# Power Allocation for Distributed Detection Systems in Wireless Sensor Networks With Limited Fusion Center Feedback

Xiaoxi Guo, *Student Member, IEEE*, Yuanyuan He, *Member, IEEE*, Saman Atapattu, *Member, IEEE*, Subhrakanti Dey, *Senior Member, IEEE*, and Jamie S. Evans, *Senior Member, IEEE*

*Abstract*—We consider a distributed detection system for a wireless sensor network over slow-fading channels. Each sensor only has knowledge of quantized channel state information (CSI) which is received from the fusion center via a limited feedback channel. We then consider transmit power allocation at each sensor in order to maximize a J-divergence based detection metric subject to a total and individual transmit power constraints. Our aim is to jointly design the quantization regions of all sensors CSI and the corresponding power allocations. A locally optimum solution is obtained by applying the generalized Lloyd algorithm (GLA). To overcome the high computational complexity of the GLA, we then propose a low-complexity near-optimal scheme which performs very close to its GLA based counterpart. This enables us to explicitly formulate the problem and to find the unique solution despite the non-convexity of the optimization problem. An asymptotic analysis is also provided when the number of feedback bits becomes large. Numerical results illustrate that only a small amount of feedback is needed to achieve a detection performance close to the full CSI case.

*Index Terms*—Distributed detection, wireless sensor networks, quantized power allocation, J-divergence, finite-rate feedback.

## I. INTRODUCTION

**W**IRELESS sensor networks (WSNs) have captured lots of attention recently. Due to low cost, robustness, and high flexibility, WSNs are widely used in current event monitoring systems. For example, small-scale, inexpensive, and low-power sensors can be deployed over a region to collect and possibly pre-process data about a particular physical phenomenon. With the fifth generation (5G) wireless networks being deployed in near future, WSNs will form a key technology and contribute to essential building blocks for emerging Internet of Things (IoT) architecture which offers a new paradigm to connect all kinds of physical objects to the Internet and provides important services to every aspect of our lives, e.g., transportation, energy, smart city and health-care [1], [2]. We adopt a WSN model that consists of a fusion center (FC) that can communicate with each sensor. Observations or local decisions from sensors are transmitted through wireless channels to the FC which can jointly process the data and make a global decision. Typically, sensors which may be geographically wide spread are powered by small batteries. Since the available energy of network is limited, WSNs employ distributed detection systems from which only the observation summary of each sensor is sent to the FC, instead of sending all collected data.

### A. Related Work

The design of a distributed detection system typically involves the design of the decision rules at local sensors as well as the fusion rule at the FC. In the literature, these rules are applied based on different objectives. The idea of 'send/no-send' has been proposed by considering the local likelihood ratio in [3]. While minimizing the error probability, sensors can choose to transmit data or keep silent to save energy. Based on other performance criteria such as Neyman-Pearson (N-P) or Ali-Silvey (A-S) distance metric, local decision strategies have been also investigated in [4]–[6]. Channel aware decision fusion rule design has been studied in [7] and [8]. The problem of optimizing detection performance is investigated for a distributed binary detection system in [9] which optimizes the performance based on J-divergence (one of A-S distance metrics) for individual and total transmitter power constraints on each sensor. These results have also been extended to a multiple-hypothesis scenario in [10]. The idea of distributed WSNs has also been extensively applied towards cooperative detection of primary users in cognitive radio (CR) to significantly improve the spectrum sensing performance. It is known as "Cooperative Spectrum Sensing (CSS)", the performance of which is characterized and evaluated by false alarm and detection probabilities [11]–[15]. In [11], analytical detection performance of an energy detector used for CSS is investigated over a multiple cognitive relay network based on two fusion strategies: data fusion and decision fusion. A new spectrum sensing scheme based on spatial-temporal opportunity detection is proposed

for a spectrum-heterogeneous cognitive wireless sensor network [12]. In [13], Energy Harvesting (EH) capability is incorporated in cognitive wireless sensor networks to enable spectrum and energy efficient heterogeneous cognitive radio sensor networks, whereas in [14], Patel *et al.* proposed schemes for cooperative spectrum sensing in multi-user MIMO cognitive radio networks considering the availability of both perfect CSI and CSI with uncertainty with closed-form analytical expressions of the probability of detection and the probability of false alarm. Recently, multiband cooperative spectrum sensing and resource allocation framework for IoT in cognitive 5G networks is presented in [15], to minimize the energy consumption for spectrum sensing while satisfying probabilities of detection and false alarm requirements.

Most of the existing research assumes that each sensor has full channel state information (CSI) available. However, acquisition of accurate CSI in such wireless scenarios is difficult and imposes unrealistic overheads. Thus, recent studies have considered sensors with no CSI, partial CSI or imperfect CSI. In [16]–[18], the fusion decisions are made based on the statistics of fading. Similarly, local decision rules are also studied with the presence of channel statistics in [19] and [20]. A decentralized detection problem with imperfect estimated CSI is studied in [21]–[23], where the minimum mean square error (MMSE) technique has been employed for channel estimation. In [24], Venkategowda *et al.* investigated precoding schemes for decentralized estimation of scalar and vector parameters in a coherent multiple access channel based WSN with the availability of only imperfect estimated CSI. However, channel estimation based on training may not be suitable for frequency division duplexing (FDD) systems due to the highly uncorrelated forward and reverse channels [25]. The incorporation of feedback from fusion center to each sensors in distributed detection networks, has been considered in the literature over decades [26, Sec. 4.4], [27]–[30]. Introduction of this additional communication capability, where the sensors use this feedback information to adapt their decision rules, improves overall system performance. Since the channel information cannot be fed back with infinite precision, it is natural to investigate the effects of quantization in channel feedback. Therefore, a viable solution is to use a limited feedback technique, where sensors can adapt the transmitted signals according to the received quantized CSI feedback from the fusion center [31]–[37].

In [31], a decentralized detection system has been considered, in which the FC separately quantizes the channel phase of each sensor uniformly, and then transmits the corresponding index to each sensor. Further, a limited feedback scheme has also been investigated in [32] where the FC divides sensors into different priority groups based on the channel quality, and sends only the group indices as the guideline for sensors to perform adaptive transmission. In [33] and [34], power-efficient sensor transmission and quantization schemes are derived based on quantized CSI fed back from the FC to the sensors, while the power allocation codebook is designed independently of sensors' CSI. However, these limited feedback schemes could only provide sub-optimal solutions. To boost the system performance, it is indispensable to jointly design

channels partitions and quantized power allocation schemes in WSN. In [35], Wang and Dey studied such problem and proposed a number of computationally efficient near-optimal power allocation algorithms for minimizing the distortion outage probability of a clustered WSN with quantized channel feedback. Similar to limited feedback strategies proposed in [38] in the context of cognitive radio, [36], [37] jointly designed channels quantization and power codebook for linear distributed estimation of WSN, using the generalized Lloyd algorithm (GLA) with the variance of linear unbiased estimator as distortion functions. But, no rigorous or tractable analysis was provided. To the best of our knowledge, the joint sensor channel quantization and power codebook design problem has not been investigated thoroughly in the context of distributed detection, and hence is the focus of our work.

