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 analog amplify and forward 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 distortion subject to a power constraint at each time step, and 3) long-term optimal power allocation (LT-OPA) where we minimize 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 large numbers 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.

I. INTRODUCTION

Wireless sensor networks have recently attracted research interests and practical implementations in many areas of human life due to the numerous applications WSNs (wireless sensor networks) 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 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.

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).
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 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 distortion constraint and minimizing distortion under 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 $D_{\text{max}}$. 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 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 such a network 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, 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 $N > 1$. 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 $\langle \mathbf{x} \rangle \triangleq \frac{1}{N} \sum_{i=1}^{N} x_i$. 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 $\sigma^2_\theta$, and is i.i.d. (independent and identically distributed) in time ($k$ denotes the discrete time index). The measurements of the $i$th sensor at time $k$ are given as

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

where $w_i$ is Gaussian with zero mean and variance $\sigma^2_i$ and denotes the sensor measurement noise. 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 (as in the block fading model). The channel noise is i.i.d. AWGN denoted as $n_c[k]$, with zero mean and variance $\sigma^2_c$. 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 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

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

(1)

Remark 1: Note that in this paper we are not claiming that such perfect synchronization at the sensor transmitters or in other words, distributed transmit beamforming is a realistic assumption. However our goal in this 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.

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

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

where $C_i = \sigma^2_\theta + \sigma^2_i$. 

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

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

Note that (2) gives the expression of the instantaneous distortion, i.e., it is a function of the channel realizations $h_i, \forall i, k$. Due to the randomness of the fading channels, the instantaneous distortion at the FC changes randomly over time. Such estimation networks usually impose a distortion threshold at the FC to guarantee acceptable estimation, and if the instantaneous distortion $D_k$ exceeds the distortion threshold $D_{\text{max}}$, 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 $P_{\text{outage}} \triangleq \Pr(D_k > D_{\text{max}})$.

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, $P_i[k](h[k])$ will be assumed to be a function of the channel gains.

**Remark:** Due to the i.i.d. (in time) nature of the network model, we will drop the time index $k$ from the rest of the 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 $P_{\text{tot}}$ 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 $P_{\text{tot}}$ 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 $P_{\text{tot}}$, the individual sensor power is then given as $P_i = P_{\text{tot}}/N, \forall i$.

#### 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

$$\min \quad D(P(h), h)$$

$$\text{s.t.} \quad \sum_{i=1}^{N} P_i(h) \leq P_{\text{tot}}, \quad P_i(h) \geq 0 \quad \forall i.$$ (3)
Problem (3) has been solved in [19]. The short-term optimal power allocation of the $i$th sensor is given by

$$P_i^*(h) = P_{tot} c_i(h_i) \left( \sum_{j=1}^{N} c_j(h_j) \right)^{-1} \quad \forall i$$  \hspace{1cm} (4)$$

where $c_i(h_i) = C_i h_i / \left( C_i + P_{tot} h_i \sigma_i^2 / \sigma_c^2 \right)^2$. From (4) we see that the optimal power of the $i$th sensor is computed by multiplying $P_{tot}$ by a ratio that is bounded between zero and one, i.e., we divide up $P_{tot}$ 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. Since the problem now deals with an extra dimension in time, an appropriate performance measure is the distortion outage probability introduced in Section II. 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

$$\min \ Pr \left( D \left( \mathbf{P}(h), h \right) > D_{max} \right)$$

$$\text{s.t.} \quad E \left( \sum_{i=1}^{N} P_i(h) \right) \leq P_{tot}, \quad P_i(h) \geq 0 \quad \forall i.$$  \hspace{1cm} (5)

Problem (5) can be solved in a similar way to [28]. First consider the following minimization problem given as

$$\min \ \langle \mathbf{P}(h) \rangle$$

$$\text{s.t.} \quad D(\mathbf{P}(h), h) \leq D_{max}, \quad P_i(h) \geq 0 \quad \forall i.$$  \hspace{1cm} (6)

We have the following lemma:

Lemma 3.1: With the knowledge of $\mathbf{h}$, the solution of problem (6) is given as

$$P_i^*(h) = P_{tot}(h) c_i(h_i) \left( \sum_{j=1}^{N} c_j(h_j) \right)^{-1}, \quad i = 1, \ldots, N$$  \hspace{1cm} (7)

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

$$\gamma_{th} = \sum_{i=1}^{N} \frac{h_i}{\left( \frac{\sigma_i^2 C_i}{P_{tot}(h)} + \frac{\sigma_c^2 \gamma_{th}}{h_i} \right)}$$  \hspace{1cm} (8)

where $\gamma_{th} = 1/D_{max} - 1/\sigma_0^2$.

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 (5):

Lemma 3.2: The long-term optimal power $\mathbf{P}^*(h) = [P_1^*(h), \ldots, P_N^*(h)]^T$ as given in (7), is a continuous function of $\mathbf{h}$. Furthermore, $\langle \mathbf{P}^*(h) \rangle$ is a non-increasing function of $h_i$ for $i = 1, \ldots, N$. 
Proof: See Appendix.

Before we give the solution to problem (5), we will also need the following definitions and notations, similar to those in [28]. We first define the regions \( \mathcal{R}_T(t) = \{ \mathbf{h} : \sum_{i=1}^{N} P_i(h) < t \} \), \( \overline{\mathcal{R}}_T(t) = \{ \mathbf{h} : \sum_{i=1}^{N} P_i(h) \leq t \} \) and \( \mathcal{B}_T(t) = \{ \mathbf{h} : \sum_{i=1}^{N} P_i(h) = t \} \). We then define two power sum quantities as \( P_T(t) = \int_{\mathcal{R}_T(t)} \sum_{i=1}^{N} P_i(h) dF(h) \) and \( \overline{P}_T(t) = \int_{\overline{\mathcal{R}}_T(t)} \sum_{i=1}^{N} P_i(h) dF(h) \), where \( F(h) \) denotes the joint c.d.f. of \( \mathbf{h} \). Finally, the power sum threshold \( t^* \) and the weight \( u^* \) are given as \( t^* = \sup \{ t : P_T(t) < \mathcal{P}_{\text{tot}} \} \) and \( u^* = \frac{\mathcal{P}_{\text{tot}} - P_T(t^*)}{P_T(t^*)} \).

With the above lemma and definitions we can now present the solution to problem (5).

**Theorem 1:** The solution of problem (5) is given as

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

while if \( \mathbf{h} \in \mathcal{B}_T(t^*) \), \( \hat{P}(\mathbf{h}) = P^*(\mathbf{h}) \) with probability \( u^* \) and \( \hat{P}(\mathbf{h}) = 0 \) with probability \( 1 - u^* \), where \( P^*(\mathbf{h}) \) is given in (7).

