \to A$ are given as $h(u) \sim h(A) + \sum_{s=0}^{\infty} a_s (u-A)^{s+\mu}$, $\varphi(u) \sim \sum_{s=0}^{\infty} b_s (u-A)^{s+\alpha-1}$. We give the values of the first few terms of the series: $h(A) = \epsilon/\sqrt{2}$, $a_0 = 1/a$, $a_1 = 0$, $\mu = 1$, $\alpha = 1$, $b_0 = \left(\frac{a\epsilon}{\sqrt{2}}\right)^3 \exp\left(-\frac{(a\epsilon)^2}{4\kappa^2}\right)$. The asymptotic approximation of $\mathcal{I}$ is given as

$$\mathcal{I}(\sqrt{n}) \sim e^{-\sqrt{n}h(A)} \sum_{s=0}^{\infty} \Gamma\left(\frac{s+\alpha}{\mu}\right) \frac{c_s}{\sqrt{n}^{(s+\alpha)/\mu}} \sim \frac{c_0}{\sqrt{n}} e^{-\epsilon\sqrt{2n}} \tag{6.71}$$

where we have simply retained the first term in the sum. Note that $c_0 = a\left(\frac{a\epsilon}{\sqrt{2}}\right)^3 \exp\left(-\frac{(a\epsilon)^2}{4\kappa^2}\right)$. Hence Lindeberg's condition becomes

$$\frac{8c_0 n\sqrt{n}}{a^4} e^{-\epsilon\sqrt{2n}} \to 0 \quad \text{as } n \to \infty. \tag{6.72}$$

This completes the proof for showing that the CLT holds for $\sqrt{h_i}$. □

6.8 Proof of Lemma 3.3.3

Proof. Let $g(t) = \int_1^\infty e^{\frac{t}{x}-cx}dx$ and $h(t) = \int_1^\infty \frac{1}{x}e^{\frac{t}{x}-cx}dx$. We use Laplace's method [70] to obtain asymptotic approximations of $g(t)$ and $h(t)$. We begin by writing $g(t)$ and $h(t)$ as $g(t) = \int_1^\infty e^{-tp(x)}q(x)dx$ and $h(t) = \int_1^\infty e^{-tp(x)}\phi(x)dx$ where $p(x) = -1/x$, $q(x) = e^{-cx}$ and $\phi(x) = \frac{e^{-cx}}{x}$. In order to apply Laplace's method, we must check four conditions (Theorem 1 in Ch 2 of [70]). The first condition is that $p(x) > p(1)$ for all $x \in (1, \infty)$, and for every $\delta > 0$ the infimum of $p(x) - p(1)$ in $[1 + \delta, \infty)$ is positive. This is true for $p(x) = -1/x$. The second condition is that $p'(x)$ and $q(x)$ and $\phi(x)$ are continuous in a neighborhood of $x = 1$, except possibly at $x = 1$. This is again true for the $p'(x)$, $q(x)$ and $\phi(x)$ defined here. The third condition says that the asymptotic Taylor series of $p(x)$, $q(x)$ and $\phi(x)$ can be obtained as $x \to 1$ from the right. This can be easily verified and we will explicitly give these expressions in what follows. The last condition is that the integral converges absolutely for all sufficiently large t. This can be shown easily for $g(t)$ and $h(t)$. We will now directly apply Laplace's method. The Taylor series for $p(x)$, $q(x)$ and $\phi(x)$ as $x \to 1$ are given as $p(x) \sim p(1) + \sum_{s=0}^\infty a_s(x-1)^{s+\mu}$, $q(x) \sim \sum_{s=0}^\infty b_s(x-1)^{s+\alpha-1}$ and $\phi(x) \sim \sum_{s=0}^\infty k_s(x-1)^{s+\beta-1}$. We give the values of the first few terms of the series: $p(1) = -1$, $a_0 = 1$, $a_1 = -1$, $\mu = 1$, $b_0 = e^{-c}$, $b_1 = -ce^{-c}$, $\alpha = 1$, $k_0 = e^{-c}$, $k_1 = -(c+1)e^{-c}$ and $\beta = 1$. The asymptotic approximation of $g(t)$ is given as $g(t) \sim e^{-tp(1)} \sum_{s=0}^\infty \Gamma\left(\frac{s+\alpha}{\mu}\right) \frac{c_s}{t^{(s+\alpha)/\mu}} \sim e^t \left(\frac{e^{-c}}{t} + \frac{(2-c)e^{-c}}{t^2}\right)$ where we have simply retained the first two terms in the sum. Note that $c_0 = \frac{b_0}{\mu a_0^{\alpha/\mu}}$ and $c_1 = \left\{\frac{b_1}{\mu} - \frac{(\alpha+1)a_1b_0}{\mu^2 a_0}\right\} \frac{1}{a_0^{(\alpha+1)/\mu}}$ [70]. In the same way we obtain the asymptotic approximation of $h(t)$ given as $h(t) \sim e^t \left(\frac{e^{-c}}{t} + \frac{(1-c)e^{-c}}{t^2}\right)$.