### B. Problem Statement and Contribution

In this paper, we consider a binary distributed detection system where a finite number of sensors communicate with the FC over slow-fading orthogonal multiple-access channels (MAC). The motivation for orthogonal MAC schemes is to avoid the perfect synchronization between all sensors and the FC, but only require pairwise synchronization between each sensor and the FC. We then study optimal power allocation schemes for sensors, when only quantized CSI is available at sensors. The system performance is evaluated based on the J-divergence metric [9], which is maximized with both individual and total transmitter power constraints at sensors. Further, we focus on a joint design of the optimal quantization regions consisting of a vector space of all sensors' CSI and the corresponding optimal power codebook solved offline at the FC based on the channel statistics of sensors. This codebook is known *a priori* by both the transmitter and receiver. We assume that the FC can perfectly estimate the full CSI of all sensors. In real-time, given a particular channel realization, the FC maps the current instantaneous channel information to one of the quantization regions and then sends the corresponding region index to the sensors via a no-delay and error-free link with finite-rate feedback. Each sensor then chooses the associated power codebook element based on received index to adapt its transmission strategy. Our main contributions of this paper are as follows:

- Despite the difficulties of finding a global optimal solution for the considered optimization problem (non-convex), we locally solve the problem by applying an appropriate Generalized Lloyd Algorithm (GLA) with a Lagrangian based distortion function [38] to jointly design the power codebook and quantization regions.
- The offline design of GLA based quantizer usually requires a large number of training samples coupled with a prolonged convergence time and high computational complexity. To overcome these drawbacks, we then propose a suboptimal solution by making an approximation on channel quantization regions. This enables us to explicitly formulate the problem and find the optimal power codebook with less computational cost. Although the problem is still non-convex after the approximation,
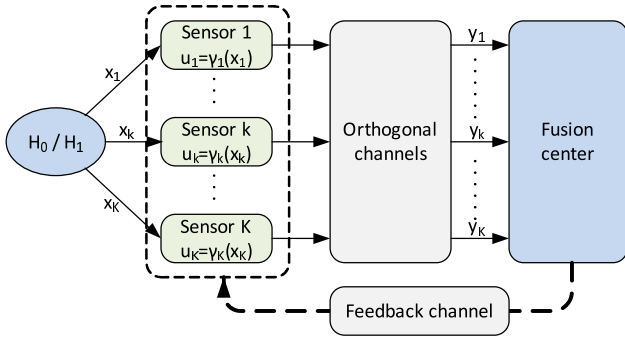
Fig. 1. Distributed detection system with limited feedback.

we are able to find the unique solution, and derive insightful properties of the optimal power codebook for each sensor.

- Asymptotic approximations for the J-divergence when the number of feedback bits grows large are also derived.

The rest of the paper is organized as follows. Section II provides the system model and the problem formulation. Section III develops a locally optimal quantized power allocation scheme based on a Lagrangian based modified GLA, a suboptimal low-complex power allocation scheme, and asymptotic analysis for high resolution quantization. Section IV provides numerical results to compare the proposed algorithms, followed by concluding remarks in Section V.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a distributed binary hypothesis testing task, where $K$ sensors are employed to identify two alternatives hypotheses $H = \{H_0, H_1\}$ where $H_0$ denotes the target-absent hypothesis and $H_1$ represents for target-present hypothesis. More specifically, as shown in Fig. 1 (the use of feedback channel will be discussed in Section III), each Sensor $k$ ($\forall k = 1, 2, ..., K$) collects its own observation $x_k$ and makes a local binary decision $u_k$ independently according to a certain local decision rule on the basis of $x_k$ only. After that, each local decision $u_k$ is sent by each sensor to the FC (who makes the final decision) via an orthogonal channel. We assign orthogonal channels using FDD. Sensor $k$ sends $u_k = 1$ to the FC if $H_1$ is decided, otherwise $u_k = 0$ is sent. We assume that local sensors do not communicate with each other.

Let $\gamma_k(\cdot)$ denote the local decision rule for Sensor $k$, $k \in \{1, 2, ..., K\}$, then the local decision, $u_k$, $\forall k$, can be given as

$$u_k = \gamma_k(x_k) = \begin{cases} 0, & \text{decide } H_0 \\ 1, & \text{decide } H_1. \end{cases} \quad (1)$$

For each Sensor $k$, the detection performance is characterized by its false-alarm probability and detection probability, which can be given as [9]

$$P_F(k) = p(u_k = 1|H_0), \text{ and } P_D(k) = p(u_k = 1|H_1), \quad (2)$$

respectively. Due to independent decisions, the joint conditional probability density function (PDF) of the local decisions

can be expressed as

$$p(u_1, \ldots, u_K|H_i) = \prod_{k=1}^{K} p(u_k|H_i), \quad i = 0, 1. \quad (3)$$

All involved channels experience slow fading by assuming that the coherence time is much longer than the time interval between two consecutive decisions made by the FC [9]. Since the local decision of Sensor $k$ is transmitted over a slow-fading orthogonal channel, the received signal at the FC from Sensor $k$, $y_k$, can be given as

$$y_k = \sqrt{g_k} a_k u_k + n_k, \quad (4)$$

where $g_k$ is the channel power gain for Sensor $k$, $a_k$ is the signal amplitude, and $n_k$ is an additive noise which is assumed to be Gaussian with zero mean and $\sigma^2$ variance. Then, the conditional PDF of the received signals under two hypotheses can be written as

$$p(\mathbf{y}|H_i) = \sum_{\mathbf{u}} p(\mathbf{y}|\mathbf{u})p(\mathbf{u}|H_i), \quad i = 0, 1, \quad (5)$$

where $\mathbf{y} = [y_1, \ldots, y_K]^T$ and $\mathbf{u} = [u_1, \ldots, u_K]^T$.