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 \( \mathcal{P}_{\text{tot}} \), then the sensors should transmit with powers given by (7) and achieve a distortion of exactly \( D_{\text{max}} \). 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 \( P_{\text{tot}}(\mathbf{h}) \). From \( P_{\text{tot}}(\mathbf{h}) \) and the definition of \( t^* \), the outage event only occurs if \( P_{\text{tot}}(\mathbf{h}) > t^* \). Hence in every transmission, the fusion center simply computes the quantity \( P_{\text{tot}}(\mathbf{h}) \) and compares it against \( t^* \). If \( P_{\text{tot}}(\mathbf{h}) > t^* \), then all sensors should be turned off to save power. Otherwise, the sensors should transmit with power given by (7). The value of \( t^* \) would depend on the value of \( \mathcal{P}_{\text{tot}} \) and it can be predetermined numerically in off-line mode via Monte-Carlo simulation. A closed-form expression of a lower bound on \( t \) is given in Section IV-C which allows one to quickly compute a lower bound of \( t^* \) given \( \mathcal{P}_{\text{tot}} \).

**IV. DIVERSITY ORDERS OF DISTORTION OUTAGE**

We are interested in seeing how the outage probability decays as the number of sensors increases. In this section we will obtain for large \( N \) asymptotic closed-form expressions of \( \log P_{\text{outage}} \), 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 for \( N \) i.i.d. orthogonal MAC. For analytical tractability, in the following theoretical analysis, we will only consider a homogeneous wireless sensor network where all the measurement noise and fading distributions are
i.i.d. As a consequence, we will denote $\sigma_i^2 = \sigma^2$ and $C_i = C = \sigma^2 + \sigma^2$, $\forall i$.

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

### A. Equal power allocation

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

$$
\frac{D}{\sigma_0^2} = \frac{\sum_{i=1}^{N} h_i}{\sum_{i=1}^{N} \sqrt{h_i}} + \frac{\sigma^2}{\sigma^2_{\text{tot}}} + \frac{\sigma^2_N}{\sigma^2} \left( \frac{\sum_{i=1}^{N} \sqrt{h_i}}{N} \right)^2.
$$

(10)

Inspecting the RHS of (10), we note that $\frac{1}{N} \sum_{i=1}^{N} h_i$ and $\frac{\sigma^2}{\sigma^2_{\text{tot}}}$ converge to $E[h]$ and $E[\sqrt{h}]$ respectively by the strong law of large numbers as $N$ gets large. However we find that $\text{var}(\frac{1}{N} \sum_{i=1}^{N} h_i) = \frac{1}{N} \text{var}[h]$ and $\text{var}(\frac{\sigma^2_N}{\sigma^2} (\frac{\sum_{i=1}^{N} \sqrt{h_i}}{N})^2) \approx \frac{4\sigma^4_N}{\sigma^4} (E[\sqrt{h}])^2 \text{var}[\sqrt{h}]$ (obtained using the Delta method [30]). We see that the variance of $\frac{1}{N} \sum_{i=1}^{N} h_i$ decreases like $1/N$, whereas the approximate variance of $\frac{\sigma^2_N}{\sigma^2} (\frac{\sum_{i=1}^{N} \sqrt{h_i}}{N})^2$ increases with $N$. We therefore choose to replace $\frac{1}{N} \sum_{i=1}^{N} h_i$ by its mean $E[h]$, and retain $\frac{\sigma^2_N}{\sigma^2} (\frac{1}{N} \sum_{i=1}^{N} \sqrt{h_i})^2$ for large $N$. This gives us the following result where the distortion converges (for large $N$) almost surely to a random variable expressed as:

$$
D \xrightarrow{a.s.} \frac{\sigma_0^2}{\sigma^2} \eta \left( \frac{\sigma^2_N}{\sigma^2} \left( \frac{\sum_{i=1}^{N} \sqrt{h_i}}{N} \right)^2 \right)^{-1}
$$

(11)

where $\eta = E[h] + \frac{\sigma^2_C}{\sigma^2_{\text{tot}}}$.  

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

$$
P_{\text{outage}} = \Pr(D > D_{\text{max}}) = \Pr \left( \frac{1}{N} \sum_{i=1}^{N} \sqrt{h_i} < \sqrt{\frac{\eta \sigma^2}{D_{\text{max}}}} \left( \frac{\sigma^2_N}{\sigma^2} - \frac{D_{\text{max}}}{\sigma^2_{\text{tot}}} \right) \right) \quad (\text{substitute (11) and re-arrange})
$$

(13)

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

(14)

where

$$
a = \sqrt{\frac{\eta \sigma^2}{D_{\text{max}}}} \left( \frac{\sigma^2_N}{\sigma^2} - \frac{D_{\text{max}}}{\sigma^2_{\text{tot}}} \right).
$$

By inspecting (14) 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 (14). Before we give the theorem we need the following definitions. The moment-generating function of the random variable $X$ is defined as $M_X(t) \triangleq E[e^{tX}]$. The cumulant-generating function of the random variable $X$ is defined as $\Lambda_X(t) \triangleq \log M_X(t)$. The rate function of the random variable $X$ is defined as $I_X(c) = \sup_t \{ct - \Lambda_X(t)\}$. 
We also define the following notations relating to the rate function as \( I^+_X(c) = \sup_{t > 0} \{ ct - \Lambda_X(t) \} \) and \( I^-_X(c) = \sup_{t < 0} \{ ct - \Lambda_X(t) \} \). Note here that \( I^+_X \) and \( I^-_X \) have the same value as \( I_X \); these two notations are introduced only to further restrict the domain of the supremum without affecting the result of \( I_X \). Hence these notations may be used interchangeably depending on whether we have extra knowledge of the domain over which the supremum is achieved.

**Theorem 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 \( M_X(t) = E[e^{tX}] \) 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

\[
    dF_{\tilde{Y}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)} dF_Y(y)
\]

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

\[
    \log \Pr\left(\frac{1}{n} \sum_{i=1}^{n} X_i \leq a_n\right) \sim -n I_X(a_n) \quad \text{as} \quad n \to \infty.
\]

**Proof:** See Appendix.

In order to apply Theorem 2 to (14), we need to verify the assumption that \( \Pr(\frac{1}{n} \sum_{i=1}^{n} \tilde{Y}_{n,i} > E[\tilde{Y}_{n,i}]) \) 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 \( Y_i = -\sqrt{h_i} + E[\sqrt{h_i}] \), where \( \sqrt{h_i} \) is Rayleigh distributed with parameter \( \kappa \) (i.e. \( f_{\sqrt{h}}(x) = \frac{x}{\kappa} e^{-x^2/2\kappa^2} \)). Denote \( \tilde{Y}_{n,i} \) as the exponential change of distribution of \( Y_i \) as defined in (15). Then

