Distortion Outage Minimization and Diversity Order Analysis for Coherent Multi-Access

Chih-Hong Wang, *Student Member, IEEE,* Alex S. Leong, *Member, IEEE,* and Subhrakanti Dey, *Senior Member, IEEE*

Abstract

In this paper we investigate the distortion outage performance of distributed estimation schemes in wireless sensor networks, where a distortion outage is defined as the event that the estimation error or distortion exceeds a pre-determined threshold. The sensors transmit their observation signals using the analog amplify and forward scheme through coherent multi-access channels to the fusion center, which reconstructs a minimum mean squared error (MMSE) estimate of the physical quantity observed. We consider three power allocation schemes - 1) equal power allocation (EPA), 2) short-term optimal power allocation (ST-OPA), where we minimize the distortion subject to a power constraint at each time step, and 3) long-term optimal power allocation (LT-OPA), where we minimize the distortion outage probability subject to a long-term average power constraint. We study their diversity orders of distortion outage in terms of increasing numbers of sensors, and show that under Rayleigh fading EPA and ST-OPA achieve the same diversity order of $N \log N$, where N is the number of sensors. This suggests that in the case of a large number of sensors, the spatial diversity gain in EPA can overcome fading equally well as in ST-OPA. On the other hand, in LT-OPA, we find that for $N > 1$ the outage probability can be driven to zero with a finite amount of total power.

Index Terms

Diversity order, multiple access channel, outage probability, power control, sensor networks

I. Introduction

Wireless sensor networks (WSNs) 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 [1]. The sensors may be configured into various ad-hoc network structures depending on the protocol

The authors are with the Department of Electrical & Electronic Engineering, University of Melbourne, Parkville, Victoria 3010 (e-mail: chwang@ee.unimelb.edu.au, [asleong,sdey]@unimelb.edu.au).

and the application being considered [2]. Examples of these such as forming clusters and electing cluster heads [3], cooperative transmission and cooperative diversity (relay nodes used to forward signals) [4]–[8] and multiple sensor transmission to achieve distributed beam-forming as in MIMO systems [1] 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 [9]–[13], power allocation schemes and cross-layer optimization [1], [4], [14] 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 [15] showed that in a Gaussian sensor network (a sensor network estimating a Gaussian source) it is asymptotically optimal to transmit using uncoded analog forwarding of measurements by multiple sensors as opposed to separate source channel coding. Later in [16] 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 [17] the authors obtained the optimal power allocation of an inhomogeneous Gaussian wireless sensor network using analog amplify-and-forward through coherent MAC (multiple access channel) 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, [18] solved the problem of minimizing power under a distortion constraint and minimizing distortion under a power constraint. The study of power allocation in distributed estimation for a vector source is given in [19] for coherent MAC and [20] for orthogonal MAC, which also studied power allocation with correlation in sensor data. Power allocation considering correlated sensor noise is studied in [21]. 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 [18]. 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 [22] obtained the optimal power allocation that minimizes the distortion outage probability subject to a long-term average power constraint in a clustered WSN using the amplify-and-forward orthogonal multi-access protocol.

The estimation diversity achieved by wireless sensor networks was first studied in [18] for equal power allocation in orthogonal multi-access channels with Rayleigh fading. They showed that a sensor network with independent and identically distributed (i.i.d.) fading channels and i.i.d. sensor noise variances can achieve an estimation diversity on the order of the number of sensors in the network. In [23] 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 [24] for a large class of fading distributions. With a fixed power per sensor, deterministic and equal sensor signal-to-noise ratios (SNR) and i.i.d. channel SNR, the authors in [24] 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 this paper 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 (FC) 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 [15]. 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 $N > 1$, which intuitively can be regarded as achieving a "diversity order of infinity". 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 paper is organized as follows. In Section II we give the network model. We define and state the three different power allocations in Section III, based on which we perform theoretical analysis to find their diversity orders of distortion outage in Section IV. Simulation results are given in Section V, followed by concluding remarks in Section VI.

In this paper, 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. (probability density function) and c.d.f. (cumulative distribution function) are denoted as $f_X(x)$ and $F_X(x)$ respectively, while $E[X]$ denotes its expectation.

II. Network model

A schematic diagram of the wireless sensor network model is shown in Fig. 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. (independent and identically distributed) in time (k denotes the discrete time index). The measurements of the ith sensor at time k are given as

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

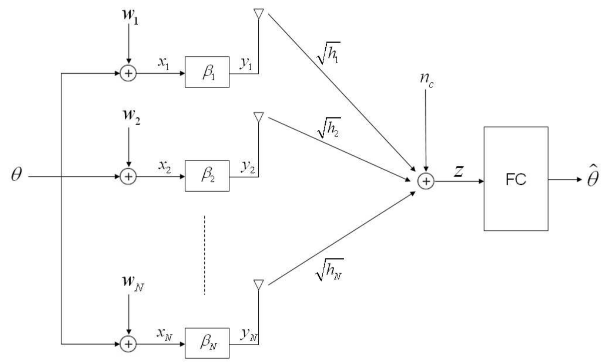


Fig. 1. Schematic diagram of the wireless sensor network using coherent MAC scheme.

where the sensor measurement noise w_i is i.i.d. Gaussian with zero mean and variance σ_i^2.[1] The sensors amplify and forward their signals to the fusion center (FC) via a coherent MAC channel [15] 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. 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 (CSIR) can be easily obtained by the use of pilot tone training from the transmitters, while CSI at the transmitter (CSIT) 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 to 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[2]

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

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

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

[1]In this paper we will mostly be assuming that the σ_i^2 are deterministic quantities (similar to [24]), due to deterministic placement of the sensors by e.g. the network designer. This differs from the model of [18] that considers i.i.d. sensor noise variances. However our diversity order analysis can also be used to treat a class of randomly distributed σ_i^2, see Section IV-D.

[2]The coherent sum (1) requires distributed transmit beamforming [25] that may be difficult to achieve for large sensor networks. This model however is commonly studied, e.g. in works such as [15], [16], [19]. Our goal in this paper 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.

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

It is well known that the optimal estimator for θ is the linear MMSE (minimum mean square error) estimator [26], given as $\hat{\theta} = \frac{E[\theta z]}{E[z^2]} z$. The mean squared error or *distortion* D_k of this estimator, is given as

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

Note that (2) gives the expression of the *instantaneous* distortion, i.e., it is a function of the channel realizations $h_i[k]$, $\forall i, k$. 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. In the next section we will consider three different power allocation schemes.

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 paper.

III. 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 IV.

Remark: In this paper 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. [27]. 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.

A. 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 $P_i = \mathcal{P}_{tot}/N$, $\forall i$.

B. 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:

$$\begin{aligned} \min \quad & D(\mathbf{P}(\mathbf{h}),\mathbf{h}) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{N} P_i(\mathbf{h}) \leq \mathcal{P}_{tot},\ P_i(\mathbf{h}) \geq 0, \quad \forall i. \end{aligned} \tag{3}$$

Problem (3) is related to outage minimization in the following way. Similar to the information outage minimization problem in communications theory [28], we can define a "short-term distortion outage minimization problem" as:

$$\begin{aligned} \min \quad & \Pr\left(D\left(\mathbf{P}(\mathbf{h}),\mathbf{h}\right) > D_{max}\right) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{N} P_i(\mathbf{h}) \leq \mathcal{P}_{tot},\ P_i(\mathbf{h}) \geq 0, \quad \forall i. \end{aligned} \tag{4}$$

where the power constraint holds for each time instant (or channel realization). By similar methods to [28], it can be shown that the solution to problem (3) is a solution to the short-term distortion outage minimization problem (4), that in general has many possible solutions.

Problem (3) has been solved in [19]. The short-term optimal power allocation of the ith sensor is given by

$$P_i^*(\mathbf{h}) = \mathcal{P}_{tot} c_i(h_i) \left(\sum_{j=1}^{N} c_j(h_j) \right)^{-1}, \quad \forall i \tag{5}$$

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 (5) 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 [18].