Substituting the asymptotic approximations of $g(t)$ and $h(t)$ back into (3.45) gives

$$g_N = 1 - \frac{h(t)}{g(t)} \sim 1 - \frac{e^t \left(\frac{e^{-c}}{t} + \frac{(1-c)e^{-c}}{t^2}\right)}{e^t \left(\frac{e^{-c}}{t} + \frac{(2-c)e^{-c}}{t^2}\right)} = \frac{1}{t+2-c} \sim \frac{1}{t} \quad \text{for large } t \qquad (6.73)$$

which completes the proof. □

6.9 Proof of Lemma 4.3.2

Proof. It can be easily show that at optimality the average distortion constraint given in problem (4.10) must be met with equality, given as

$$\frac{1}{N}\text{tr}\left(\mathbf{R}^* + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1} = \bar{D}_{max} \tag{6.74}$$

where $\mathbf{R}^* = \text{diag}(r_1^*, \ldots, r_N^*)$. Differentiating (6.74) both sides w.r.t. h_n gives

$$\begin{aligned}
&\frac{1}{N}\frac{\partial}{\partial h_n}\text{tr}\left(\mathbf{R}^* + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1} = 0 \\
\Rightarrow &\text{tr}\left(\frac{\partial}{\partial h_n}\left(\mathbf{R}^* + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1}\right) = 0 \\
\Rightarrow &-\text{tr}\left(\frac{\partial \mathbf{R}^*}{\partial h_n}\left(\mathbf{R}^* + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-2}\right) = 0 \\
\Rightarrow &\sum_{i=1}^{N}\frac{\partial r_i^*}{\partial h_n}\cdot \boldsymbol{e}_i^T\left(\mathbf{R}^* + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-2}\boldsymbol{e}_i = 0
\end{aligned} \tag{6.75}$$

where $\frac{\partial \mathbf{R}^*}{\partial h_n} = \text{diag}\left(\frac{\partial r_1^*}{\partial h_n}, \ldots, \frac{\partial r_N^*}{\partial h_n}\right)$.

Let $k_i = \boldsymbol{e}_i^T\left(\mathbf{R}^* + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-2}\boldsymbol{e}_i$, the last line of (6.75) becomes

$$\sum_{i=1}^{N}\frac{\partial r_i^*}{\partial h_n}k_i = 0 \tag{6.76}$$

From the KKT conditions given in (4.11), we can obtain the following expression (by eliminating γ_i) given as

$$r_i^* = \frac{1}{\sigma_i^2}\left(1 - \sqrt{\frac{N\alpha_i^2}{h_i\lambda_0^*\left(\boldsymbol{e}_i^T\left(\mathbf{R}^* + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-2}\boldsymbol{e}_i\right)}}\right)^+ \tag{6.77}$$

for $i = 1, \ldots, N$. Assuming that $r_i^* \geq 0$, we can obtain

$$k_i = \frac{N\left(P_i^2 h_i \sigma_i^2\right)^2}{h_i\lambda_0^*\alpha_i^2} \tag{6.78}$$