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

**Proof:** See Appendix.

Applying Theorem 2 to (14) we have

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

where

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

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

\[
    M_{\sqrt{h}}(-\frac{\sqrt{2}x}{\kappa}) = 1 - \sqrt{\pi} x e^{x^2} \text{erfc} (x)
\]

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

We need to find the value of \( \theta \) that attains the supremum in the rate function \( I^-_{\sqrt{h}}(a/\sqrt{N}) \). This value of \( \theta \)
can be found by using the stationary condition (first derivative) given as
\[
\frac{dI}{d\theta} \left( \frac{a/\sqrt{N}}{\sqrt{h}} \right) = 0, \quad \theta < 0
\]
(21)
\[
\Rightarrow \frac{\sqrt{N}}{a} = \psi(\theta)
\] 
(22)
where
\[
\psi(\theta) = \left( \Lambda'_{\sqrt{h}}(\theta) \right)^{-1} = M_{\sqrt{h}}(\theta)/M'_{\sqrt{h}}(\theta).
\] 
(23)

After substituting \( \theta = -\sqrt{2}x/\kappa \) in (22) and some algebraic manipulation, it is possible to obtain
\[
\frac{\sqrt{N}}{2} = \psi \left( -\frac{\sqrt{2}x}{\kappa} \right)
\] 
(24)
where
\[
\psi \left( -\frac{\sqrt{2}x}{\kappa} \right) = \frac{\sqrt{2}}{\kappa} \frac{xM_{\sqrt{h}} \left( -\frac{\sqrt{2}x}{\kappa} \right)}{1 - M_{\sqrt{h}} \left( -\frac{\sqrt{2}x}{\kappa} \right) - 2x^2M_{\sqrt{h}} \left( -\frac{\sqrt{2}x}{\kappa} \right)}
\] 
(25)
Note that \( \psi \left( -\frac{\sqrt{2}x}{\kappa} \right) \) is a continuous non-decreasing function of \( x \) since
\[
\frac{d\psi \left( -\frac{\sqrt{2}x}{\kappa} \right)}{dx} = \frac{\sqrt{2}}{\kappa} \left( \frac{\Lambda''_{\sqrt{h}} \left( -\frac{\sqrt{2}x}{\kappa} \right)}{\left( \Lambda'_{\sqrt{h}} \left( -\frac{\sqrt{2}x}{\kappa} \right) \right)^2} \right) \geq 0
\] 
(26)
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 (23); 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 \( \theta \) to \( x \) preserves the continuity of the function.

Hence from (24), large \( N \) corresponds to the case of large \( x \). We now show that \( \psi \left( -\frac{\sqrt{2}x}{\kappa} \right) \) 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 (20) and obtain
\[
M_{\sqrt{h}} \left( -\frac{\sqrt{2}x}{\kappa} \right) = \frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots.
\] 
(27)
We then substitute (27) into (24) to obtain the following
\[
\frac{\sqrt{N}}{a} = \frac{\sqrt{2}}{\kappa} \frac{x \left( \frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots \right)}{1 - \frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots} - 2x^2 \left( \frac{1}{2x^2} - \frac{3}{4x^4} + \frac{15}{8x^6} + \cdots \right)
\] 
(28)
\[
= \frac{\sqrt{2}}{\kappa} \frac{x}{2x^2} \frac{1}{1 - \frac{3}{4x^2} + \frac{15}{8x^4} + \cdots} \sim \frac{\sqrt{2}}{\kappa} \frac{x}{2} \quad \text{for large } x
\] 
(29)
Hence for large \( N \),
\[
\theta \sim -\frac{2\sqrt{N}}{a}.
\] 
(30)
Substituting this asymptotic expression for \( \theta \) back into the rate function gives

\[
I_{\sqrt{N}} \left( \frac{a}{\sqrt{N}} \right) \sim - \frac{a}{\sqrt{N}} \frac{2\sqrt{N}}{a} - \log M_{\sqrt{N}} \left( - \frac{2}{aN} \right) = -2 - \log M_{\sqrt{N}} \left( - \frac{2}{aN} \right)
\]

\[
\sim -2 - \log \left( \frac{a^2}{2\kappa^2} \right)
\]

\[
= -2 - \log \left( \frac{a^2}{2\kappa^2} \right) + \log N.
\]

Hence from (18) the outage probability for large \( N \) satisfies

\[
\log P_{\text{outage}} \sim -NI_{\sqrt{N}} \left( \frac{a}{\sqrt{N}} \right)
\]

\[
\sim -N \left( -2 - \log \left( \frac{a^2}{2\kappa^2} \right) + \log N \right)
\]

\[
\sim -N \log N
\]

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.

**Remark**: Note that if the total power scales linearly with the number of sensors, then a diversity order of \( N \log N \) for orthogonal MAC can also be achieved [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.

### B. Short-term optimal power allocation

We first give the expression of distortion using ST-OPA. Substituting (4) into (2) gives

\[
D = \left( \frac{1}{\sigma_\theta^2} + \frac{\left( \sum_{i=1}^{N} \sqrt{h_iP_i^*} \right)^2}{\sigma^2 \sum_{i=1}^{N} h_iP_i^* + \sigma_c^2 C} \right)^{-1} = \frac{\sigma_\theta^2 \sigma^2}{\sigma^2 + \sigma_g^2 \sum_{i=1}^{N} Z_i}
\]

(37)

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

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

(38)

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

Denote \( Z \) as the random variable distributed according to the common distribution of \( Z_i \). We now apply Theorem 2 to (38). 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 (15). Then

$$\Pr \left( \frac{1}{n} \sum_{i=1}^{n} \tilde{Y}_{n,i} > E \left[ \tilde{Y}_{n,i} \right] \right) \to 0.5 \quad \text{as} \quad n \to \infty \quad (39)$$

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

Applying Theorem 2 to (38) we have

$$\log P_{outage} \sim -NI_Z(g_N) \quad \text{as} \quad N \to \infty \quad (40)$$

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 $Z_i = \left( 1 + \frac{\nu}{h_i} \right)^{-1}$, where $h_i$ are i.i.d. exponentially distributed random variables with parameter $\lambda$. Note that the domain of $Z_i$ is $[0,1)$. The c.d.f. and p.d.f. of $Z$ are given by $F_Z(z) = 1 - e^{-\frac{\lambda z}{1-z}}$ and $f_Z(z) = \lambda z e^{-\frac{\lambda z}{1-z}}$ respectively. The mean of $Z$ is given as $\mu_Z = 1 - \lambda \rho e^{\lambda \rho} E_1(\lambda \rho)$, where $E_1(x) = \int_x^{\infty} \frac{e^{-z}}{z} \, dz$ is the exponential integral. The moment generating function of $Z$ is given as $M_Z(\theta) = \lambda \rho \int_0^{1} \frac{1}{(1-z)^2} e^{\theta z - \lambda \rho \frac{z}{1-z}} \, dz$.