C. 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 (e.g. at some time instance k_1 the power usage could be greater than the average power, while at another time k_2 the power usage could be less than the average power), see also [28] for the information outage minimization problem in communications theory. 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

$$\begin{aligned} \min \quad & \Pr\left(D\left(\mathbf{P}(\mathbf{h}),\mathbf{h}\right) > D_{max}\right) \\ \text{s.t.} \quad & E\left[\textstyle\sum_{i=1}^{N} P_i(\mathbf{h})\right] \leq \mathcal{P}_{tot},\ P_i(\mathbf{h}) \geq 0, \quad \forall i. \end{aligned} \tag{6}$$

Problem (6) can be solved in a similar way to [28]. First consider the following minimization problem given as

$$\begin{aligned} \min \quad & \langle \mathbf{P}\left(\mathbf{h}\right) \rangle \\ \text{s.t.} \quad & D(\mathbf{P}(\mathbf{h}),\mathbf{h}) \leq D_{max},\ P_i(\mathbf{h}) \geq 0, \quad \forall i. \end{aligned} \tag{7}$$

We have the following lemma:

Lemma 3.1: With the knowledge of $\mathbf{h}$, the solution of problem (7) is given as

$$P_i^*(\mathbf{h}) = P_{tot}(\mathbf{h}) c_i(h_i) \left(\sum_{j=1}^{N} c_j(h_j) \right)^{-1}, \quad i = 1, \ldots, N, \tag{8}$$

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

$$\gamma_{th} = \sum_{i=1}^{N} \frac{h_i}{\left(\frac{\sigma_c^2 C_i}{P_{tot}(\mathbf{h})} + \sigma_i^2 h_i \right)} \tag{9}$$

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

The proof of this lemma can be found in [19] and is hence omitted. One also has the following Lemma which is necessary to find the optimal solution of problem (6):

Lemma 3.2: The long-term optimal power $\mathbf{P}^*(\mathbf{h}) = [P_1^*(\mathbf{h}), \ldots, P_N^*(\mathbf{h})]^T$ as given in (8), 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 (6), we will also need the following definitions and notations, similar to those in [28]. 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 $P_T(t) = \int_{\mathcal{R}_T(t)} \sum_{i=1}^N P_i^*(\mathbf{h}) dF(\mathbf{h})$ and $\overline{P}_T(t) = \int_{\bar{\mathcal{R}}_T(t)} \sum_{i=1}^N P_i^*(\mathbf{h}) dF(\mathbf{h})$, 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 $u^* = \frac{\mathcal{P}_{tot} - P_T(t^*)}{\bar{P}_T(t^*) - P_T(t^*)}$.

With the above lemma and definitions we can now present the solution to problem (6).

Theorem 1: The solution of problem (6) is given as

$$\hat{\mathbf{P}}(\mathbf{h}) = \begin{cases} \mathbf{P}^*(\mathbf{h}), & \text{if } \mathbf{h} \in \mathcal{R}_T(t^*) \\ \mathbf{0}, & \text{if } \mathbf{h} \notin \overline{R}_T(t^*), \end{cases} \tag{10}$$

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 (8).

The proof follows using similar techniques as in [28] 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 (8) 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 (8). 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 is given in Section IV-C which allows one to quickly compute a lower bound of t^* given $\mathcal{P}_{tot}$.

IV. DIVERSITY ORDERS OF DISTORTION OUTAGE

We are interested in seeing how the outage probability decays as the number of sensors N 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 III. Such expressions characterize the *diversity order of distortion outage* introduced in [18], who showed that the outage probability decays exponentially with the number of sensors N for i.i.d. orthogonal MAC. For analytical tractability, in the following theoretical analysis, we will first 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 $C_i = C = \sigma_\theta^2 + \sigma^2$, $\forall i$. These results will then be used in Section IV-D for more general cases of different sensor noise variances and/or fading channels.

Notation: For two functions $f(\cdot)$ and $g(\cdot)$, we will use the standard asymptotic notation (see for example [29], [30]) and say that $f \sim g$ as $t \to t_0$, if $\frac{f(t)}{g(t)} \to 1$ as $t \to t_0$.

Notation: A summary of some important notation can be found in Table I.

TABLE I
TABLE OF NOTATIONS

σ_θ^2	Variance of source θ
σ_i^2	Variance of i^{th} sensor measurement noise w_i
σ^2	Variance of sensor measurement noise in homogeneous case, i.e. $\sigma_i^2 = \sigma^2$, $\forall i$
σ_c^2	Variance of channel noise n_c
C_i	$\sigma_\theta^2 + \sigma_i^2$
C	$\sigma_\theta^2 + \sigma^2$
$c_i\left(h_i\right)$	$C_i h_i / \left(C_i + \mathcal{P}_{tot} h_i \sigma_i^2 / \sigma_c^2\right)^2$
D	Instantaneous distortion
$D_{\max}$	Maximum allowable distortion threshold which if exceeded results in an outage

A. Equal power allocation

Substituting $P_i = \mathcal{P}_{tot}/N$ into (2), after some algebraic manipulation we obtain

$$\frac{D}{\sigma_\theta^2} = \frac{\frac{\sum_{i=1}^N h_i}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}}{\frac{\sum_{i=1}^N h_i}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}} + \frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h_i}}{N}\right)^2}. \tag{11}$$

Inspecting the right hand side (RHS) of (11), 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 law of large numbers as N gets large. However we find that var $\left(\frac{1}{N}\sum_{i=1}^N h_i\right) =$

$\frac{1}{N}\text{var}[h]$ and $\text{var}\left(\frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h_i}}{N}\right)^2\right) \approx \frac{4\sigma_\theta^4 N}{\sigma^4}(E[\sqrt{h}])^2\text{var}[\sqrt{h}]$ (obtained using the Delta method [31]). 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. Let us call

$$\tilde{D} = \frac{\sigma_\theta^2 \eta}{\eta + \frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h_i}}{N}\right)^2}$$

where $\eta = E[h] + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}$. Now by the weak law of large numbers and stochastic order properties on pp.12-13 of [32], we can show the following convergence result:

$$D - \tilde{D} \xrightarrow{p} 0 \tag{12}$$

where $\xrightarrow{p}$ denotes convergence in probability.

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

$$\begin{aligned}
P_{outage} &= \Pr\left(D > D_{max}\right) = \Pr(D - \tilde{D} > D_{max} - \tilde{D}) \\
&\sim \Pr(0 > D_{max} - \tilde{D}) = \Pr(\tilde{D} > D_{max}) \\
&= \Pr\left(\frac{\sigma_\theta^2 \eta}{\eta + \frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h_i}}{N}\right)^2} > D_{max}\right) \\
&= \Pr\left(\frac{1}{N}\sum_{i=1}^N \sqrt{h_i} < \sqrt{\frac{\eta\sigma^2\left(\sigma_\theta^2 - D_{max}\right)}{D_{max}\sigma_\theta^2 N}}\right) \\
&= \Pr\left(\frac{1}{N}\sum_{i=1}^N \sqrt{h_i} < \frac{a}{\sqrt{N}}\right)
\end{aligned} \tag{13}$$

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

To verify the accuracy in our use of the asymptotic approximation (12), in Fig. 2 we plot the expression of $D - \tilde{D}$ on the left hand side of (12), where h_i is exponentially distributed with parameter λ, for a single realization. We can readily see the convergence to zero as N increases. In Fig. 3 (see Section V) we also compare between Monte Carlo simulations of $\log P_{outage}$ and $\log Pr\left(\frac{1}{N}\sum_{i=1}^N \sqrt{h_i} < \frac{a}{\sqrt{N}}\right)$. The results show almost no difference between using the actual outage probability P_{outage} and the asymptotic approximation (13).