There are three common criteria for measuring detection performance: i) the detection probability; ii) asymptotic relative efficiency (ARE); and iii) distance related bounds [9], [10], [39]. As we mentioned before, false alarm and detection probabilities have been critical performance criteria for cooperative spectrum sensing [11]–[15], where the closed-form analytical expressions of these two probabilities can be derived in certain scenarios. The ARE is usually used under the conditions of a large sample size at sensors and weak signal [9], [10]. For those reasons, distance related bounds, such as Ali-Silvey class of distance measures, are commonly used as tractable design criteria. Both the J-divergence (known as symmetric Kullback-Leibler (KL) divergence) and the Bhattacharyya bound belong to the Ali-Silvey class of distance. But J-divergence is more favorable, as it provides a lower bound on the detection error probability and also has fundamental connections with the ARE [39]. We thus choose the J-divergence as detection performance metric.

Let $P_k = a_k^2, \forall k = 1, \ldots, K$ indicate the transmitting power of Sensor $k$. Given a channel realization $\mathbf{g} = \{g_1, \ldots, g_K\}$, we assume that the amount of channel state information (CSI) available at the sensors is given as $\phi(\mathbf{g})$. We consider the following optimal power allocation that maximizes the J-divergence detection (denoted as $J(P_1(\phi(\mathbf{g})), \ldots, P_K(\phi(\mathbf{g})))$) under a total average power constraint and an individual power constraint, given as

$$\max_{\{P_1(\phi(\mathbf{g})), \ldots, P_K(\phi(\mathbf{g}))\}} E[J(P_1(\phi(\mathbf{g})), \ldots, P_K(\phi(\mathbf{g})))]$$

$$= \sum_{k=1}^{K} \mathbb{E}[N(g_k, P_k(\phi(\mathbf{g})))]$$

$$s.t. \sum_{k=1}^{K} \mathbb{E}[P_k(\phi(\mathbf{g}))] \leq \mathcal{P}_{\text{tot}};$$

$$0 \leq P_k(\phi(\mathbf{g})) \leq P_{\text{max}}^k,$$

$$\forall k = 1, \ldots, K, \text{ almost sure}, \quad (6)$$

where

$$N(g_k, P_k(\phi(\mathbf{g}))) \triangleq \frac{\sigma^2 + \alpha_F(k)g_k P_k(\phi(\mathbf{g}))}{\sigma^2 + \beta_F(k)g_k P_k(\phi(\mathbf{g}))} + \frac{\sigma^2 + \alpha_D(k)g_k P_k(\phi(\mathbf{g}))}{\sigma^2 + \beta_D(k)g_k P_k(\phi(\mathbf{g}))}, \quad (7)$$

and

$$\alpha_F(k) = P_F(k)(1 - P_D(k)) + P_D(k)(P_D(k) - P_F(k)),$$
$$\alpha_D(k) = P_D(k)(1 - P_F(k)) - P_F(k)(P_D(k) - P_F(k)),$$
$$\beta_F(k) = P_D(k)(1 - P_D(k)), \quad \beta_D(k) = P_F(k)(1 - P_F(k)).$$

with $0 < P_F(k) < P_D(k) < 1$ (otherwise the sensors do not provide useful information). When full CSI is available at sensors, i.e., $\phi(\mathbf{g}) = \mathbf{g}$, the optimization problem in (6) is studied in [9], and it is convex when $(P_D(k), P_F(k)) \in \mathcal{S}$ where the set $\mathcal{S}$ is defined as

$$\mathcal{S} = \Big\{ (P_D, P_F) \ \Big| \ \frac{3}{4} - \frac{1}{2}P_F - \frac{1}{4}\sqrt{1 + 12P_F - 12P_F^2}$$
$$\leq P_D \leq \frac{3}{4} - \frac{1}{2}P_F + \frac{1}{4}\sqrt{1 + 12P_F - 12P_F^2},$$
$$0 < P_F < P_D < 1 \Big\}. \quad (8)$$

For the rest of this paper, we also assume $(P_D(k), P_F(k)) \in \mathcal{S}$ is always true. The allocated power can thus be found by the water-filling algorithm [9], i.e., $P_k^*(\mathbf{g}) = \min\left[ P_{\max}^k, [P_k']^+ \right]$, $\forall k = 1, \ldots, K$, with $P_k'$ satisfying $\xi(g_k, P_k') - \lambda^* = 0$, where

$$\xi(g_k, P_k')$$
$$\triangleq \frac{(\alpha_F(k) - \beta_F(k))\sigma^2 g_k}{(\sigma^2 + \beta_F(k)g_k P_k')^2} + \frac{(\alpha_D(k) - \beta_D(k))\sigma^2 g_k}{(\sigma^2 + \beta_D(k)g_k P_k')^2}, \quad (9)$$

and $\lambda^*$ is the optimal nonnegative Lagrange multiplier associated with total transmitted power constraint, which can be obtained by solving $\lambda^*\left( \mathcal{P}_{\text{tot}} - \sum_{k=1}^K \mathbb{E}[P_k^*(\mathbf{g})] \right) = 0$.

*Lemma 1:* From $\xi(g_k, P_k) - \lambda = 0$, we have $P_k$ which is first increasing in $g_k \in [0, \hat{g}_k)$ (where $\hat{g}_k$ is the unique solution of $\frac{\partial \xi(g_k, P_k)}{\partial g_k} = 0$), and then decreasing over $g_k \in [\hat{g}_k, \infty)$.

*Proof:* See Appendix A.

However, the assumption of full CSI available at the sensors is generally unrealistic in practical systems, due to the limited bandwidth of the feedback channel. Therefore, in the next section, we will address this concern by investigating the J-divergence maximization problem with quantized CSI.

## III. OPTIMUM QUANTIZED POWER ALLOCATION WITH FINITE-RATE FEEDBACK

In this section, we focus on designing a power control strategy for each sensor based on quantized $\mathbf{g}$ acquired via a no-delay and error-free feedback link with limited rate.[1]

---

[1] Since we consider a slow block-fading environment and only a few bits of CSI feedback are required, the feedback delay is significantly less than the coherence time of the fading channels concerned. Thus, the effect of feedback delay can be negligible. Moreover, with such small data rate, error-free feedback can be easily achieved by using efficient error control coding over the feedback link [40], [41].