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

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

$$\Rightarrow g_N = \frac{\int_{0}^{1} z g(z,t) \, dz}{\int_{0}^{1} g(z,t) \, dz} \quad (41)$$

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

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 $\theta$ by $-t$ and taking the derivative of $\varphi(-t)$ w.r.t. $t$ yields $\frac{d\varphi(-t)}{dt} = -\Lambda_Z''(-t) \leq 0$, where the inequality arises due to the cumulant generating function being a convex function. Hence $\varphi(-t)$ is a continuous non-increasing function of $t$ (the continuity of $\varphi(-t)$ is evident by inspecting the RHS of (42)). Hence large $N$ corresponds to the case of large $t$ in (42). Let $x = 1/(1-z)$. It can be easily shown that (42) can be written as

$$g_N = 1 - \frac{\int_{1}^{\infty} \frac{1}{x} e^{-cx} \, dx}{\int_{1}^{\infty} e^{-cx} \, dx} \quad (43)$$

where $c = \lambda \rho$.

**Lemma 4.3:**

$$g_N \sim \frac{1}{t} \quad \text{as} \quad t \to \infty. \quad (44)$$

**Proof:** See Appendix.
Hence for large $N$, we have

$$\theta \sim -\frac{1}{g_N}. \quad (45)$$

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

$$M_Z(\theta) = \lambda \rho e^{-t+c} \int_1^\infty e^{-tp(x)}q(x)dx \sim \lambda \rho e^{-t+c} \frac{e^{t-c}}{t} \sim \lambda \rho g_N \quad (46)$$

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

$$I_Z(a_N) \sim -g_N \frac{1}{g_N} - \log \left( \frac{\lambda \rho g}{N} \right) \quad \text{for large } N \quad (47)$$

$$= -1 - \log (\lambda \rho g) + \log N. \quad (48)$$

Hence from (40) the outage probability for large $N$ is asymptotically

$$\log P_{\text{outage}} \sim -NI_Z(g_N) \quad (49)$$

$$\sim -N(-1 - \log (\lambda \rho g) + \log N) \quad (50)$$

$$\sim -N \log N. \quad (51)$$

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 $N$.

We first analyze the power required to achieve zero outage. For $N = 1$, the sum power expression in (8) can be re-arranged and expressed as $P_{\text{tot}}(h) = \frac{K_1}{h}$ where $K_1 = \frac{\gamma_1 \sigma_2^2 C}{(1-\sigma_1^2 \gamma_1)}$. The region $\mathcal{R}_T(t)$ can be easily found directly from the definition as $\mathcal{R}_T(t) = \{h : P_{\text{tot}}(h) < t\} = \{h : h > \frac{K_1}{t}\}$. The average power sum, $P_T(t)$, becomes

$$P_T(t) = \int_{\mathcal{R}_T(t)} P_{\text{tot}}(h)dF(h) = \int_{\frac{K_1}{t}}^\infty \frac{K_1}{h} \lambda e^{-\lambda h} dh = \lambda K_1 \int_{\frac{K_1}{t}}^{\infty} \frac{e^{-u}}{u} du \quad (52)$$

$$= \lambda K_1 E_1 \left( \frac{\lambda K_1}{t} \right) \quad (53)$$

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 \( P_{\text{outage}} = \Pr(h \notin \mathcal{R}_T(t)) \). 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 \( \hat{P}(h_{\text{max}}) \) where \( h_{\text{max}} = \max(h_1, \ldots, h_N) \). \( \hat{P}(h) \) can be obtained from the distortion constraint and it is given as \( \hat{P}(h_{\text{max}}) = \frac{\gamma_{\text{th}} \sigma^2_C}{(1 - \sigma^2_{\text{th}})h_{\text{max}}} \). 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 \( f \) is concave in \( \gamma_{\text{th}} \). It is straightforward to show that power is proportional to the inverse of the channel realization and over time is then given as

\[
E[\hat{P}(h_{\text{max}})] = \int_0^\infty \frac{\gamma_{\text{th}} \sigma^2_C}{(1 - \sigma^2_{\text{th}})h} \cdot N \lambda \left(1 - e^{-\lambda h}\right)^{N-1} e^{-\lambda h} dh.
\]

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

We now proceed to find an approximation for the maximum number of sensors \( N_{\text{max}} \) that still has non-zero outage for a given \( P_{\text{tot}} \) for LT-OPA. Then \( N_{\text{max}} + 1 \) 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 \( P_{\text{tot}}(h) \). We begin with the equation we need to solve to obtain \( P_{\text{tot}}(h) \), given as \( \sigma^2 \gamma_{\text{th}} = \sum_{i=1}^N \left( \frac{\sigma^2_C}{\sigma^2 P_{\text{tot}}(h) h_i} + 1 \right)^{-1} \). Let \( f(h_i) = \left( \frac{\sigma^2_C}{\sigma^2 P_{\text{tot}}(h) h_i} + 1 \right)^{-1} \). It is straightforward to show that \( f \) is concave in \( h_i \) \( \forall i \). Applying Jensen’s inequality we have

\[
\sigma^2 \gamma_{\text{th}} = \frac{\sum_{i=1}^N f(h_i)}{N} \leq f \left( \frac{\sum_{i=1}^N h_i}{N} \right) = \frac{\sigma^2 \gamma_{\text{th}}}{N} \leq \frac{1}{\sigma^2 P_{\text{tot}}(h) \left( \sum_{i=1}^N h_i \right) + 1} \tag{55}
\]

\[
\Rightarrow \frac{\sigma^2 \gamma_{\text{th}}}{N} \leq \frac{\sigma^2 C}{\sigma^2 P_{\text{tot}}(h) \frac{1}{N} \sum_{i=1}^N h_i} \leq 1 - \frac{\sigma^2 \gamma_{\text{th}}}{N} \tag{56}
\]

\[
\Rightarrow P_{\text{tot}}(h) \geq \frac{\frac{1}{N} \sum_{i=1}^N h_i}{K_N} \tag{57}
\]

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

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

$$\mathcal{R}_T(\bar{\ell}) = \left\{ h : \tilde{P}_{\text{tot}}(h) < \bar{\ell} \right\} = \left\{ h : \sum_{i=1}^{N} h_i > \frac{K_N}{\ell} \right\}. \quad (58)$$

The definition of $P_T(\bar{\ell})$ becomes

$$P_T(\bar{\ell}) = \int_{\mathcal{R}_T(\bar{\ell})} \tilde{P}_{\text{tot}}(h) dF(h) = K_T \int_{\sum_{i=1}^{N} h_i > \frac{K_N}{\ell}} \frac{1}{\lambda} \sum_{i=1}^{N} h_i \cdot dh_1 \cdots dh_N. \quad (59)$$