By inspecting (13) 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 2 we will provide a generalized version of Cramer's Theorem which can be applied to (13). 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

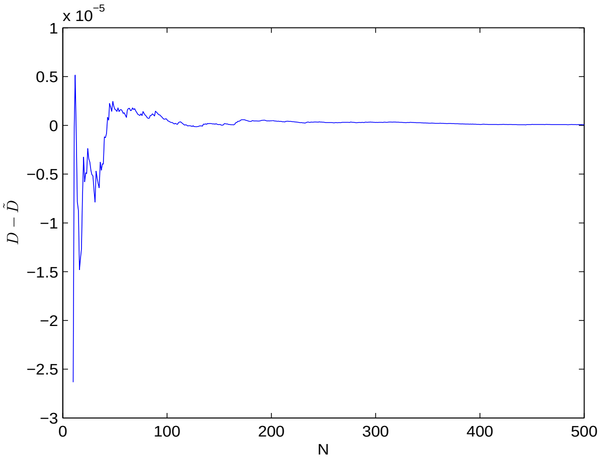


Fig. 2. EPA with total power constraint $\mathcal{P}_{tot} = 10m$W. Plots of $D - \tilde{D}$. Simulation parameters: $\lambda = 250,000$, $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$.

$\Lambda_X(t) \triangleq \log M_X(t)$. 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 2: 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

$$dF_{\tilde{Y}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)} dF_Y(y). \tag{14}$$

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

$$\log \Pr\left(\frac{1}{n}\sum_{i=1}^n X_i \leq a_n\right) \sim -nI_X(a_n) \quad \text{as } n \to \infty. \tag{15}$$

Proof: See Appendix. ■

In order to apply Theorem 2 to (13), 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 4.1: 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 (14). Then

$$\Pr\left(\frac{1}{n}\sum_{i=1}^n \tilde{Y}_{n,i} > E\left[\tilde{Y}_{n,i}\right]\right) \to 0.5 \quad \text{as } n \to \infty. \tag{16}$$

Proof: See Appendix. ■

Applying Theorem 2 to (13) we have

$$\log P_{outage} \sim -NI^{-}_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) \quad \text{as } N \to \infty \tag{17}$$

where

$$I^{-}_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) = \sup_{\theta<0}\left(\frac{a}{\sqrt{N}}\theta - \log M_{\sqrt{h}}(\theta)\right). \tag{18}$$

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

$$M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right) = 1 - \sqrt{\pi}xe^{x^2}\text{erfc}\,(x) \tag{19}$$

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

$$\frac{dI^{-}_{\sqrt{h}}\left(a/\sqrt{N}\right)}{d\theta} = 0, \quad \theta < 0 \tag{20}$$

$$\Rightarrow \frac{\sqrt{N}}{a} = \psi(\theta) \tag{21}$$

where

$$\psi(\theta) = \left(\Lambda'_{\sqrt{h}}(\theta)\right)^{-1} = M_{\sqrt{h}}(\theta)/M'_{\sqrt{h}}(\theta). \tag{22}$$

After substituting $\theta = -\sqrt{2}x/\kappa$ in (21) and some algebraic manipulation, it is possible to obtain

$$\frac{\sqrt{N}}{2} = \psi\left(-\frac{\sqrt{2}x}{\kappa}\right) \tag{23}$$

where

$$\psi\left(-\frac{\sqrt{2}x}{\kappa}\right) = \frac{\sqrt{2}}{\kappa}\frac{xM_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)}{1 - M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right) - 2x^2M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)}. \tag{24}$$

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

$$\frac{d\psi\left(-\frac{\sqrt{2}x}{\kappa}\right)}{dx} = \frac{\sqrt{2}}{\kappa}\frac{\Lambda''_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)}{\left(\Lambda'_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right)\right)^2} \geq 0, \tag{25}$$

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 (22); 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 (23), 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 $\text{erfc}(x) = \frac{e^{-x^2}}{x\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{(2n)!}{n!(2x)^{2n}}$ into the moment generating function (19) and obtain

$$M_{\sqrt{h}}\left(-\frac{\sqrt{2}x}{\kappa}\right) = \frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots. \tag{26}$$

We then substitute (26) into (23) to obtain the following

$$\begin{aligned}\frac{\sqrt{N}}{a} =& \frac{\sqrt{2}}{\kappa} \frac{x\left(\frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots\right)}{1 - \left(\frac{1}{2x^2} - \frac{3}{4x^4} + \cdots\right) - 2x^2\left(\frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots\right)} \\ =& \frac{\sqrt{2}}{\kappa} \frac{\frac{1}{2x} - \frac{3}{4x^3} + \frac{15}{8x^5} + \cdots}{\frac{1}{x^2} - \frac{3}{x^4} + \cdots} \sim \frac{\sqrt{2}}{\kappa}\frac{x}{2} \quad \text{for large } x.\end{aligned}$$

Hence for large N,

$$\theta \sim -\frac{2\sqrt{N}}{a}. \tag{27}$$

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

$$I_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) \sim -\frac{a}{\sqrt{N}}\frac{2\sqrt{N}}{a} - \log M_{\sqrt{h}}\left(-\frac{2}{a_N}\right) \tag{28}$$

$$= -2 - \log M_{\sqrt{h}}\left(-\frac{2}{a_N}\right) \tag{29}$$

$$\sim -2 - \log\left(\frac{a^2}{2\kappa^2 N}\right) \tag{30}$$

$$= -2 - \log\left(\frac{a^2}{2\kappa^2}\right) + \log N. \tag{31}$$

Hence from (17) the outage probability for large N satisfies

$$\log P_{outage} \sim -NI_{\sqrt{h}}\left(\frac{a}{\sqrt{N}}\right) \tag{32}$$

$$\sim -N\left(-2 - \log\left(\frac{a^2}{2\kappa^2}\right) + \log N\right) \tag{33}$$

$$\sim -N\log N, \tag{34}$$

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 [18], 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. 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 [24]. 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. Similar improvements in performance of the coherent MAC over the orthogonal MAC has also been previously observed (for different performance criteria) in e.g. [19], where for a fixed total power the distortion decays to zero at the rate $1/N$ as N increases for coherent MAC, but the distortion is bounded away from zero for orthogonal MAC. This is due to the fact that in coherent combination, the received signal to noise ratio scales with the number of sensors due to the correlation among transmitted messages, even when the total transmit power is finite [19]. However if the total power scales linearly with the number of sensors, then the distortion will decay to zero for orthogonal MAC.

B. Short-term optimal power allocation

We first give the expression of distortion using ST-OPA. Substituting (5) into (2) gives

$$D = \left(\frac{1}{\sigma_\theta^2} + \frac{\left(\sum_{i=1}^N \sqrt{h_i P_i^*} \right)^2}{\sigma^2 \sum_{i=1}^N h_i P_i^* + \sigma_c^2 C} \right)^{-1} = \frac{\sigma_\theta^2 \sigma^2}{\sigma^2 + \sigma_\theta^2 \sum_{i=1}^N Z_i} \tag{35}$$

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

$$P_{outage} = \Pr(D > D_{max}) = \Pr\left(\frac{1}{N} \sum_{i=1}^N Z_i < g_N \right) \tag{36}$$

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

Denote Z as the random variable distributed according to the common distribution of Z_i. We now apply Theorem 2 to (36). We have the following lemma needed for verifying one of the assumptions in Theorem 2 (similar to lemma 4.1).

Lemma 4.2: 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 (14). Then

$$\Pr\left(\frac{1}{n} \sum_{i=1}^n \tilde{Y}_{n,i} > E\left[\tilde{Y}_{n,i} \right] \right) \to 0.5 \quad \text{as } n \to \infty. \tag{37}$$

This lemma can be proved in a similar manner to Lemma 4.1 and is excluded to avoid repetition.