As depicted in Fig.1, we assume that the FC can perfectly estimate CSI of all sensors $\mathbf{g} = \{g_1, \ldots, g_K\}$ via a training-based channel estimation, and forward some appropriately quantized CSI to sensors through a finite-rate feedback link. Thus the sensors can adjust their transmission power based on the received quantized CSI. More specifically, given $B$ bits of limited feedback link between the FC and sensors, the space of the channel vector $\mathbf{g}$ is partitioned into $L = 2^B$ disjoint regions $\mathbb{G}_1, \ldots, \mathbb{G}_L$. A power codebook $\{\mathbf{P}_1, \ldots, \mathbf{P}_L\}$ of cardinality $L = 2^B$, with each $\mathbf{P}_j$ representing the power level using in the region $\mathbb{G}_j$, is designed offline purely based on the statistical information of $\mathbf{g}$, where $\mathbf{P}_j = [P_{1,j}, \ldots, P_{K,j}]^{\mathrm{T}}$ with integer $j \in \{1, 2, \ldots, L\}$. This power codebook is known *a priori* by both the FC and the sensors. Given an instantaneous channel realization, the FC employs a deterministic index mapping $\Theta$ from the current instantaneous $\mathbf{g}$ to one of $L$ indices, defined as $\Theta(\mathbf{g}) = j$, if $\mathbf{g} \in \mathbb{G}_j, j = 1, \ldots, L$. The index $j$ is then sent to sensors via a feedback link. Each sensor uses the associated power codebook element to adapt its transmission strategy.

Let $\Pr(\mathbb{G}_j), \mathbb{E}[\cdot | \mathbb{G}_j]$ represent the probability $\Pr(\mathbf{g} \in \mathbb{G}_j)$ and the expectation $\mathbb{E}[\cdot | \mathbf{g} \in \mathbb{G}_j]$, respectively. Then, the Problem (6) with limited feedback can be formulated as

$$\max_{\{\mathbf{P}_j\}, \{\mathbb{G}_j\}} \sum_{j=1}^L \mathbb{E}\left[ \sum_{k=1}^K N(g_k, P_{k,j}) \Big| \mathbb{G}_j \right] \Pr(\mathbb{G}_j),$$
$$\text{s.t.} \sum_{j=1}^L \left( \sum_{k=1}^K P_{k,j} \right) \Pr(\mathbb{G}_j) \leq P_{\text{tot}},$$
$$0 \leq P_{k,j} \leq P_{\max}^k,$$
$$\forall k = 1, \ldots, K, \quad \forall j = 1, \ldots, L. \quad (10)$$

Our goal is to jointly design quantization regions and the corresponding power codebook for the above optimization problem. Due to the non-convexity of problem (10), it is very challenging to find the globally optimal solution. However, in the next section, we provide an analysis to find a locally optimal solution for problem (10).

### A. A Locally Optimal Quantized Power Scheme

A local optimum of optimization problem (10) can be obtained by the Lagrange duality method, i.e., by solving the dual problem of (10), given as,

$$\min_{\lambda \geq 0} g(\lambda), \quad (11)$$

where $\lambda$ is a nonnegative Lagrange multiplier, and the Lagrange dual function $g(\lambda)$ is defined as,

$$g(\lambda) = \max_{0 \leq P_{k,j} \leq P_{\max}^k, \ \mathbb{G}_j \ \forall j} \mathcal{L}(\{\mathbf{P}_j\}, \{\mathbb{G}_j\}, \lambda) + \lambda P_{\text{tot}}, \quad (12)$$

with the Lagrangian

$$\mathcal{L}\left( \{\mathbf{P}_j\}_{j=1}^L, \{\mathbb{G}_j\}_{j=1}^L, \lambda \right)$$
$$= \sum_{j=1}^L \mathbb{E}\left[ \sum_{k=1}^K (N(g_k, P_{k,j}) - \lambda P_{k,j}) \Big| \mathbb{G}_j \right] \Pr(\mathbb{G}_j). \quad (13)$$

In order to solve the dual problem (11), we first assign an arbitrary initial value to $\lambda$, and then iteratively apply the

following two steps, Step 1 and Step 2, until a pre-specified convergence criterion is reached:

- **Step 1** (Inner loop): With a fixed $\lambda$, find locally optimal power codebook $\{\mathbf{P}_j\}_{j=1}^{L}$ and quantization regions $\{\mathbb{G}_j\}_{j=1}^{L}$ of the Lagrange dual function (12).
- **Step 2** (Outer loop): With the resulting power codebook and channel partitions, update the optimal value $\lambda$ by solving the dual problem (11) using a subgradient search method, i.e., updating $\lambda$ using $\lambda^{l+1} = \left[\lambda^l - \beta^l \left(\mathcal{P}_{\text{tot}} - \sum_{j=1}^{L} \left(\sum_{k=1}^{K} P_{k,j}^l\right) \Pr(\mathbb{G}_j)\right)\right]^{+}$, where $l$ is the iteration number and $\beta^l$ is a positive scalar step size for the $l$-th iteration satisfying $\sum_{l=1}^{\infty} \beta^l = \infty$ and $\sum_{l=1}^{\infty} (\beta^l)^2 < \infty$. Then go back to Step 1 until convergence (i.e., $\frac{|\lambda^{l+1}-\lambda^l|}{\lambda^l} < \epsilon$ or $\frac{|\mathcal{P}_{\text{tot}}-\mathcal{P}_{\text{new}}^l|}{\mathcal{P}_{\text{tot}}} < \epsilon$ with $\mathcal{P}_{\text{new}}^l \triangleq \sum_{j=1}^{L} \left(\sum_{k=1}^{K} P_{k,j}^l\right) \Pr(\mathbb{G}_j)$).

*Remark 1 (The Convergence of the Outer Loop): One cannot guarantee the feasibility of the primal variables in each Step 1 iteration, because the constraint may be violated at a particular step. However, in the next step (i.e., Step 2), the Lagrange multiplier value $\lambda$ will be adjusted accordingly using the above sub-gradient search method. In particular, if the power constraint is exceeded, the value $\lambda$ should increase, and decrease otherwise. Only in the limiting step when $\lambda$ converges, one can say that the primal variables are feasible. Note that the above sub-gradient update is guaranteed to converge to the optimum $\lambda$ as long as the step size $\beta^l$ is chosen satisfying the conditions stated above, due to the convexity of the dual problem (11) over $\lambda$ even though the primal problem is non-convex [42]. Due to the non-convexity of the primal problem however, the final obtained primal variables are only locally optimum. Fig.8 shows an example of the convergence rate of the outer loop iteration in a 4 sensor system. Please see Section IV for more details.*

For Step 1, we employ Generalized Lloyd Algorithm (GLA) with a modified distortion measure (i.e., a Lagrangian based cost function) to solve problem (12), as stated in [38]. This modified GLA is implemented using a sufficiently large number of training samples (channel realizations) and is designed based on following two necessary optimality conditions:

1) Given a power codebook $\mathbf{P} = \{\mathbf{P}_1, \ldots, \mathbf{P}_L\}$, the optimal channel partitions are determined by the nearest neighbor condition, i.e. $\forall j = 1, \ldots, L$,

$$\mathbb{G}_j = \{\mathbf{g} : \Phi(\mathbf{g}, \mathbf{P}_j) \geq \Phi(\mathbf{g}, \mathbf{P}_n), \ \forall n \in \{1, \ldots, L\}, \ n \neq j\}, \quad (14)$$

where $\Phi(\mathbf{g}, \mathbf{P}_j) = \sum_{k=1}^{K} N(g_k, P_{k,j}) - \lambda P_{k,j}$.