Differentiating (6.78) w.r.t. h_n yields

$$\frac{\partial r_i^*}{\partial h_n} = \begin{cases} \frac{\alpha_n^2 P_n^*}{(P_n^* h_n \sigma_n^2 + \alpha_n^2)^2} + \frac{h_n^2 \alpha_n^2}{(P_n^* h_n \sigma_n^2 + \alpha_n^2)^2} \frac{\partial P_n^*}{\partial h_n}, & \text{if } i = n \\ \frac{\alpha_i^2 h_i}{\left(P_i^2 h_i \sigma_i^2 + \alpha_i^2\right)^2} \frac{\partial P_i^*}{\partial h_n}, & \text{if } i \neq n \end{cases} \tag{6.79}$$

Substituting (6.78) and (6.79) into (6.76) gives

$$\begin{aligned} & \sum_{i=1}^{N} \frac{\partial r_i^*}{\partial h_n} k_i = 0 \\ \Rightarrow & \sum_{\substack{i=1 \\ i \neq n}}^{N} \frac{\partial r_i^*}{\partial h_n} k_i + \frac{\partial r_n^*}{\partial h_n} k_n = 0 \\ \Rightarrow & \frac{N}{\lambda_0^*} \sum_{\substack{i=1 \\ i \neq n}}^{N} \frac{\partial P_i^*}{\partial h_n} + \frac{N}{\lambda_0^*} \frac{P_n^*}{h_n} + \frac{N}{\lambda_0^*} \frac{\partial P_n^*}{\partial h_n} = 0 \\ \Rightarrow & \frac{\partial \left\langle \mathbf{P}^*(\mathbf{h}) \right\rangle}{\partial h_n} = -\frac{1}{N} \frac{P_n^*}{h_n} \end{aligned} \tag{6.80}$$

which completes the proof. □

Bibliography

[1] (2010, Aug.) 'Smart' water quality sensor network wins iAward. CSRIO Media Release. [Online]. Available: http://www.csiro.au/news/smart-water-quality-sensor-network-wins-iAward.html

[2] A. Ephremides, "Energy concerns in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 45–59, Aug. 2002.

[3] R. Hafez, I. Haroun, and I. Lambadaris. (2005, Sep.) Building wireless sensor networks. Microwaves & RF. [Online]. Available: http://www.mwrf.com/Article/ArticleID/11071/11071.html

[4] (2004, Dec.) Sun microsystems researchers unveil world's smallest secure web server, win best paper award at PerCom 2005. Oracle Corporation. [Online]. Available: http://labs.oracle.com/spotlight/2004-12-20_vgupta.html

[5] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.

[6] C. Verikoukis, L. Alonso, and T. Giamalis, "Cross-layer optimization for wireless systems: a european research key challenge," in *Global Communications Newsletter*, Jul. 2005, pp. 1–3.

[7] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun. Mag.*, vol. 9, no. 4, pp. 8–27, Aug. 2002.

[8] E. Uysal-Biyikoglu and A. El Gamal, "On adaptive transmission for energy efficiency in wireless data networks," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3081–3094, Dec. 2004.

[9] P. Nuggehalli, V. Srinivasan, and R. R. Rao, "Energy dfficient transmission scheduling for delay constrained wireless networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 531–539, Mar. 2006.

[10] H. Ren, M. Meng, and X. Chen, "Cross-layer optimization schemes for wireless biosensor networks," in *Proc. of the 6th World Congress on Intelligent Control and Automation*, Dalian, China, Jun. 2006, pp. 369–374.

[11] R. Madan, S. Cui, S. Lall, and A. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3142–3152, Nov. 2006.

[12] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1135–1149, May 2002.

[13] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[14] J. Luo, R. S. Blum, L. J. Cimini, L. J. Greenstein, and A. M. Haimovich, "Decode-and-forward cooperative diversity with power allocation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 793–799, Mar. 2007.

[15] A. R. Nigara, M. Qin, and R. S. Blum, "On the performance of wireless ad hoc networks using amplify-and-forward cooperative diversity," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3204–3214, Nov. 2006.

[16] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[17] G. Li and H. Liu, "Resource allocation for OFDMA relay networks with fairness constraints," *IEEE J. Select. Areas Commun.*, vol. 24, no. 11, pp. 2061–2069, Nov. 2006.

[18] A. D. Coso, U. Spagnolini, and C. Ibars, "Cooperative distributed mimo channels in wireless sensor networks," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 402–414, Feb. 2007.

[19] S. Cui and A. J. Goldsmith, "Cross-layer optimization of sensor networks based on cooperative MIMO techniques with rate adaptation," *2005 IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, pp. 960–964, Jun. 2005.