Applying Theorem 2 to (36) we have

$$\log P_{outage} \sim -N I_Z^-(g_N) \quad \text{as } N \to \infty \tag{38}$$

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

$$g_N - \frac{M'_Z(\theta)}{M_Z(\theta)} = 0 \Rightarrow g_N = \frac{\int_0^1 \frac{z}{(1-z)^2} e^{\theta z - \lambda\rho\frac{z}{1-z}} dz}{\int_0^1 \frac{1}{(1-z)^2} e^{\theta z - \lambda\rho\frac{z}{1-z}} dz} \tag{39}$$

$$\Rightarrow g_N = \frac{\int_0^1 z g(z,t) dz}{\int_0^1 g(z,t) dz} \tag{40}$$

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 $\varphi(-t)$ with respect to t yields $\frac{d\varphi(-t)}{dt} = -\Lambda''_Z(-t) \le 0$, 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 (40)). Hence large N corresponds to the case of large t in (40). Let $x = 1/(1-z)$. It can be easily shown that (40) can be written as

$$g_N = 1 - \frac{\int_1^\infty \frac{1}{x} e^{\frac{t}{x} - cx} dx}{\int_1^\infty e^{\frac{t}{x} - cx} dx} \tag{41}$$

where $c = \lambda\rho$.

Lemma 4.3:

$$g_N \sim \frac{1}{t} \text{ as } t \to \infty. \tag{42}$$

Proof: See Appendix. ■

Hence for large N, we have

$$\theta \sim -\frac{1}{g_N}. \tag{43}$$

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

$$\begin{aligned} M_Z(\theta) &= \lambda\rho e^{-t+c} \int_1^\infty e^{-tp(x)} q(x) dx \\ &\sim \lambda\rho e^{-t+c} \frac{e^{t-c}}{t} \sim \lambda\rho g_N. \end{aligned}$$

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

$$I_Z(a_N) \sim -g_N \frac{1}{g_N} - \log\left(\frac{\lambda\rho g}{N}\right) \quad \text{for large } N \tag{44}$$

$$= -1 - \log(\lambda\rho g) + \log N. \tag{45}$$

Hence from (38) the outage probability for large N is asymptotically

$$\log P_{outage} \sim -N I_Z(g_N) \tag{46}$$

$$\sim -N(-1 - \log(\lambda\rho g) + \log N) \tag{47}$$

$$\sim -N \log N. \tag{48}$$

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.

C. 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. Intuitively these results could be regarded as saying that one can achieve a "diversity order of infinity" if using the long-term optimal power allocation, as plots of $\log P_{outage}$ vs N will approach a vertical asymptote, see also [28] for similar situations in the context of information outage minimization.

We first analyze the power required to achieve zero outage. For $N = 1$, the sum power expression in (9) 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

$$P_T(t) = \int_{\mathcal{R}_T(t)} P_{tot}(h) dF(h) \tag{49}$$

$$= \int_{\frac{K_1}{t}}^{\infty} \frac{K_1}{h} \lambda e^{-\lambda h} dh = \lambda K_1 \int_{\frac{\lambda K_1}{t}}^{\infty} \frac{e^{-u}}{u} du \tag{50}$$

$$= \lambda K_1 E_1\left(\frac{\lambda K_1}{t}\right) \tag{51}$$

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

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

The integral above is well-known to be finite for $N > 1$, see e.g. [33]. 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

$$\sigma^2\gamma_{th} = \frac{\sum_{i=1}^N f(h_i)}{N} \leq f\left(\frac{\sum_{i=1}^N h_i}{N}\right) \tag{52}$$

$$\Rightarrow \frac{\sigma^2\gamma_{th}}{N} \leq \frac{1}{\frac{\sigma_c^2 C}{\sigma^2 P_{tot}(\mathbf{h})\frac{1}{N}\sum_{i=1}^N h_i} + 1} \tag{53}$$

$$\Rightarrow \frac{\sigma^2\gamma_{th}}{N} \frac{\sigma_c^2 C}{\sigma^2 P_{tot}(\mathbf{h})\frac{1}{N}\sum_{i=1}^N h_i} \leq 1 - \frac{\sigma^2\gamma_{th}}{N} \tag{54}$$

$$\Rightarrow P_{tot}(\mathbf{h}) \geq \frac{K_N}{\sum_{i=1}^N h_i} \tag{55}$$

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 IV-C. The definition of $\mathcal{R}_T(\breve{t})$ becomes

$$\breve{\mathcal{R}}_T(\breve{t}) = \left\{\mathbf{h} : \breve{P}_{tot}(\mathbf{h}) < \breve{t}\right\} = \left\{\mathbf{h} : \sum_{i=1}^N h_i > \frac{K_N}{\breve{t}}\right\}. \tag{56}$$

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

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

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

$$\breve{N}_{max} = \left\lfloor \frac{(1+\sigma^2\gamma_{th})\mathcal{P}_{tot} + \gamma_{th}\sigma_c^2 C\lambda + \sqrt{[(1+\sigma^2\gamma_{th})\mathcal{P}_{tot} + \gamma_{th}\sigma_c^2 C\lambda]^2 - 4\mathcal{P}_{tot}^2\gamma_{th}\sigma^2}}{2\mathcal{P}_{tot}} \right\rfloor \tag{59}$$

---

distributed with parameters $k = N$, $\theta = \frac{1}{\lambda}$. Hence $\breve{P}_T(t)$ becomes

$$\begin{aligned}
\breve{P}_T(t) &= K_N \int_{\sum_{i=1}^N h_i > \frac{K_N}{\breve{t}}} \frac{1}{\sum_{i=1}^N h_i} e^{-\lambda \sum_{i=1}^N h_i} dh_1 \cdots dh_N \\
&= K_N \frac{1}{\Gamma(k)\theta^k} \int_{\frac{K_N}{\breve{t}}}^{\infty} T^{k-2} e^{-\frac{T}{\theta}} dT \\
&= \frac{K_N}{\Gamma(N)\lambda^{-n}} \int_{\frac{K_N}{\breve{t}}}^{\infty} T^{N-2} e^{-\lambda T} dT \\
&= \frac{K_N \lambda}{N-1} \cdot \frac{\Gamma\left(N-1, \lambda K_N/\breve{t}\right)}{\Gamma(N-1)}.
\end{aligned}$$

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

$$\frac{K_N \lambda}{N-1} \cdot \frac{\Gamma\left(N-1, \lambda K_N/\breve{t}^*\right)}{\Gamma(N-1)} = \mathcal{P}_{tot}. \tag{57}$$

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 (57) 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

$$\frac{K_N \lambda}{N-1} = \mathcal{P}_{tot}. \tag{58}$$

Equation (58) 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 (58) can be found in closed-form and is given as (59) where $\lfloor x \rfloor$ denotes the floor function of x.

D. General Parameters

The previous subsections have analyzed the diversity orders of distortion outage for "symmetric" sensor networks. Here we show how these results can be extended to the case where the sensor noise variances are not necessarily identical, and the case where the fading channels are not necessarily identically distributed. The idea is to upper and lower bound the distortion and hence the distortion outage probability, and show that asymptotically they have the same diversity orders of distortion outage. Similar techniques have been previously used in [34] and [35]. To avoid repetition, we will only treat the case of equal power allocation (EPA).

1) General sensor noise variances: We consider first the case where the sensor noise variances $\sigma_i^2, i = 1, \ldots, M$ are not necessarily identical, with the fading channels still assumed to be i.i.d. Rayleigh across sensors. We assume that the sensor noise variances can be bounded from both above and below, i.e.

$$0 < \sigma_{min}^2 \leq \sigma_i^2 \leq \sigma_{max}^2 < \infty, \forall i.$$

Such an assumption can cover the situation where sensors are placed deterministically but with different distances from the source, as well as the situation where the sensor noise variances σ_i^2 are random (but are upper and lower bounded, though not necessarily i.i.d.) due to random placement of the sensor nodes.

Then, since the distortion D is an increasing function of σ_i^2 for all i, a larger σ_i^2 will lead to a higher outage probability. Hence an upper bound on the outage probability is the case when we take $\sigma_i^2 = \sigma_{max}^2, \forall i$. From our results in Section IV-A, we obtain (taking the leading term only)

$$\log P_{outage} \leq -N \log N(1 + o(1))$$

as $N \to \infty$.

Similarly, a lower bound on the outage probability is the case when we take $\sigma_i^2 = \sigma_{min}^2, \forall i$. In this case, we obtain $\log P_{outage} \geq -N \log N(1 + o(1))$ as $N \to \infty$.