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

$$P_T(\bar{\ell}) = K_N \int_{\sum_{i=1}^{N} h_i > \frac{K_N}{\ell}} \frac{1}{\lambda} \sum_{i=1}^{N} h_i \cdot dh_1 \cdots dh_N = K_N \frac{1}{\Gamma(k)\theta^k} \int_{\frac{K_N}{\ell}}^\infty T^{k-2} e^{-\frac{T}{\theta}} dT \quad (60)$$

$$= \frac{K_N}{\Gamma(N)\lambda^{-n}} \int_{\frac{K_N}{\ell}}^\infty T^{N-2} e^{-\lambda T} dT = \frac{K_N^N \lambda^N}{N-1} \cdot \frac{\Gamma(N-1, \lambda K_N/\bar{\ell})}{\Gamma(N-1)}. \quad (61)$$

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

$$\frac{K_N^N \lambda^N}{N-1} \cdot \frac{\Gamma(N-1, \lambda K_N/\bar{\ell}^*)}{\Gamma(N-1)} = \mathcal{P}_{\text{tot}}. \quad (62)$$

The outage event becomes $P_{\text{outage}} = \left\{ h : \tilde{P}_{\text{tot}}(h) > \bar{\ell}^* \right\} = \left\{ h : \frac{1}{N} \sum_{i=1}^{N} h_i < \frac{K_N}{\bar{\ell}^*} \right\}$. If we let $\bar{\ell}^* \to \infty$ in (62) for a given finite $N$ then $K_N/\bar{\ell}^* \to 0$, $\frac{\Gamma(N-1, \lambda K_N/\bar{\ell}^*)}{\Gamma(N-1)} \to 1$ and

$$\frac{K_N \lambda}{N-1} = \mathcal{P}_{\text{tot}}. \quad (63)$$

Equation (63) allows us to solve for $N$, and it gives an approximation $\tilde{N}_{\text{max}}$ to the maximum number of sensors that has non-zero outage probability for a given $\mathcal{P}_{\text{tot}}$. The solution of (63) can be found in closed-form and is given as

$$\tilde{N}_{\text{max}} = \left\lceil \frac{(1 + \sigma^2 \chi^2) P_{\text{tot}} + \gamma \sigma^2 C \lambda + \sqrt{[(1 + \sigma^2 \chi^2) P_{\text{tot}} + \gamma \sigma^2 C \lambda]^2 - 4 P_{\text{tot}}^2 \chi^2 \sigma^2}}{2 P_{\text{tot}}} \right\rceil \quad (64)$$

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

V. SIMULATION RESULTS

The following results, if not computed directly from the equations, are obtained via Monte Carlo simulation over 1,000,000 channel realizations. We first present the diversity order of distortion outage for EPA. We simulated the case where $\mathcal{P}_{\text{tot}} = 10 \text{mW}$ and plotted the results in Fig. 2. The lines plotted in plus signs shown are plots of $\log P_{\text{outage}}$ obtained via Monte Carlo simulation, where $\log$ is the natural log. The lines plotted in triangles are the exact values of $-NI_0(a/\sqrt{N})$ where the values of $I_0(a/\sqrt{N})$ are obtained by solving (19) numerically. The squares are plots of (35). The figure shows that as $N$ gets large, the asymptotic expression
Fig. 2. EPA with $P_{tot} = 10\text{mW}$. Squares: (35) against $N$. Triangles: $-NI_{\sqrt{h}}(a/\sqrt{N})$ against $N$. Plus signs: $\log P_{\text{outage}}$ from Monte Carlo simulation. Simulation parameters: $\sigma = 0.0014$, $a = 0.003$, $\sigma_\theta^2 = 10^{-3}$, $\sigma_c^2 = 10^{-8}$, $D_{\text{max}} = 0.1$.

(35) converges to $-NI_{\sqrt{h}}(a/\sqrt{N})$. Note that the asymptotic results $I_{\sqrt{h}}(a/\sqrt{N})$ and (35) only give us the slope of the outage probability when plotted on a log scale; these two lines may not necessarily converge to $\log P_{\text{outage}}$ but their gradients should coincide for large $N$, as can be seen in Fig. 2.

We now look at ST-OPA. Fig. 3 shows the log of the outage probability using ST-OPA as a function of $N$ in circles and $-NI_Z(g)/N$ obtained numerically in squares for $P_{tot} = 10\text{mW}$. It shows that $-NI_Z(g_n)$ gives a similar gradient as $\log P_{\text{outage}}$. In Fig. 3 we also show the asymptotic expression of $-NI_Z(g_n)$ plotted in plus signs. We see that as $N$ increases, the asymptotic expression gives very similar gradients as $-NI_Z(g_n)$.

With the long-term OPA, we first present the relation between $t^*$ and $P_{tot}$ for a fixed total power constraint shown in Fig. 4. The circles and squares are obtained via Monte Carlo simulation. The solid lines on the graph are obtained by solving (62) numerically. We see that the results match closely. Note that $P_{tot}$ is a monotonically increasing function of $t^*$ for any fixed $N$, and as $t^*$ gets large, the value of $P_{tot}$ saturates and approaches some asymptotic value. The saturation behavior is due to $\Gamma(N-1,\lambda \zeta) / \Gamma(N-1) \rightarrow 1$ as $t \rightarrow \infty$ for fixed $N$, implying the existence of a finite $P_{tot}$ achieving zero outage.
Fig. 4. $P_{tot}$ versus $t^*$. Circles and squares: from Monte Carlo simulation with 1,000,000 channel realizations. Solid lines: numerical solution of (62). Simulation parameters: $\sigma_2^2 = 1$, $\sigma_2^2 = 10^{-3}$, $\sigma_2^2 = 10^{-8}$, $D_{max} = 0.1$, $\lambda = 250,000$.

Fig. 5. $N_{max}$ versus $P_{tot}$. Circles: (64) against $P_{tot}$. Solid line: $N_{max}$ from Monte Carlo simulation. Simulation parameters: $\sigma_0^2 = 1$, $\sigma_2^2 = 10^{-3}$, $\sigma_2^2 = 10^{-8}$, $D_{max} = 0.1$ and $\lambda = 250,000$.

An approximate relationship between $N_{max}$ and $P_{tot}$ for LT-OPA has been obtained in (64). To see how good the approximation is, we plot (64) together with $N_{max}$ obtained via Monte Carlo simulation, where we compute $E[\langle P_{tot}(h) \rangle]$ for a given $N$ over 1,000,000 channel realizations. The results are shown in Fig. 5.

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 $P_{tot} = 1,600\mu W$. Note that for LT-OPA, due to the existence of $N_{max}$, the outage probability for $N > N_{max}$ is zero and hence we cannot show results for $N > N_{max}$ on the graph. In this example, $N_{max} = 15$. From this figure we can see that the gradients of EPA and ST-OPA are similar for large $N$, while the outage probability curve for LT-OPA approaches a vertical asymptote located at $N_{max}$.