2) Given the channel partition $\mathbb{G}_1, \ldots, \mathbb{G}_L$, the optimal power codebook is updated by generalized centroid condition, namely,

$$\{P_{k,j}^*\}_{k=0}^{K} = \arg \max_{P_{k,j} \in [0, P_{\max}^k], \forall k=1,\ldots,K} \\ \mathrm{E}[\Phi(\mathbf{g}, \mathbf{P}_j)|\mathbb{G}_j]\Pr(\mathbb{G}_j), \ \forall j = 1, \ldots, L. \quad (15)$$

Note that $\mathbb{G}_j$ in the first condition is typically not unique since ties may break arbitrarily, and the second condition is a convex optimization problem [9]. Therefore, by beginning

with an arbitrary initial power codebook, one can repeatedly apply the two optimality conditions until convergence (i.e., $\frac{\|\mathbf{P}^{(n+1)}-\mathbf{P}^{(n)}\|}{\|\mathbf{P}^{(n)}\|} < \zeta$ or $\frac{\mathcal{L}^{(n+1)}-\mathcal{L}^{(n)}}{\mathcal{L}^{(n)}} < \zeta$, where $n$ is the iteration number) to obtain a locally optimal power codebook and the corresponding channel partitions for problem (12).

*Remark 2 (The Convergence of the Inner Loop): While convergence (to a local optimum) of our modified GLA follows immediately by noting that the Lagrangian $\mathcal{L}\left(\{\mathbf{P}_j\}_{j=1}^{L}, \{\mathbb{G}_j\}_{j=1}^{L}, \lambda\right) = \sum_{j=1}^{L} \mathrm{E}[\sum_{k=1}^{K}(N(g_k, P_{k,j}) - \lambda P_{k,j})|\mathbb{G}_j]\Pr(\mathbb{G}_j)$ is non-decreasing at each iteration and is upper bounded due to the finite average transmit power constraint. A more formal convergence proof can be obtained as well by alluding to [38, Th. 2], where it was shown that the above Lagrangian based modified GLA satisfies the global convergence and empirical consistency properties as defined in [43]. Fig. 7 shows an example of the convergence rate of the inner loop iteration in a 4-sensor system. Please see IV for more details.*

Despite the offline design of the GLA based quantizer, for each inner loop, it usually requires a large number of training samples, which leads to a prolonged convergence time of the outer loop. As the number of quantization regions $L$ increases, the number of training samples required will also increase. Thus it is computationally demanding even for the fusion center, especially when the number of feedback bits is large. For example, as shown in Fig.8 in a 4 sensors system, with 4 bits feedback, for each inner loop simulation, $10^5$ randomly generated channel realizations are used for each sensor. The outer loop takes $12305.808074$ seconds (i.e., $\approx 3.5$ hours) to converge. Please see Section IV for more details. In order to alleviate computational burden of the above GLA-based optimum scheme for large values of $L$ case, in the next section, we will derive a low complexity suboptimal scheme (similar to [44]) for implementing **Step 1** by approximating the channel quantization regions. It enables us to simplify the dual problem with an explicit formulation.

### B. Suboptimal Quantized Power Allocation (SQPA)

For each channel $g_k$, $k = 1 \ldots, K$, let $\{X_{k,1}, \ldots, X_{k,L_k}\}$ denote the associated quantization thresholds on the $g_k$-axis, where $0 = X_{k,0} < X_{k,1} \cdots < X_{k,L_k} = \infty$ and $L_k$ is the number of partition intervals on the $g_k$-axis (satisfying $\prod_{k=1}^{K} L_k = L$). In order to avoid the complexity of implementing modified GPA, we simply approximate the quantized channel regions (14) as $\mathbb{G}_{(l_1,\ldots,l_K)} \triangleq \{\mathbf{g}| g_k \in [X_{k,l_k-1}, X_{k,l_k}], \forall k = 1, \ldots, K\}$, with $\forall l_1 \in [1, L_1], \cdots, \forall l_K \in [1, L_K]$. Let $\mathbf{P}_{(l_1,\ldots,l_K)}$ represent the corresponding allocated power level for the current suboptimal setting, where $\mathbf{P}_{(l_1,\ldots,l_K)} \triangleq \left[P_{1,(l_1,\ldots,l_K)}, \ldots, P_{K,(l_1,\ldots,l_K)}\right]^{T}$. Then for given $L_1, \ldots, L_K$ (note that to find the optimal $L_1, \ldots, L_K$, we can exhaustively search over all the possible combinations of $L_1, \ldots, L_K$ and pick the combination that achieve the best performance), the dual function (12) becomes

$$g(\lambda) = \max_{0 \leq P_{k,(l_1,\ldots,l_K)} \leq P_{\max}^k, \forall k, \forall l_1,\ldots,\forall l_K \ \{X_{k,1},\ldots,X_{k,L_k}\},\forall k} \\ \times \mathcal{L}\left(\{\mathbf{P}_{(l_1,\ldots,l_K)}\}, \{\mathbb{G}_{(l_1,\ldots,l_K)}\}, \lambda\right) + \lambda \mathcal{P}_{\text{tot}}, \quad (16)$$

where

$$
\begin{aligned}
&\mathcal{L}\left(\left\{\mathbf{P}_{(l_1,\ldots,l_K)}\right\}, \left\{\mathbb{G}_{(l_1,\ldots,l_K)}\right\}, \lambda\right) \\
&= \sum_{l_1=1}^{L_1} \cdots \sum_{l_K=1}^{L_K} E\left[\sum_{k=1}^{K}\left(N(g_k, P_{k,(l_1,\ldots,l_K)}) - \lambda P_{k,(l_1\ldots l_K)}\right)\Bigg| \right.\\
&\qquad\qquad\qquad\qquad \left. \mathbb{G}_{(l_1,\ldots,l_K)}\right] \Pr\left(\mathbb{G}_{(l_1,\ldots,l_K)}\right) \\
&= \sum_{l_1=1}^{L_1} \cdots \sum_{l_K=1}^{L_K}\left\{\sum_{m=1}^{K}\left(\prod_{i\neq m}^{K} \mathcal{F}(X_{i,l_i})\right) \mathbb{N}(X_{m,l_m}, P_{m,(l_1\ldots l_K)})\right.\\
&\qquad\qquad\qquad\quad \left. - \lambda \prod_{i=1}^{K} \mathcal{F}(X_{i,l_i}) \sum_{j=1}^{K} P_{j,(l_1,\ldots,l_K)}\right\}, \qquad (17)
\end{aligned}
$$

with $\mathbb{N}(X_{k,l_k}, P_{k,(l_1\ldots l_K)}) = \int_{X_{k,l_k-1}}^{X_{k,l_k}} N(g_k, P_{k,(l_1\ldots l_K)}) f(g_k)\,dg_k$, and $f(g_k)$ being the PDf of $g_k$, and $\mathcal{F}(X_{k,l_k}) = F(X_{k,l_k}) - F(X_{k,l_k-1})$ with $F(\cdot)$ being the cumulative distribution function (CDF).