Since the upper and lower bounds on $\log P_{outage}$ both have the asymptotic behaviour $-N \log N$ as $N \to \infty$, the general situation will also do so. Hence the diversity order of $N \log N$ is also obtained in the case of general sensor noise variances.

2) Non-identically distributed fading channels: Here we consider the case where the sensor noise variances are identical, and the fading channels are independent but not necessarily identically distributed. In particular, we analyze the situation satisfying the following assumption:

Assumption 4.1: The channel gains h_i can be written as

$$h_i = \mu_i h_i', \forall i$$

where $\mu_i > 0$ are constants satisfying

$$0 < \mu_{min} \leq \mu_i \leq \mu_{max} < \infty,$$

and the h_i''s are identically distributed.

For instance, if h_i is exponentially distributed with mean $1/\lambda_i$, then we can take $\mu_i = 1/\lambda_i$, and h_i' will be exponentially distributed with mean 1. Thus Rayleigh fading channels with different means will satisfy Assumption 4.1.

We first derive an upper bound on the outage probability. From the distortion expression (11) we have

$$\begin{aligned}
\frac{D}{\sigma_\theta^2} &= \frac{\frac{\sum_{i=1}^N \mu_i h_i'}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}}{\frac{\sum_{i=1}^N \mu_i h_i'}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}} + \frac{\sigma_\theta^2 N}{\sigma^2} \left(\frac{\sum_{i=1}^N \sqrt{\mu_i h_i'}}{N} \right)^2} \\
&\leq \frac{\frac{\sum_{i=1}^N \mu_{max} h_i'}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}}{\frac{\sum_{i=1}^N \mu_{min} h_i'}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}} + \frac{\sigma_\theta^2 N}{\sigma^2} \left(\frac{\sum_{i=1}^N \sqrt{\mu_{min} h_i'}}{N} \right)^2}.
\end{aligned}$$

and the result

$$\frac{\frac{\sum_{i=1}^N \mu_{max} h'_i}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}}{\frac{\sum_{i=1}^N \mu_{min} h'_i}{N} + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}} + \frac{\sigma_\theta^2 N}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{\mu_{min} h'_i}}{N}\right)^2} - \frac{\eta_1}{\eta_2 + \frac{\sigma_\theta^2 N \mu_{min}}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h'_i}}{N}\right)^2} \xrightarrow{p} 0,$$

where $\eta_1 = \mu_{max} E[h'] + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}$ and $\eta_2 = \mu_{min} E[h'] + \frac{\sigma_c^2 C}{\sigma^2 \mathcal{P}_{tot}}$. Then

$$\begin{aligned} & P_{outage} \\ & \leq \Pr\left(\frac{\sigma_\theta^2 \eta_1}{\eta_2 + \frac{\sigma_\theta^2 N \mu_{min}}{\sigma^2}\left(\frac{\sum_{i=1}^N \sqrt{h'_i}}{N}\right)^2} > D_{max}\right)(1 + o(1)) \\ & = \Pr\left(\frac{1}{N}\sum_{i=1}^N \sqrt{h'_i} < \sqrt{\frac{\sigma^2(\sigma_\theta^2 \eta_1 - D\eta_2)}{D\sigma_\theta^2 N \mu_{min}}}\right)(1 + o(1)). \end{aligned}$$

From our results in Section IV-A, for Rayleigh fading we obtain

$$\begin{aligned} & \log P_{outage} \\ & \leq \log \Pr\left(\frac{1}{N}\sum_{i=1}^N \sqrt{h'_i} < \sqrt{\frac{\sigma^2(\sigma_\theta^2 \eta_1 - D\eta_2)}{D\sigma_\theta^2 N \mu_{min}}}\right)(1 + o(1)) \\ & \sim -N \log N \end{aligned}$$

as $N \to \infty$.

Similarly, we can derive a lower bound on the outage probability and show that

$$\log P_{outage} \geq -N \log N(1 + o(1)).$$

Since the upper and lower bounds have the same asymptotic behaviour, the general situation will also do so. Hence the diversity order of $N \log N$ is also obtained in the case of Rayleigh fading channels with different means.

3) General sensor noise variances and non-identically distributed fading channels: Combining the results in the previous subsections, we can see that if we have both different sensor noise variances and Rayleigh fading channels with different means, the diversity order of $N \log N$ is still achieved.

V. SIMULATION RESULTS

In this section, we show comparisons between Monte Carlo simulations and some of the asymptotic expressions for the diversity order that have been derived in this paper. The Monte Carlo simulations are obtained by averaging over 1,000,000 channel realizations.

We first present the diversity order of distortion outage for EPA. The parameters are chosen as follows.

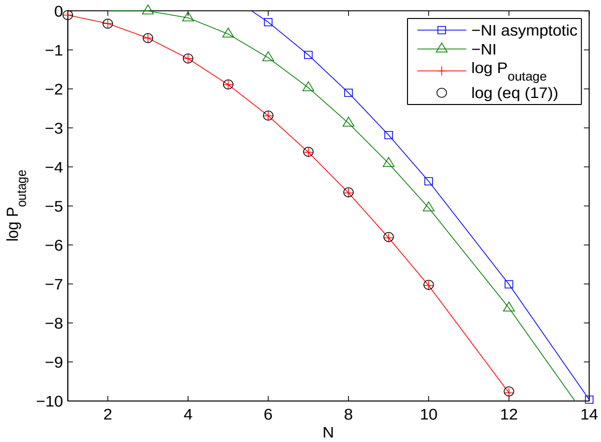


Fig. 3. EPA with total power constraint $\mathcal{P}_{tot} = 10m$W. Squares: Asymptotic expression (33). Triangles: $-NI_{\sqrt{h}}(a/\sqrt{N})$. Plus signs: $\log P_{outage}$. Circles: $\log\left(\Pr\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{h_i} < \frac{a}{\sqrt{N}}\right)\right)$. Simulation parameters: $\lambda = 250,000$, $a = 0.003$, $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$.

For simplicity we consider the source θ to be distributed as $N(0,1)$. The sensor measurement noise variance $\sigma^2 = 10^{-3}$ is chosen to represent the sensitivity of the measurement (small compared to the variance of θ), while the variance of the channel noise $\sigma_c^2 = 10^{-8}$ is chosen to be much smaller than the measurement noise. The parameter $\lambda = 250,000$ for the fading channel is used based on the loss (square law) at a distance of 500m (average distance between sensors to the FC). D_{max} is chosen to be 10% of the maximum distortion, which is equal to 1. The powers $\mathcal{P}_{tot}$ were chosen to be in the range of milliwatts, a reasonable transmission power in wireless sensor nodes. We simulated the case where $\mathcal{P}_{tot} = 10m$W and plotted the results in Fig. 3. We compare between 1) plots of $\log P_{outage}$ obtained via Monte Carlo simulation, where $\log$ is the natural log, 2) Monte Carlo simulations of $\log\left(\Pr\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{h_i} < \frac{a}{\sqrt{N}}\right)\right)$ from our approximation (13), 3) the exact values of $-NI_{\sqrt{h}}(a/\sqrt{N})$ obtained by solving (18) numerically, and 4) plots of the asymptotic expression (33). As mentioned before in Section IV-A, $\log\left(\Pr\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{h_i} < \frac{a}{\sqrt{N}}\right)\right)$ is a very good approximation to $\log P_{outage}$. Fig. 3 also shows that as N gets large, all four plots have similar gradients. Note that the asymptotic results $I_{\sqrt{h}}(a/\sqrt{N})$ and expression (33) 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. This is due to the use of asymptotic approximations to derive our expressions, e.g. the approximations made in going from equation (32) to (33).[3]

We next present the diversity order results for ST-OPA. Using $\mathcal{P}_{tot} = 10$mW, in Fig. 4 we compare between 1) Monte Carlo simulations of $\log P_{outage}$ using ST-OPA, 2) numerical computation of $-NI_Z(g_N)$, and 3) the asymptotic expression (47). We again see that as N increases, all three plots have very similar gradients.