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

VII. APPENDIX

Proof: Lemma 3.2: In the first statement it is immediate to see that $P^*(h)$ is a continuous function of $h$. In the second statement we need to show that $\langle P(h) \rangle$ 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

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

(65)

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

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

(66)
Taking the partial derivative w.r.t. $h_i$ on both sides of (66) gives

\[
\frac{\partial}{\partial h_i} \sum_{j=1}^{N} \frac{\nu h_j}{C + \nu h_j \sigma^2} = 0 \Rightarrow \frac{\partial}{\partial h_i} \frac{\nu h_i}{C + \nu h_i \sigma^2} + \sum_{j=1}^{N} \frac{\partial}{\partial \nu} \frac{\nu h_j}{C + \nu h_j \sigma^2} \frac{\partial \nu}{\partial h_i} = 0
\]

\[
\Rightarrow \frac{\nu C}{(C + \nu h_i \sigma^2)^2} + \sum_{j=1}^{N} \frac{C h_j}{(C + \nu h_j \sigma^2)^2} \frac{\partial \nu}{\partial h_i} = 0
\]

\[
\Rightarrow \frac{\partial \nu}{\partial h_i} = - \frac{\nu C}{(C + \nu h_i \sigma^2)^2} \left( \sum_{j=1}^{N} \frac{C h_j}{(C + \nu h_j \sigma^2)^2} \right)^{-1} < 0 \Rightarrow \frac{\partial \mathbb{P}(\mathbf{h})}{\partial h_i} = \frac{\sigma_n^2}{N} \frac{\partial \nu}{\partial h_i} < 0
\]

which completes the proof.

**Proof: Theorem 2:** We prove the theorem by obtaining upper and lower bounds of $\log \Pr \left( \frac{1}{n} \sum_{i=1}^{n} X_i \leq a_n \right)$, which asymptotically are equivalent for large $n$. The proof uses similar techniques to those provided in the proof of Theorem 5.11.4 in [31].

**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 $\mu_X$ as the mean of $X_i$. Let $Y_i = -X_i + \mu_X$, hence $E[Y_i] = \mu_Y = 0$. The transformation allows us to obtain the following relationships $M_Y(t) = e^{\mu_X t} M_X(-t)$, $\Lambda_Y(t) = \mu_X t + \Lambda_X(-t)$ and

\[
I_Y(c_n) = \sup_{t \rightarrow -} \{ (\mu_X - c_n) t - \Lambda_X(t) \}.
\]

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

We make two notes for future use. First, since $\Lambda_Y(t)$ is convex with $\Lambda_Y'(0) = E[Y] = 0$, and since $c_n > \mu_Y = 0$ for $n \geq N$ (the value of $N$ 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 $t \in \mathbb{R}$ is unchanged by the restriction $t > 0$, which is to say that

\[
I_Y(c_n) = \sup_{t>0} \{ c_n t - \Lambda_Y(t) \}, \quad c_n > 0 \text{ for } n \geq N.
\]

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

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

where the inequality is due to the Cauchy-Schwartz inequality applied to the random variables $Y e^{\frac{1}{2} tY}$ and $e^{\frac{1}{2} tY}$.
We have the following
\[
\Pr \left( \frac{1}{n} \sum_{i=1}^{n} X_i \leq a_n \right) = \Pr \left( \sum_{i=1}^{n} Y_i \geq c_n \right) = \Pr \left( e^{\sum_{i=1}^{n} Y_i} \geq e^{nc_n t} \right) \quad \text{for } t > 0
\] (70)
\[
\leq \frac{E[\exp \left( t \sum_{i=1}^{n} Y_i \right)]}{e^{nc_n t}} = e^{-nc_n t} \int Y^n(t) = e^{-n(c_n t - \Lambda_Y(t))}
\] (71)
where the inequality is due to Markov’s inequality. Taking log on both sides gives
\[
\log \Pr \left( \frac{1}{n} \sum_{i=1}^{n} X_i \leq a_n \right) \leq -n \left( c_n t - \Lambda_Y(t) \right) \quad \forall t > 0
\] (72)
Since the upper bound in (72) is true for all \( t > 0 \) and we are looking for the tightest bound, we can further bound the LHS by taking the infimum on the RHS
\[
\log \Pr \left( \frac{1}{n} \sum_{i=1}^{n} X_i \leq a_n \right) \leq \inf_{t > 0} \left\{ -n \left( c_n t - \Lambda_Y(t) \right) \right\} = -n \sup_{t > 0} \left\{ c_n t - \Lambda_Y(t) \right\}
\] (73)
\[
=-n I^+_Y(c_n) = -n I^-_X(a_n) \quad \text{from (67)}
\] (74)

**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 \( \Pr (Y_i > c_n) > 0 \) 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 \( t > 0 \),
\[
c_n t - \Lambda_Y(t) = c_n t - \log E \left[ e^{tY} \right] = c_n t - \log \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy
\] (75)
\[
= c_n t - \log \left\{ \int_{-\infty}^{b_n} e^{ty} f_Y(y) dy + \int_{b_n}^{\infty} e^{ty} f_Y(y) dy \right\}
\] (76)
\[
\leq c_n t - \log \int_{b_n}^{\infty} e^{ty} f_Y(y) dy \leq c_n t - \log \left\{ e^{t b_n} \int_{b_n}^{\infty} f_Y(y) dy \right\}
\] (77)
\[
= c_n t - \log \left\{ e^{t b_n} \Pr(Y_i > b_n) \right\}
\] (78)
\[
= - (b_n - c_n) t - \log \Pr(Y_i > b_n) \to -\infty \quad \text{as } t \to \infty
\] (79)
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
\[
dF_{\tilde{Y}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)} dF_Y(y)
\] (80)
which can also be interpreted as \( F_{\tilde{Y}_n}(y) = \frac{1}{M_Y(\tau_n)} \int_{-\infty}^{y} e^{\tau_n u} dF_Y(u) \). Let \( \tilde{Y}_{n,1}, \tilde{Y}_{n,2}, \ldots \) be i.i.d. distributed with
c.d.f. \( F_{\tilde{Y}_n} \). We note the following properties of \( \tilde{Y}_{n,i} \). The moment generating function of \( \tilde{Y}_{n,i} \) is

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

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

\[
E[\tilde{Y}_{n,i}] = M'_{\tilde{Y}_n}(0) = \frac{M_Y'(\tau_n)}{M_Y(\tau_n)} = \Lambda_Y(\tau_n) = c_n,
\]

\[
\text{var}(\tilde{Y}_{n,i}) = E \left[ (\tilde{Y}_{n,i} - E[\tilde{Y}_{n,i}])^2 \right] - (E[\tilde{Y}_{n,i}])^2 = M''_{\tilde{Y}_n}(0) - M'_Y(0)^2 = \Lambda''_Y(\tau_n) \in (0, \infty).
\]