[20] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," *Lecture Notes in Computer Science*, vol. 2634, pp. 162–177, Apr. 2003.

[21] J. J. Xiao, Z. Q. Luo, S. Cui, and A. J. Goldsmith, "Power-efficient analog forwarding transmission in an inhomogeneous gaussian sensor network," *2005 IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, pp. 121–125, Jun. 2005.

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

[23] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Processing*, vol. 56, no. 2, pp. 757–770, Feb. 2008.

[24] A. Krasnopeev, J. J. Xiao, and Z. Q. Luo, "Minimum energy decentralized estimation in wireless sensor network with correlated sensor noises," *EURASIP Journal on Wireless Communications and Networking*, no. 4, pp. 473–482, Sep. 2005.

[25] S. Dey and J. Evans, "Optimal power control in wireless data networks with outage-based utility guarantees," in *Proc. of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, USA, Dec. 2003, pp. 570–575.

[26] ——, "Outage capacity and optimal power allocation for multiple time-scale paralle fading channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2369–2373, Jul. 2007.

[27] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.

[28] N. Ahmed, M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Outage minimization with limited feedback for the fading relay channel," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 659–669, Apr. 2006.

[29] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[30] L. Lin, R. D. Yates, and P. Spasojevic, "Adaptive transmission with discrete code rates and power levels," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2115–2125, Dec. 2003.

[31] A. Khoshnevis and A. Sabharwal, "Performance of quantized power control in multiple antenna systems," in *Proc. IEEE International Conference on Communications (ICC'04)*, Jun. 2004, pp. 803–807 Vol.2.

[32] M. Banavar, C. Tepedelenlioglu, and A. Spanias, "Estimation over fading channels with limited feedback using distributed sensing," *IEEE Trans. Signal Processing*, vol. 58, no. 1, pp. 414–425, Jan. 2010.

[33] T. T. Kim, M. Skoglund, and G. Caire, "On cooperative source transmission with partial rate and power control," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1408–1418, Oct. 2008.

[34] R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1754–1763, May 2007.

[35] T. T. Kim and M. Skoglund, "Diversity-multiplexing tradeoff in MIMO channels with partial CSIT," *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 2743–2759, Aug. 2007.

[36] M. A. Khojastepour, G. Yue, X. Wang, and M. Madihian, "Optimal power control in MIMO systems with quantized feedback," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4859–4866, Dec. 2008.

[37] T. T. Kim and M. Skoglund, "Partial power control for slowly fading MIMO channels," in *Proc. IEEE International Conference on Communications (ICC'06)*, vol. 3, Jun. 2006, pp. 1362–1367.

[38] J. C. Spall, "Implementation of the simultaneous perturbation algorithm for stochastic optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 3, pp. 817–823, Jul. 1998.

[39] S. Nadarajah, "Simplified expressions for the outage and error rate performance of DS-CDMA with MRC in Nakagami-m fading," *International Journal of Electronics*, vol. 96, no. 4, pp. 365–369, Apr. 2009.

[40] P. G. Moschopoulos, "The distribution of the sum of independent gamma random variables," *Annals of the Institute of Statistical Mathematics*, vol. 37, no. 1, pp. 541–544, 1985.

[41] J. Reig and N. Cardona, "Approximation of outage probability on Nakagami fading channels with multiple interferers," *Electron. Lett.*, vol. 36, no. 19, pp. 1649–1650, Sep. 2000.

[42] M. Pausini and G. J. M. Janssen, "Performance analysis of UWB autocorrelation receivers over Nakagami-fading channels," *IEEE J. Select. Topics Signal Processing*, vol. 1, no. 3, pp. 443–455, Oct. 2007.

[43] V. A. Aalo, T. Piboongungon, and G. P. Efthymoglou, "Another look at the performance of MRC schemes in Nakagami-m fading channels with arbitrary parameters," *IEEE Trans. Commun.*, vol. 53, no. 12, pp. 2002–2005, Dec. 2005.

[44] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Massachusetts, USA: Kluwer Academic Publishers, 1992.

[45] J. Gentle, W. Härdle, and Y. Mori, (Editors), *Handbook of Computational Statistics.* Germany: Springer, 2004.