We next present results for LT-OPA. An approximate relationship between N_{max} and $\mathcal{P}_{tot}$ for LT-OPA has

[3]Note also that changing the values of parameters such as the distortion threshold D_{max} or sensor measurement noise variance σ^2 will shift the curves up or down, however the diversity order is related to the slopes of the curves, which from our analysis in Section IV has leading term behaviour $\log P_{outage} \sim -N \log N$ that does not depend on the value of D_{max} or σ^2. For brevity we will omit these additional graphs in the paper.

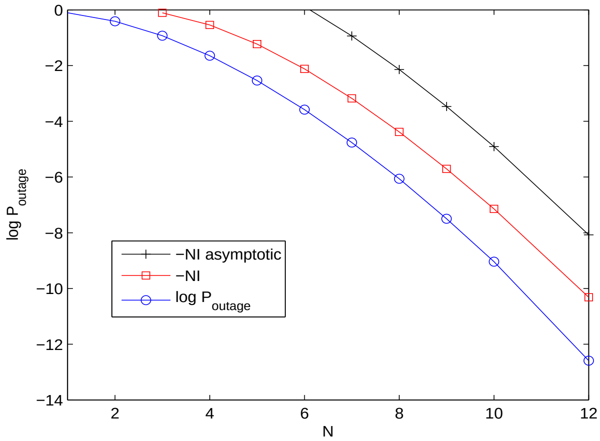


Fig. 4. ST-OPA with $\mathcal{P}_{tot} = 10mW$. Plus signs: Asymptotic expression (47). Squares: $-NI_Z(g/N)$. Circles: $\log P_{outage}$. Simulation parameters: $\lambda = 250,000$, $g = 0.09$, $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$.

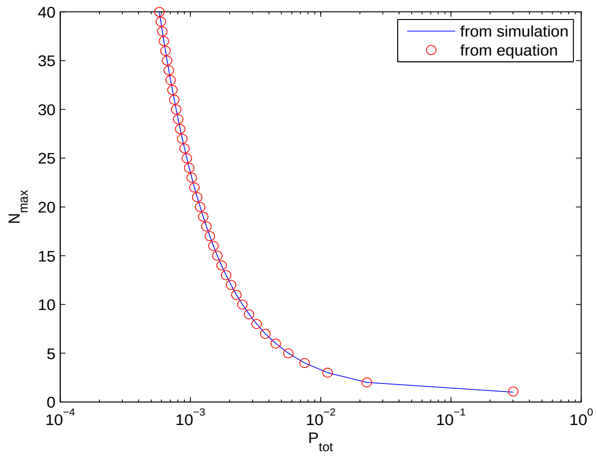


Fig. 5. N_{max} versus $\mathcal{P}_{tot}$. Circles: Approximation (59). Solid line: N_{max} from Monte Carlo simulation. Simulation parameters: $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$ and $\lambda = 250,000$.

been obtained in (59). To see how good this approximation is, we plot the approximation (59) and compare this with N_{max} obtained via Monte Carlo simulation, where we compute the average total power usage for a given N by averaging over 1,000,000 channel realizations. The results are shown in Fig. 5.

Finally, in Fig. 6 we compare the outage performance as a function of N for the three different power allocation schemes considered in this paper, 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. 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 $N_{max}+1$, where in this example $N_{max} = 15$.

VI. Conclusions

In this paper 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-

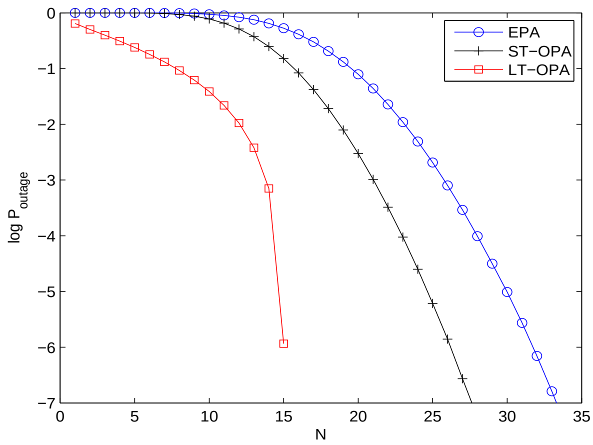


Fig. 6. $\log P_{outage}$ versus N. Simulation parameters: $\sigma_\theta^2 = 1$, $\sigma^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{max} = 0.1$, $\lambda = 250,000$ and $\mathcal{P}_{tot} = 1,600\mu$W.

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 [18] 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 deriving the diversity orders of distortion outage for different fading distributions. One may also extend this work to dynamical systems where the source is a time-varying Gauss Markov random process.

VII. Appendix

Proof: Lemma 3.2: 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

$$\frac{\partial \langle \mathbf{P}(\mathbf{h}) \rangle}{\partial h_i} = \frac{\partial}{\partial h_i} \frac{P_{tot}(\mathbf{h})}{N} = \frac{\sigma_c^2}{N} \frac{\partial \nu}{\partial h_i} \tag{60}$$

where $\nu = \frac{\mathcal{P}_{tot}}{\sigma_c^2}$ is the Lagrangian multiplier in one of the KKT conditions (see [19]). Also from the KKT conditions [19] we have

$$\sum_{i=1}^{N} \frac{\nu h_i}{C + \nu h_i \sigma^2} = \gamma_{th}. \tag{61}$$

Taking the partial derivative with respect to h_i on both sides of (61) gives

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

which completes the proof. ■

Proof: Theorem 2: We prove the theorem by obtaining upper and lower bounds of $\log \Pr\left(\frac{1}{n}\sum_{i=1}^{n}X_i\le a_n\right)$, which asymptotically are equivalent for large n. The proof uses similar techniques to those provided in the proof of Theorem 5.11.4 in [36].

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 $E\left[Y_i\right]=\mu_Y=0$. 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

$$I_Y(c_n)=\sup_{-t}\left\{\left(\mu_X-c_n\right)t-\Lambda_X(t)\right\}. \tag{62}$$

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 $c_n t-\Lambda(t)=\log\left(\frac{e^{c_n t}}{M_Y(t)}\right)=\log\left(\frac{1+c_n t+o(t)}{1+\frac{1}{2}\sigma_Y^2t^2+o(t^2)}\right)$ 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, $1+c_n t+o(t)>1+\frac{1}{2}\sigma_Y^2t^2+o(t^2)$, 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 $n\ge\mathcal{N}$ (the value of $\mathcal{N}$ 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

$$I_Y(c_n)=\sup_{t>0}\left\{c_n t-\Lambda_Y(t)\right\},\quad c_n>0 \text{ for } n\ge\mathcal{N}. \tag{63}$$

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

$$\Lambda''_Y(t) = \frac{M_Y(t)M''_Y(t) - M'_Y(t)^2}{M_Y(t)^2} \tag{64}$$

$$= \frac{E\left[e^{tY}\right]E\left[Y^2 e^{tY}\right] - E\left[Ye^{tY}\right]^2}{M_Y(t)^2} > 0 \tag{65}$$

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

We have the following

$$\begin{aligned}\Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \leq a_n\right) &= \Pr\left(\sum_{i=1}^{n} Y_i \geq nc_n\right) \\ = \Pr\left(e^{t\sum_{i=1}^{n} Y_i} \geq e^{nc_n t}\right) &\quad \text{for } t > 0 \\ \leq \frac{E\left[\exp\left(t\sum_{i=1}^{n} Y_i\right)\right]}{e^{nc_n t}} &= e^{-nc_n t} M_Y(t)^n = e^{-n(c_n t - \Lambda_Y(t))}\end{aligned}$$

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

$$\log \Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \leq a_n\right) \leq -n\left(c_n t - \Lambda_Y(t)\right), \quad \forall t > 0 \tag{66}$$

Since the upper bound in (66) 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

$$\begin{aligned}&\log \Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \leq a_n\right) \\ &\leq \inf_{t>0}\left\{-n\left(c_n t - \Lambda_Y(t)\right)\right\} \\ &= -n\sup_{t>0}\left\{c_n t - \Lambda_Y(t)\right\} \\ &= -nI_Y^{+}\left(c_n\right) = -nI_X^{-}\left(a_n\right) \quad \text{from (62).}\end{aligned}$$

Lower bound. We first show that the problem falls under the regular case, i.e., that the supremum of the rate function $I_Y(c_n)$, $n \geq \mathcal{N}$ 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 $\Pr\left(Y_i > c_n\right) > 0$ for $n \geq \mathcal{N}$, there exists $b_n \in (c_n, \infty)$ such that $\Pr(Y_i > b_n) > 0$.