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

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

where \( F_{\tilde{S}_n} \) 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

\[
dF_{\tilde{S}_n}(y) = \frac{e^{\tau_n y}}{M_Y(\tau_n)^n}dF_{S_n}(y).
\]

Let \( d > 0 \). We have

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

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

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

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

Since \( E[\tilde{Y}_{n,i}] = c_n \) and \( \text{var}(\tilde{Y}_{n,i}) > 0 \), 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 [32]. Therefore,

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

\[
\sim -n (\tau_n(c_n+d) - \Lambda_Y(\tau_n)) \quad \text{as} \quad n \to \infty
\]

\[
\sim -n (\tau_n c_n - \Lambda_Y(\tau_n)) \quad \text{as} \quad d \to 0
\]

\[
= -n \Lambda_Y(c_n) = -n I_X(a_n)
\]

\[
\blacksquare
\]
Proof: Lemma 4.1: Here we want to show that
\[ \Pr \left( \frac{1}{n} \tilde{S}_n > c_n \right) \to 0.5 \] (91)
as \( n \to \infty \). We note that the L.H.S. of (91) 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 \( \tilde{Z}_n = \tilde{Y}_n - E[\tilde{Y}_n] \). Hence \( E[\tilde{Z}_n] = 0 \) and \( \text{var}(\tilde{Z}_n) = \text{var}(\tilde{Y}_n) \). Note also that \( E[\tilde{Y}_n] = c_n \) and \( \text{var}(\tilde{Y}_n) = \Lambda_Y'(\tau_n) \). Lindeberg’s condition is hence given as
\[
\frac{1}{\sigma_{\tilde{Z}_n}^2} \int_{\left\{ |\tilde{Z}_n| > \epsilon \sqrt{n} \sigma_{\tilde{Z}_n} \right\}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} \to 0 \quad \text{as} \quad n \to \infty \quad (92)
\]
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 \( \tau_n \). 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 \( \text{var}(\tilde{Y}_n) \) as \( n \to \infty \) for the Rayleigh distribution. We have the following results:
\[
\frac{d}{d\theta} \Lambda_Y(\theta) = \mu_X + \frac{1}{M_{\sqrt{\varphi}(\theta)}} \frac{dM_{\sqrt{\varphi}(\theta)}}{d\theta}
\]
\[
\frac{d^2}{d\theta^2} \Lambda_Y(\theta) = \frac{d^2 M_{\sqrt{\varphi}(\theta)}}{d\theta^2} M_{\sqrt{\varphi}(\theta)} - \left( \frac{dM_{\sqrt{\varphi}(\theta)}}{d\theta} \right)^2 \left( M_{\sqrt{\varphi}(\theta)} \right)^2 \quad (93)
\]
Note that
\[
\frac{dM_{\sqrt{\varphi}(\theta)}}{d\theta} = \left( \kappa^2 \theta^2 + 1 \right) M_{\sqrt{\varphi}(\theta)} - 1 \quad \text{and} \quad \frac{d^2 M_{\sqrt{\varphi}(\theta)}}{d\theta^2} = \kappa^2 \left[ \left( \kappa^2 \theta^2 + 3 \right) M_{\sqrt{\varphi}(\theta)} - 1 \right] \quad (94)
\]
Substituting (94) and (95) into (93) gives
\[
\frac{d^2}{d\theta^2} \Lambda_Y(\theta) = \frac{(\kappa^2 \theta^2 - 1) M_{\sqrt{\varphi}(\theta)}^2 + (\kappa^2 \theta^2 + 2) M_{\sqrt{\varphi}(\theta)} - 1}{\theta^2 M_{\sqrt{\varphi}(\theta)}} \quad (96)
\]
Using the asymptotic expansion of \( M_{\sqrt{\varphi}(\theta)} \) (since \( \theta \to \infty \) as \( n \to \infty \))
\[
M_{\sqrt{\varphi}(\theta)} = \frac{1}{\kappa^2 \theta^2} - \frac{3}{(\kappa^2 \theta^2)^2} + \frac{15}{(\kappa^2 \theta^2)^3} - \cdots
\]
we obtain \( \frac{d^2}{d\theta^2} \Lambda_Y(\theta) \sim \frac{2}{\theta^2} \) and hence
\[
\text{var}(\tilde{Y}_n) = \Lambda_Y'(\tau_n) \sim \frac{a^2}{2n} \quad (97)
\]
The expression of \( f_{\tilde{Z}_n}(\tilde{z}) \) can be easily found and is given as
\[
f_{\tilde{Z}_n}(\tilde{z}) = \frac{a_n - \tilde{z}}{\kappa^2 M_Y(\tau_n)} e^{-\frac{(a_n - \tilde{z})^2}{2\kappa^2}} + \tau_n(\tilde{z} + c_n)
\] (98)

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

We are now ready to look at Lindeberg’s condition (92) after obtaining the expressions (97) and (98). We have
\[
\frac{1}{\sigma_{\tilde{Z}_n}^2} \int_{\{|\tilde{z}_n| < \sqrt{n\sigma^2_\tilde{Z}_n} \}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} = \frac{1}{\sigma_{\tilde{Z}_n}^2} \left( \int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_\tilde{Z}_n}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} + \int_{c_n}^{a_n} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} \right)
\]
\[
\approx \frac{1}{\sigma_{\tilde{Z}_n}^2} \int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_\tilde{Z}_n}} \tilde{z}^2 f_{\tilde{Z}_n}(\tilde{z}) d\tilde{z} = \frac{1}{\kappa^2 \sigma_{Z_n}^2 M_Y(\tau_n)} \int_{-\infty}^{-\epsilon\sqrt{n\sigma^2_\tilde{Z}_n}} (a_n - \tilde{z})^2 e^{-\frac{(a_n - \tilde{z})^2}{2\kappa^2}} + \tau_n(\tilde{z} + c_n) d\tilde{z}
\]
\[
= \frac{e^{-\mu\pi\tau_n}}{\kappa^2 \sigma_{Z_n}^2 M_Y(\tau_n)} \int_{c_n}^{\infty} (a_n + u) u^2 e^{-\frac{(a_n + u)^2}{4\kappa^2}} - \tau_n u + \tau_n c_n - \tau_n \mu \pi d\mu
\]
\[
\leq \frac{1}{\kappa^2 \sigma_{Z_n}^2 M_Y(\tau_n)} \int_{c_n}^{\infty} (a_n + u) u^2 e^{-\frac{(a_n + u)^2}{4\kappa^2}} - \tau_n u + \tau_n c_n - \tau_n \mu \pi d\mu
\]
where \( \mu_{\sqrt{n}} = E[\sqrt{n}] \), \( u = -\tilde{z} \) and step (a) is due to the second integral vanishing to zero as the integration interval becomes null, since \( a_n \to 0 \) and \( \epsilon\sqrt{n\sigma^2_\tilde{Z}_n} \to \alpha \sqrt{2} \) as \( n \to \infty \). Also note that we have the following asymptotic expressions (as \( n \to \infty \))
\[
a_n = a / \sqrt{n} \to 0
\]
\[
c_n = \mu_{\sqrt{n}} - a_n \to 0
\]
\[
\tau_n \sim 2\sqrt{n}
\] (from (30) and \( \tau = -\theta \))
\[
M(\sqrt{n})(-\tau_n) \sim \frac{a^2}{4\kappa^2 n}
\]
\[
\sigma_{\tilde{Z}_n}^2 \sim \frac{a^2}{2n}.
\]