Now for **Step 1**, with a fixed $\lambda$, our aim becomes to find the jointly optimal quantization thresholds and the corresponding allocated power for problem (16). In the following, we employ a simple iterative algorithm (alternating optimization algorithm), similar to [44], to solve the problem. Let $q$ denote the iteration number. For $k = 1, \ldots, K$, the algorithm starts from an initial guess of $\left\{X_{k,1}^{(0)}, \ldots, X_{k,L_k-1}^{(0)}\right\}$, and then produces a sequence of iterates

$$
\begin{aligned}
&\left(\left\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q+1)}\right\}, \left\{\left\{X_{k,1}^{(q+1)}, \ldots, X_{k,L_k-1}^{(q+1)}\right\}_{k=1}^{K}\right\}\right) \\
&= \theta\left(\left\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q)}\right\}, \left\{\left\{X_{k,1}^{(q)}, \ldots, X_{k,L_k-1}^{(q)}\right\}_{k=1}^{K}\right\}\right),
\end{aligned}
$$

where $\theta$ operates as follows:

- For given partition thresholds

$$
\left\{\left\{X_{k,1}^{(q)}, \ldots, X_{k,L_k-1}^{(q)}\right\}_{k=1}^{K}\right\},
$$

a power codebook $\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q+1)}\}$ is chosen to maximize $\mathcal{L}(\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q+1)}\}, \{\mathbb{G}_{(l_1,\ldots,l_K)}^{(q)}\}, \lambda)$ in (17);
- And for the fixed $\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q+1)}\}$, $\{\{X_{k,1}^{(q)}, \ldots, X_{k,L_k-1}^{(q)}\}_{k=1}^{K}\}$ is updated by maximizing the objective function $\mathcal{L}(\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q+1)}\}, \{\mathbb{G}_{(l_1,\ldots,l_K)}^{(q+1)}\}, \lambda)$, until the resulting $\mathcal{L}(\{\mathbf{P}_{(l_1,\ldots,l_K)}\}, \{\mathbb{G}_{(l_1,\ldots,l_K)}\}, \lambda)$ converges within a pre-specified accuracy, i.e. $\frac{\mathcal{L}^{(q+1)} - \mathcal{L}^{(q)}}{\mathcal{L}^{(q)}} < \zeta$.

More specifically, the following two steps are carried out for each iteration $q$ (we omit the iteration index for the simplicity):

1) Given the channel partitions, the optimal power codebook is obtained by maximizing $\mathcal{L}\left(\left\{\mathbf{P}_{(l_1,\ldots,l_K)}\right\},\right.$ $\left\{\mathbb{G}_{(l_1,\ldots,l_K)}\right\}, \lambda)$, i.e., $\forall k, \forall l_1, \ldots, \forall l_K$,

$$
\begin{aligned}
P_{k,(l_1,\ldots,l_K)}^* &= \arg \max_{0 \leq P_{k,(l_1,\ldots,l_K)} \leq P_{\max}^k} \\
&\times E\left[\sum_{k=1}^{K} N(g_k, P_{k,(l_1,\ldots,l_K)}) - \lambda P_{k,(l_1\ldots l_K)}\Bigg| \right. \\
&\qquad\qquad \left. \mathbb{G}_{(l_1,\ldots,l_K)}\right] \Pr\left(\mathbb{G}_{(l_1,\ldots,l_K)}\right). \qquad (18)
\end{aligned}
$$

As Problem (18) is convex, the optimal power codebook can be easily solved by the water-filling algorithm:

$$
P_{k,(l_1,\ldots,l_K)}^* = \min\left[P_{\max}^k, [P_{k,(l_1,\ldots,l_K)}']^+\right], \quad (19)
$$

where $P_{k,(l_1,\ldots,l_K)}'$ satisfies

$$
\begin{aligned}
E\left[\xi\left(g_k, P_{k,(l_1,\ldots,l_K)}'\right) - \lambda\Big| \mathbb{G}_{(l_1,\ldots,l_K)}\right] \\
\times \Pr\left(\mathbb{G}_{(l_1,\ldots,l_K)}\right) = 0. \quad (20)
\end{aligned}
$$

From (20), we know that $P_{k,(l_1,\ldots,l_k,\ldots,l_K)}$ only depends on $X_{k,l_k-1}$ and $X_{k,l_k}$, thus, for the notational simplicity, we denote $P_{k,(l_1,\ldots,l_k,\ldots,l_K)}$ by $P_{k,l_k}$.

*Remark 3: If we assume the channel power gain $g_k, \forall k$ follows an exponential distribution with $E[g_k] = \frac{1}{\hat{\lambda}_k}$, we can obtain a closed-form expression for (20) as*

$$
\begin{aligned}
&\frac{\sigma^2 \hat{\lambda}_k}{P_{k,l_k}'^2}\left\{\frac{(\alpha_F(k)-\beta_F(k))}{\beta_F^2(k)}\left[\delta_F\left(\tilde{X}_{k,l_k}\right) - \delta_F\left(\tilde{X}_{k,l_k-1}\right)\right]\right.\\
&\qquad \left. + \frac{(\alpha_D(k)-\beta_D(k))}{\beta_D^2(k)}\left[\delta_D\left(\tilde{X}_{k,l_k}\right) - \delta_D\left(\tilde{X}_{k,l_k-1}\right)\right]\right\}\\
&= \lambda\left(e^{-\hat{\lambda}_k X_{k,l_k-1}} - e^{-\hat{\lambda}_k X_{k,l_k}}\right), \qquad (21)
\end{aligned}
$$

*where $\tilde{X}_{k,l_k} = \hat{\lambda}_k X_{k,l_k}$, $\tilde{X}_{k,l_k-1} = \hat{\lambda}_k X_{k,l_k-1}$,*

$$
\delta_D(x) \triangleq \frac{e^{-x}\sigma_D^2}{\sigma_D^2 + x} + e^{\sigma_D^2}\left(1 + \sigma_D^2\right)\text{Ei}\left[-\sigma_D^2 - x\right],
$$

$$
\delta_F(x) \triangleq \frac{e^{-x}\sigma_F^2}{\sigma_F^2 + x} + e^{\sigma_F^2}\left(1 + \sigma_F^2\right)\text{Ei}\left[-\sigma_F^2 - x\right],
$$