It follows that for $t > 0$,

$$\begin{aligned}
&c_n t - \Lambda_Y(t) \\
=&c_n t - \log E\left[e^{tY}\right] = c_n t - \log \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy \\
=&c_n t - \log \left[\int_{-\infty}^{b_n} e^{ty} f_Y(y) dy + \int_{b_n}^{\infty} e^{ty} f_Y(y) dy\right] \\
\leq&c_n t - \log \int_{b_n}^{\infty} e^{ty} f_Y(y) dy \leq c_n t - \log \left\{e^{tb_n} \int_{b_n}^{\infty} f_Y(y) dy\right\} \\
=&c_n t - \log \left\{e^{tb_n} \Pr(Y_i > b_n)\right\} \\
=&-\left(b_n - c_n\right) t - \log \Pr(Y_i > b_n) \to -\infty \text{ as } t \to \infty
\end{aligned}$$

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

$$dF_{\tilde{Y}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)} dF_Y(y), \tag{67}$$

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

$$M_{\tilde{Y}_n}(t) = \int_{-\infty}^{\infty} e^{tu} dF_{\tilde{Y}_n}(u) \tag{68}$$

$$= \int_{-\infty}^{\infty} \frac{e^{(t+\tau_n)u}}{M_Y(\tau_n)} dF_Y(u) = \frac{M_Y(t+\tau_n)}{M_Y(\tau_n)}. \tag{69}$$

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

$$E\left[\tilde{Y}_{n,i}\right] = M'_{\tilde{Y}_n}(0) = \frac{M'_Y(\tau_n)}{M_Y(\tau_n)} = \Lambda'_Y(\tau_n) = c_n, \tag{70}$$

$$\text{var}\left(\tilde{Y}_{n,i}\right) = E\left[\left(\tilde{Y}_{n,i}\right)^2\right] - \left(E\left[\tilde{Y}_{n,i}\right]\right)^2 \tag{71}$$

$$= M''_{\tilde{Y}_n}(0) - M'_{\tilde{Y}_n}(0)^2 \tag{72}$$

$$= \Lambda''_Y(\tau_n) \in (0, \infty). \tag{73}$$

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

$$M_{\tilde{S}_n}(t) = \left(\frac{M_Y(t+\tau_n)}{M_Y(\tau_n)}\right)^n = \frac{E\left[e^{(t+\tau_n)\tilde{S}_n}\right]}{M_Y(\tau_n)^n} \tag{74}$$

$$= \frac{1}{M_Y(\tau_n)^n} \int_{-\infty}^{\infty} e^{(t+\tau_n)u} dF_{S_n}(u) \tag{75}$$

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

$$dF_{\tilde{S}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)^n} dF_{S_n}(y). \tag{76}$$

Let $d > 0$. We have

$$\begin{aligned}
&\Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \le a_n\right) = \Pr\left(\sum_{i=1}^{n} Y_i \ge nc_n\right) \\
=&\int_{nc_n}^{\infty} dF_{S_n}(u) = \int_{nc_n}^{\infty} M_Y(\tau_n)^n e^{-\tau_n u} dF_{\tilde{S}_n}(u) \\
\ge& M_Y(\tau_n)^n \int_{nc_n}^{n(c_n+d)} e^{-\tau_n u} dF_{\tilde{S}_n}(u) \\
\ge& M_Y(\tau_n)^n e^{-n(c_n+d))\tau_n} \int_{nc_n}^{n(c_n+d)} dF_{\tilde{S}_n}(u) \\
=& e^{-n(\tau_n(c_n+d)-\Lambda_Y(\tau_n))} \Pr\left(nc_n < \tilde{S}_n < n\left(c_n+d\right)\right) \\
=& e^{-n(\tau_n(c_n+d)-\Lambda_Y(\tau_n))} \Pr\left(c_n < \frac{1}{n}\tilde{S}_n < c_n + d\right).
\end{aligned}$$

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 [37]. Therefore,

$$\begin{aligned}
&\log \Pr\left(\frac{1}{n}\sum_{i=1}^{n} X_i \le a_n\right) \\
\ge& -n\left(\tau_n\left(c_n+d\right) - \Lambda_Y(\tau_n)\right) \\
&\quad + \log \Pr\left(c_n < \frac{1}{n}\tilde{S}_n < c_n + d\right) \\
\sim& -n\left(\tau_n\left(c_n+d\right) - \Lambda_Y(\tau_n)\right) \quad \text{as } n \to \infty \\
\sim& -n\left(\tau_n c_n - \Lambda_Y(\tau_n)\right) \quad \text{as } d \to 0 \\
=& -nI_Y^{+}\left(c_n\right) = -nI_X^{-}\left(a_n\right).
\end{aligned}$$

■

Proof: Lemma 4.1: Here we want to show that

$$\Pr\left(\frac{1}{n}\tilde{S}_n > c_n\right) \to 0.5 \tag{77}$$

as $n \to \infty$. We note that the L.H.S. of (77) 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

$$\frac{1}{\sigma_{\tilde{Z}_n}^2} \int_{\left\{|\tilde{Z}_n| > \epsilon\sqrt{n\sigma_{\tilde{Z}_n}^2}\right\}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} \to 0 \text{ as } n \to \infty \tag{78}$$

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:

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

$$\frac{d^2\Lambda_Y(\theta)}{d\theta^2} = \frac{\frac{d^2 M_{\sqrt{h}}(-\theta)}{d\theta^2} M_{\sqrt{h}}(-\theta) - \left(\frac{dM_{\sqrt{h}}(-\theta)}{d\theta}\right)^2}{M_{\sqrt{h}}(-\theta)^2}. \tag{79}$$

Note that

$$\frac{dM_{\sqrt{h}}(-\theta)}{d\theta} = \frac{\left(\kappa^2\theta^2 + 1\right) M_{\sqrt{h}}(-\theta) - 1}{\theta} \tag{80}$$

$$\frac{d^2 M_{\sqrt{h}}(-\theta)}{d\theta^2} = \kappa^2 \left[\left(\kappa^2\theta^2 + 3\right) M_{\sqrt{h}}(-\theta) - 1\right]. \tag{81}$$

Substituting (80) and (81) into (79) gives

$$\frac{d^2\Lambda_Y(\theta)}{d\theta^2} = \frac{\left(\kappa^2\theta^2 - 1\right) M_{\sqrt{h}}(-\theta)^2 + \left(\kappa^2\theta^2 + 2\right) M_{\sqrt{h}}(-\theta) - 1}{\theta^2 M_{\sqrt{h}}(-\theta)}. \tag{82}$$

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

$$\text{var}\left(\tilde{Y}_n\right) = \Lambda_Y''\left(\tau_n\right) \sim \frac{a^2}{2n}. \tag{83}$$

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

$$f_{\tilde{Z}_n}\left(\tilde{z}\right) = \frac{a_n - \tilde{z}}{\kappa^2 M_Y(\tau_n)} e^{-\frac{(a_n - \tilde{z})^2}{2\kappa^2} + \tau_n(\tilde{z} + c_n)}. \tag{84}$$