We now show that (102) goes to zero as \( n \to \infty \) by using an upper bound of (102) 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 (102): 1) \( e^{-\frac{(a_n + u)^2}{4\kappa^2}} \leq e^{-\frac{u^2}{2\kappa^2}} \) (since \( a_n > 0 \)), 2) \( e^{\tau_n c_n - \tau_n \mu \pi} = e^{\tau_n \mu \pi - \tau_n a_n - \tau_n \mu \pi} = e^{-\tau_n a_n} = O(1) \) (from (103) and (105)) \( \Rightarrow e^{\tau_n c_n - \tau_n \mu \pi} \leq C \) (for sufficiently large \( n \)), 3) \( e^{-\tau_n u} = e^{-\frac{\sqrt{2\kappa^2 a}{\alpha}u(1+o(1))} \leq e^{-\frac{\sqrt{2\kappa^2 a}}{\alpha}u} \) (for sufficiently large \( n \)), where \( C \) is a constant.

Hence we substitute the upper bounds obtained above and the asymptotic expressions (103), (106) and (107) into (102) and obtain the following upper bound
\[
\frac{1}{\kappa^2 \sigma_{Z_n}^2 M_Y(\tau_n)} \int_{c_n}^{\infty} (a_n + u) u^2 e^{-\frac{(a_n + u)^2}{4\kappa^2}} - \tau_n u + \tau_n c_n - \tau_n \mu \pi d\mu
\]
\[
\leq 8C n^2 \frac{a^4}{a^4} \int_{\alpha \sqrt{2}}^{\infty} u^2 e^{-\frac{u^2}{2\kappa^2}} e^{-\frac{u^2}{2\kappa^2} \alpha} d\alpha (1 + o(1))
\]
We may use Laplace’s method [33] to obtain an asymptotic approximation of
\[ I(\sqrt{n}) = \int_{a\sqrt{2}/\mu}^{\infty} u^3 e^{-u^2} e^{-\sqrt{n}u^2} du \] (110)
in (109). Let \( h(u) = u/a \) and \( \varphi(u) = u^3 e^{-u^2} \). Hence the integral becomes
\[ I(\sqrt{n}) = \int_{A}^{\infty} \varphi(u)e^{-\sqrt{n}h(u)} du \] (111)
where \( A = a\sqrt{2}/\mu \). It is straightforward to see that \( h(u) \) and \( \varphi(u) \) satisfy the four conditions necessary for using Laplace’s method. The Taylor series for \( h(u) \) and \( \varphi(u) \) as \( u \to A \) are given as \( h(u) \sim h(A) + \sum_{s=0}^\infty a_s(u-A)^{s+\mu}, \varphi(u) \sim \sum_{s=0}^\infty b_s(u-A)^{s+\alpha-1} \). We give the values of the first few terms of the series:
\[ h(A) = \epsilon/\sqrt{2}, a_0 = 1/\alpha, a_1 = 0, \mu = 1, \alpha = 1, b_0 = \left(\frac{a_0}{\sqrt{2}}\right)^3 \exp\left(-\frac{(\alpha x)^2}{4\epsilon^2}\right) \]. The asymptotic approximation of \( I \) is given as
\[ I(\sqrt{n}) \sim e^{-\sqrt{n}h(A)} \sum_{s=0}^\infty \Gamma\left(\frac{s+\alpha}{\mu}\right) \frac{c_s}{\sqrt{n}^{s+\alpha}/\mu} \sim \frac{c_0}{\sqrt{n}} e^{-c\sqrt{n}} \] (112)
where we have simply retained the first term in the sum. Note that \( c_0 = a \left(\frac{a}{\sqrt{2}}\right)^3 \exp\left(-\frac{(\alpha x)^2}{4\epsilon^2}\right) \). Hence Lindeberg’s condition becomes
\[ \frac{8c_0n\sqrt{n}}{a^4} e^{-c\sqrt{n}} \to 0 \quad \text{as} \quad n \to \infty. \] (113)
This completes the proof for showing that the CLT holds for \( \sqrt{T_n} \).

**Proof:** Lemma 4.3: Let \( g(t) = \int_1^\infty e^{t-x}e^{\mu}dx \) and \( h(t) = \int_1^\infty \frac{1}{x} e^{t-x}dx \). We use Laplace’s method [33] to obtain asymptotic approximations of \( g(t) \) and \( h(t) \). We begin by writing \( g(t) \) and \( h(t) \) as \( g(t) = \int_1^\infty e^{t-p(x)}q(x)dx \) and \( h(t) = \int_1^\infty e^{t-p(x)}\phi(x)dx \) where \( p(x) = -1/x, q(x) = e^{-c\epsilon} \) and \( \phi(x) = e^{-c\epsilon} \). In order to apply Laplace’s method, we must check four conditions (Theorem 1 in Ch 2 of [33]). 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 \( g(t) \sim e^{-t-p(1)} \sum_{s=0}^\infty \Gamma\left(\frac{s+\alpha}{\mu}\right) \frac{c_s}{e^{t-s+\alpha}/\mu} \sim e^{t} \left( e^{c-x} + \frac{(2-c)e^{-c}}{t^2} \right) \) where we have simply retained the first two terms in the sum. Note that \( c_0 = b_0 \mu/a^3 \) and \( c_1 = \left\{ b_1 - \frac{(a+1)a_1b_0}{a^2a_0} \right\} \frac{1}{a^2a_3} \) [33]. 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 (43) gives

\[
 g_N = 1 - \frac{h(t)}{g(t)} \sim 1 - \frac{e^t \left( \frac{e^{-c}}{t} + \frac{(1-c)e^{-c}}{t^2} \right)}{e^t \left( \frac{e^{-c}}{t} + \frac{(2-c)e^{-c}}{t^2} \right)} = \frac{1}{t + 2 - c} \sim \frac{1}{t} \quad \text{for large } t
\] (114)

which completes the proof.

REFERENCES