[46] J. Pan, L. Cai, Y. T. Hou, Y. Shi, and S. X. Shen, "Optimal base-station locations in two-tiered wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 5, pp. 458–473, Sep. 2005.

[47] O. Younis and S. Fahmy, "HEED: A hybrid, enenrgy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, Oct. 2004.

[48] V. P. Mahtre, C. Rosenberg, D. Kofman, R. Mazumdar, and N. Shroff, "A minimum cost heterogeneous sensor network with a lifetime constraint," *IEEE Transactions on Mobile Computing*, vol. 4, no. 1, pp. 4–15, Jan. 2005.

[49] Y. Ma and J. H. Aylor, "System lifetime optimization for heterogeneous sensor networks with a hub-spoke topology," *IEEE Transactions on Mobile Computing*, vol. 3, no. 3, pp. 286–294, Jul. 2004.

[50] T. Himsoon, W. P. Siriwongpairat, Z. Han, and K. J. R. Liu, "Lifetime maximization via cooperative nodes and relay deployment in wireless networks," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 306–317, Feb. 2007.

[51] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian "sensor" network," *IEEE Trans. Inform. Theory*, vol. 54, no. 11, pp. 5427–5251, Nov. 2008.

[52] I. Bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor networks with otpimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1146–1156, Jul. 2008.

[53] C. Wang and S. Dey, "Power allocation for distortion outage minimization in clustered wireless sensor networks," in *Proc. IEEE International Wireless Com-*

munications and Mobile Computing Conference (IWCMC'08), Crete, Greece, Aug. 2008, pp. 395–400.

[54] H. Senol and C. Tepedelenlioglu, "Performance of distributed estimation over unknown parallel fading channels," *IEEE Trans. Signal Processing*, vol. 56, no. 12, pp. 6057–6068, Dec. 2008.

[55] K. Bai, H. Senol, and C. Tepedelenlioglu, "Outage scaling laws and diversity for distributed estimation over parallel fading channels," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 3182–3192, Aug. 2009.

[56] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* New Jersey: Prentice Hall, 1993.

[57] K. Liu, H. El Gamal, , and A. M. Sayeed, "Decentralized inference over multiple-access channels," *IEEE Trans. Signal Processing*, vol. 5, no. 7, pp. 3445–3455, Jul. 2007.

[58] F. W. J. Olver, *Asymptotics and Special Functions.* New York: Academic Press, 1974.

[59] G. Casella and R. L. Berger, *Statistical Inference.* USA: Duxbury, 2002.

[60] M. C. Vuran and I. F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 2, pp. 316–329, Apr. 2006.

[61] J. O. Berger, V. D. Oliveira, and B. Sanso, "Objective bayesian analysis of spatially correlated data," *Journal of the American Statistical Association*, vol. 96, no. 456, pp. 1361–1374, Dec. 2001.

[62] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables andEntropy.* USA: American Mathematical Society, 2000.

[63] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices.* New York: Cambridge University Press, 2010.

[64] M. J. M. Peacock, I. B. Collings, and M. L. Honig, "Eigenvalue distributions of sums and products of large random matrices via incremental matrix expansions," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 2123–2138, May 2008.

[65] A. Bottcher and S. M. Grudsky, *Spectral Properties of Banded Toeplitz Matrices*. Philadelphia: SIAM, 2005.

[66] G. Meurant, "A review on the inverse of symmetric tridiagonal and block tridiagonal matrices," *SIAM J. Matrix Anal. Appl.*, vol. 13, no. 3, pp. 707–728, Jul. 1992.

[67] H. A. Yamani and M. S. Abdelmonem, "The analytic inversion of any finite symmetric tridiagonal matrix," *J. Phys. A: Math. Gen.*, vol. 30, no. 8, pp. 2889–2893, Apr. 1997.

[68] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. New York, USA: Oxford University Press, 2001.

[69] T. Hu, F. Móricz, and R. L. Taylor, "Strong laws of large numbers for arrays of rowwise independent random variables," *Acta Math. Hung.*, vol. 54, no. 1-2, pp. 153–162, 1989.

[70] R. Wong, *Asymptotic Approximations of Integrals*. USA: SIAM, 2001.