*with $\sigma_D^2 = \frac{\sigma^2 \hat{\lambda}_k}{\beta_D(k) P_{k,l_k}'}$, $\sigma_F^2 = \frac{\sigma^2 \hat{\lambda}_k}{\beta_F(k) P_{k,l_k}'}$, and $\text{Ei}[\cdot]$ is the exponential integral defined as $\text{Ei}[z] = -\int_{-z}^{\infty} \frac{e^{-t}}{t}dt$.*

2) With the resulting power codebook, we then update quantization thresholds by maximizing $\mathcal{L}(\{\mathbf{P}_{(l_1,\ldots,l_K)}\},$ $\left\{\mathbb{G}_{(l_1,\ldots,l_K)}\right\}, \lambda)$, i.e.,

$$
\begin{aligned}
&\{X_{k,1}, \ldots, X_{k,L_k-1}\}_{k=1}^{K} \\
&= \arg \max_{0 < X_{k,1} < \cdots < X_{k,L_k-1} < \infty, \forall k} \sum_{l_1=1}^{L_1} \cdots \sum_{l_K=1}^{L_K} \\
&\times \left\{\sum_{m=1}^{K}\left(\prod_{i\neq m}^{K} \mathcal{F}(X_{i,l_i})\right) \mathbb{N}(X_{m,l_m}, P_{m,l_m})\right.\\
&\qquad\quad \left. - \lambda \prod_{i=1}^{K} \mathcal{F}(X_{i,l_i}) \sum_{j=1}^{K} P_{j,l_j}\right\}. \qquad (22)
\end{aligned}
$$

Problem (22) in general is a non-convex optimization problem. However, we will prove in Lemma 4 that there

exists a unique maximum point for Problem (22), and this optimal solution $X_{k,l_k}^*, \forall l_k = 1, \ldots, L_k$, is obtained by solving the following equation and inequality:

$$\begin{cases} \lambda - \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right) = 0; \\ \left(P_{k,l_k+1} - P_{k,l_k}\right) \frac{\partial \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)}{\partial X_{k,l_k}} < 0, \end{cases} \quad (23)$$

where

$$\begin{aligned} &\psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right) \\ &= \frac{\left(\alpha_F(k) - \beta_F(k)\right)\sigma^2 X_{k,l_k}}{\left(\sigma^2 + \beta_F(k)X_{k,l_k}P_{k,l_k}\right)\left(\sigma^2 + \beta_F(k)X_{k,l_k}P_{k,l_k+1}\right)} \\ &+ \frac{\left(\alpha_D(k) - \beta_D(k)\right)\sigma^2 X_{k,l_k}}{\left(\sigma^2 + \beta_D(k)X_{k,l_k}P_{k,l_k}\right)\left(\sigma^2 + \beta_D(k)X_{k,l_k}P_{k,l_k+1}\right)}. \end{aligned} \quad (24)$$

We now present some important intermediate results :
*Lemma 2: For Sensor k, let $P_{k,m_k}$ be the largest power level, then based on Lemma 1 the optimal power levels obtained in (20) satisfies: $0 \leq P_{k,1} < \cdots < P_{k,m_k-1} < P_{k,m_k} > P_{k,m_k+1} > \cdots > P_{k,L_k}$.*
   *Proof:* See Appendix C.
*Lemma 3: $\frac{\partial \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)}{\partial X_{k,l_k}} = 0$ has a unique solution, and $\frac{\partial \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)}{\partial P_{k,l_k}} < 0$, $\frac{\partial \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)}{\partial P_{k,l_k+1}} < 0$.*
   *Proof:* See Appendix B.
*Lemma 4: Based on Lemma 3 there exists a unique optimal solution $\{X_{k,1} \ldots X_{k,L_k-1}\}$ for Problem (22), which is obtained by solving (23). And the resulting $\{X_{k,l_k}\}_{l_k=1}^{L_k-1}$ automatically satisfies $0 < X_{k,1} < \cdots < X_{k,L_k-1} < \infty$.*
   *Proof:* See Appendix D.

Therefore, with a fixed $\lambda$, a scheme which finds a suboptimal power codebook and the corresponding quantization regions for Problem (16) is summarized in Algorithm 1:

---

**Algorithm 1** Suboptimal Quantized Power Allocation (SQPA)

---

1: Initialize $q = 0$, choose arbitrary initial value of $\left\{X_{k,1}^{(0)}, \ldots, X_{k,L_k-1}^{(0)}\right\}$ satisfying $0 < X_{k,1}^{(0)} < \cdots < X_{k,L_k-1}^{(0)} < \infty$ for $k = 1, \ldots, K$
2: **repeat**
   1) Given $\left\{X_{k,1}^{(q)}, \ldots, X_{k,L_k-1}^{(q)}\right\}_{k=1}^{K}$, optimal $\left\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q+1)}\right\}$ is given by (19), for $l_k = 1, \ldots, L_k$
   2) Given $\left\{\mathbf{P}_{(l_1,\ldots,l_K)}^{(q+1)}\right\}$, optimal $\left\{X_{k,1}^{(q+1)}, \ldots, X_{k,L_k-1}^{(q+1)}\right\}_{k=1}^{K}$ is given by (23), for $l_k = 1, \ldots, L_k - 1$
   3) $q = q + 1$
3: **until** Convergence: $\frac{\mathcal{L}^{(q+1)} - \mathcal{L}^{(q)}}{\mathcal{L}^{(q+1)}} < \zeta$

---

As we mentioned in **Step 2**, the optimal $\lambda$ is determined by solving the dual problem, satisfying

$$\lambda\left(\mathcal{P}_{\text{tot}} - \sum_{l_1=1}^{L_1} \cdots \sum_{l_K=1}^{L_K}\left(\sum_{k=1}^{K} P_{k,(l_1\ldots l_K)}\right)\Pr(\mathbb{G}_{(l_1\ldots l_K)})\right) = 0.$$

Since the average power $\sum_{l_1=1}^{L_1} \cdots \sum_{l_K=1}^{L_K}\left(\sum_{k=1}^{K} P_{k,(l_1\ldots l_K)}\right)$ $\Pr(\mathbb{G}_{(l_1\ldots l_K)})$ can be shown to be a monotonically decreasing function of $\lambda$ using a similar proof of Lemma 1 in [45], we can find the unique optimal $\lambda$ via a simple bisection method.