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

We are now ready to look at Lindeberg's condition (78) after obtaining the expressions (83) and (84). We have

$$\begin{aligned}
&\frac{1}{\sigma^2_{\tilde{Z}_n}} \int_{\left\{|\tilde{Z}_n|>\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}\right\}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} \\
=&\frac{1}{\sigma^2_{\tilde{Z}_n}} \left(\int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} + \int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{a_n} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} \right) \\
\overset{(a)}{\sim}&\frac{1}{\sigma^2_{\tilde{Z}_n}} \int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} \\
=&\frac{1}{\kappa^2 \sigma^2_{\tilde{Z}_n} M_Y(\tau_n)} \\
&\quad \times \int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}} (a_n - \tilde{z})\, \tilde{z}^2 e^{-\frac{(a_n-\tilde{z})^2}{2\kappa^2}+\tau_n(\tilde{z}+c_n)} d\tilde{z} \\
=&\frac{e^{-\mu_{\sqrt{h}}\tau_n}}{\kappa^2 \sigma^2_{\tilde{Z}_n} M_{\sqrt{h}}(-\tau_n)} \\
&\quad \times \int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{\infty} (a_n + u)\, u^2 e^{-\frac{(a_n+u)^2}{2\kappa^2}+\tau_n(c_n-u)} du \\
=&\frac{1}{\kappa^2 \sigma^2_{\tilde{Z}_n} M_{\sqrt{h}}(-\tau_n)} \\
&\quad \times \int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{\infty} (a_n + u)\, u^2 e^{-\frac{(a_n+u)^2}{2\kappa^2}-\tau_n u+\tau_n c_n-\tau_n\mu_{\sqrt{h}}} du
\end{aligned} \tag{85}$$

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 $n \to \infty$)

$$a_n = a/\sqrt{n} \to 0 \tag{86}$$

$$c_n = \mu_{\sqrt{h}} - a_n \to \mu_{\sqrt{h}} \tag{87}$$

$$\tau_n \sim \frac{2\sqrt{n}}{a} \quad \text{(from (27) and } \tau = -\theta) \tag{88}$$

$$M_{\sqrt{h}}(-\tau_n) \sim \frac{a^2}{4\kappa^2 n} \tag{89}$$

$$\sigma^2_{\tilde{Z}_n} \sim \frac{a^2}{2n}. \tag{90}$$

We now show that (85) goes to zero as $n \to \infty$ by using an upper bound of (85) 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 (85): 1) $e^{-\frac{(a_n+u)^2}{2\kappa^2}} \leq e^{-\frac{u^2}{2\kappa^2}}$ (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 (86) and (88)) $\Rightarrow e^{\tau_n c_n - \tau_n \mu_{\sqrt{h}}} \leq C$ (for sufficiently large n), 3) $e^{-\tau_n u} = e^{-\frac{2\sqrt{n}}{a}u(1+o(1))} \leq e^{-\frac{\sqrt{n}}{a}u}$ (for sufficiently large n) , where C is a constant.

Hence we substitute the upper bounds obtained above and the asymptotic expressions (86), (89) and (90) into (85) and obtain the following upper bound

$$\begin{aligned}
&\frac{1}{\kappa^2\sigma^2_{\tilde{Z}_n}M_{\sqrt{h}}(-\tau_n)} \\
&\qquad\times\int_{\epsilon\sqrt{n\sigma^2_{\tilde{Z}_n}}}^{\infty}(a_n+u)\,u^2e^{-\frac{(a_n+u)^2}{2\kappa^2}-\tau_n u+\tau_n c_n-\tau_n\mu_{\sqrt{h}}}du \\
\leq&\frac{8Cn^2}{a^4}\int_{a\epsilon/\sqrt{2}}^{\infty}u^3e^{-\frac{u^2}{2\kappa^2}}e^{-\frac{\sqrt{n}}{a}u}du\,(1+o(1))\,.
\end{aligned} \tag{91}$$

We may use Laplace's method [30] to obtain an asymptotic approximation of

$$\mathcal{I}(\sqrt{n})=\int_{a\epsilon/\sqrt{2}}^{\infty}u^3e^{-\frac{u^2}{\kappa^2}}e^{-\sqrt{n}\frac{u}{a}}du \tag{92}$$

in (91). Let $h(u)=u/a$ and $\varphi(u)=u^3e^{-\frac{u^2}{2\kappa^2}}$. Hence the integral becomes

$$\mathcal{I}(\sqrt{n})=\int_{A}^{\infty}\varphi(u)e^{-\sqrt{n}h(u)}du \tag{93}$$

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 (Theorem 1 in Ch 2 of [30]). 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

$$\mathcal{I}(\sqrt{n})\sim e^{-\sqrt{n}h(A)}\sum_{s=0}^{\infty}\Gamma\left(\frac{s+\alpha}{\mu}\right)\frac{c_s}{\sqrt{n}^{(s+\alpha)/\mu}} \tag{94}$$

$$\sim\frac{c_0}{\sqrt{n}}e^{-\epsilon\sqrt{2n}} \tag{95}$$

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

$$\frac{8c_0n\sqrt{n}}{a^4}e^{-\epsilon\sqrt{2n}}\to 0\quad\text{as } n\to\infty. \tag{96}$$

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

Proof: Lemma 4.3: 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 [30] 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 [30]). 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 $c_1 = \left\{\frac{b_1}{\mu} - \frac{(\alpha+1)a_1 b_0}{\mu^2 a_0}\right\} \frac{1}{a_0^{(\alpha+1)/\mu}}$ [30]. 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 (41) gives

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

which completes the proof. ■

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] 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, June 2005.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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, Sept. 2005.

[10] 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.

[11] 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.

[12] 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, July 2004.

[13] 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.

[14] 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.

[15] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," *Lecture Notes in Computer Science*, vol. 2634, p. 162177, Apr. 2003.

[16] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian "Sensor" network," *IEEE Trans. Inform. Theory*, vol. 54, no. 11, pp. 5247–5251, Nov. 2008.

[17] 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*, p. 121125, June 2005.

[18] 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, Sept. 2007.

[19] 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.

[20] I. Bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, pp. 1146–1156, July 2008.

[21] 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*, pp. 473–482, Sept. 2005.

[22] C. Wang and S. Dey, "Power allocation for distortion outage minimization in clustered wireless sensor networks," in *Proc. IEEE International Wireless Communications and Mobile Computing Conference (IWCMC'08)*, Crete, Greece, Aug. 2008, pp. 395–400.

[23] 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.

[24] 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.

[25] 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.

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

[27] 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, July 2007.

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

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

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

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

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

[33] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1165–1181, July 1999.

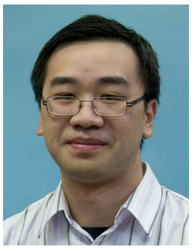
[34] A. S. Leong, S. Dey, and J. S. Evans, "Asymptotics and power allocation for state estimation over fading channels," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 611–633, Jan. 2011.

[35] A. S. Leong and S. Dey, "On scaling laws of diversity schemes in decentralized estimation," *IEEE Trans. Inform. Theory*, vol. 57, no. 7, pp. 4740–4759, July 2011.

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

[37] T.-C. 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.

Chih-Hong Wang was born in Kaohsiung, Taiwan in 1984. He received the B.E. degree in Electrical and Electronic Engineering (First-Class Hons.) in 2006 from the University of Auckland, New Zealand. He is currently pursuing the Ph.D. degree in the Department of Electrical and Electronic Engineering at the University of Melbourne, Australia. His research interests include wireless communications, statistical signal processing and optimisation.



Alex Leong was born in Macau in 1980. He received the B.S. degree in mathematics and B.E. degree in electrical engineering in 2003, and the Ph.D. degree in electrical engineering in 2008, all from the University of Melbourne, Parkville, Australia. He is currently a research fellow in the Department of Electrical and Electronic Engineering, University of Melbourne. His research interests include statistical signal processing, signal processing for sensor networks, and networked control systems. Dr. Leong was the recipient of the L. R. East Medal from Engineers Australia in 2003, and an Australian Postdoctoral Fellowship from the Australian Research Council in 2009.

Subhrakanti Dey was born in Calcutta, India, in 1968. He received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, in 1996.

He has been with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, since February 2000, where he is currently a full Professor. From September 1995 to September 1997 and September 1998 to February 2000, he was a postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a post-doctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive estimation and control. Prof. Dey currently serves on the Editorial Board of Elsevier Systems and Control Letters. He was also an Associate Editor for the IEEE Transactions on Signal Processing and the IEEE Transactions on Automatic Control. He is a Senior Member of IEEE.