*C. Asymptotically Large Number of Feedback Bits Scenario*

In this section, we look at the SQPA algorithm and analyze the system performance in the high rate quantization region, i.e., for a case with a large number of feedback bits $B$. Under this circumstance, the design of optimal channel partitions becomes less significant. Thus, with high rate quantization, we can apply the idea of equal probability per region (EPrPR) to design a simple suboptimal channel partition [46], [41]. Assuming $g_k, \forall k$ follows an exponential distribution with the mean of $\frac{1}{\hat{\lambda}_k}$, the corresponding CDF can be written as $F(X_{k,l_k}) = 1 - e^{-\hat{\lambda}_k X_{k,l_k}}$. Then, by applying EPrPR to each $g_k$-axis, we can obtain $F(X_{k,l_k}) = \frac{l_k}{L_k}, \forall l_k = 1, \ldots, L_k - 1$. Hence, the quantization threshold $\left\{X_{k,l_k}^*\right\}$ is given as,

$$X_{k,l_k}^* = \frac{1}{\hat{\lambda}_k} \ln \frac{L_k}{L_k - l_k}, \quad \forall l_k = 1, \ldots, L_k - 1. \quad (25)$$

which depends only on the number of partitions $L_k$.

With a sufficient large number of feedback bits, by applying the mean value theorem (MVT) [47], the left hand side of (20) can be approximated as

$$\begin{aligned} &\mathrm{E}\left[\xi\left(g_k, P_{k,l_k}'\right) - \lambda \big| \mathbb{G}_{(l_1,\ldots,l_K)}\right]\Pr\left(\mathbb{G}_{(l_1,\ldots,l_K)}\right) \\ &= \left(\prod_{j\neq k}^{K} \frac{1}{L_j}\right)\int_{X_{k,l_k-1}}^{X_{k,l_k}}\left[\xi\left(g_k, P_{k,l_k}'\right) - \lambda\right]f(g_k)dg_k \\ &\approx \left(\prod_{j=1}^{K} \frac{1}{L_j}\right)\left[\xi\left(\frac{1}{\hat{\lambda}_k}\ln\frac{L_k+1}{L_k-l_k+1}, P_{k,l_k}'\right) - \lambda\right]. \quad (26) \end{aligned}$$

Substituting (26) into (20), the power codebook can be explicitly derived as,

$$P_{k,l_k}^* = \begin{cases} y\left(P_{k,l_k}^*\right), & \text{if } \frac{d_0\ln\frac{L_k+1}{L_k-l_k+1} - \lambda\hat{\lambda}_k\sigma^2}{\lambda\hat{\lambda}_k\sigma^2 - (\alpha_D(k)-\beta_D(k))\ln\frac{L_k+1}{L_k-l_k+1}} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (27)$$

where $d_0 = \alpha_F(k) + \alpha_D(k) - \beta_F(k) - \beta_D(k)$ [9], and

$$\begin{aligned} y\left(P_{k,l_k}^*\right) = &\frac{\sigma^2\hat{\lambda}_k}{\beta_F(k)\ln\frac{L_k+1}{L_k-l_k+1}} \\ &\times \left(\sqrt{\frac{\alpha_F(k) - \beta_F(k)}{\frac{\lambda\sigma^2\hat{\lambda}_k}{\ln\frac{L_k+1}{L_k-l_k+1}} - \frac{\alpha_D(k)-\beta_D(k)}{\left(1+\frac{\beta_D(k)\ln\frac{L_k+1}{L_k-l_k+1}}{\sigma^2\hat{\lambda}_k}P_{k,l_k}^*\right)^2}}} - 1\right). \end{aligned} \quad (28)$$

*Remark 4: Note that* $y\left(P_{k,l_k}^*\right)$ *in* (28) *decreases as* $P_{k,l_k}^*$ *increases. We can thus apply a fixed-point iteration technique to obtain the optimal* $P_{k,l_k}^*$ *[48], if* $\dfrac{d_0 \ln \frac{L_k+1}{L_k-l_k+1} - \lambda\hat{\lambda}_k\sigma^2}{\lambda\hat{\lambda}_k\sigma^2 - (\alpha_D(k) - \beta_D(k))\ln\frac{L_k+1}{L_k-l_k+1}} > 0.$

Thus, the J-divergence can be explicitly approximated as

$$J \approx \sum_{k=1}^{K}\left\{\frac{1}{L_k}\sum_{l_k=1}^{L_k}\left[\frac{\sigma^2\hat{\lambda}_k + \alpha_F(k)\ln\frac{L_k+1}{L_k-l_k+1}P_{k,l_k}^*}{\sigma^2\hat{\lambda}_k + \beta_F(k)\ln\frac{L_k+1}{L_k-l_k+1}P_{k,l_k}^*}\right.\right.$$
$$\left.\left.+\frac{\sigma^2\hat{\lambda}_k + \alpha_D(k)\ln\frac{L_k+1}{L_k-l_k+1}P_{k,l_k}^*}{\sigma^2\hat{\lambda}_k + \beta_D(k)\ln\frac{L_k+1}{L_k-l_k+1}P_{k,l_k}^*}\right]\right\} \quad (29)$$

*Remark 5: With average individual power constraints, we can rewrite the Problem* (10) *as*

$$\max_{\mathbf{P}_j,\ \mathbb{G}_j,\ \forall j} \sum_{j=1}^{L}\mathrm{E}\left[\sum_{k=1}^{K}N(g_k,P_{k,j})\middle|\mathbb{G}_j\right]\mathrm{Pr}(\mathbb{G}_j)$$

$$s.t.\ \sum_{j=1}^{L}P_{k,j}\mathrm{Pr}(\mathbb{G}_j) \leq p_k, \quad \forall k = 1,\dots,K,$$
$$0 \leq P_{k,j} \leq P_{\max}^k, \quad \forall j = 1,\dots,L, \quad (30)$$

*where* $p_k$ *is the average power constraint for Sensor* $k$. *The Lagrange dual function* $g(\{\lambda_k\})$ *is defined as:*

$$g(\{\lambda_k\}) = \max_{0\leq P_{k,j}\leq P_{\max}^k,\ \mathbb{G}_j,\forall j}\mathcal{L}\left(\{\mathbf{P}_j\},\{\mathbb{G}_j\},\{\lambda_k\}\right)$$
$$+\sum_{k=1}^{K}\lambda_k p_\mathrm{k}, \quad (31)$$

*with*

$$\mathcal{L}\left(\{\mathbf{P}_j\},\{\mathbb{G}_j\},\{\lambda_k\}\right)$$
$$=\sum_{j=1}^{L}\mathrm{E}\left[\sum_{k=1}^{K}(N(g_k,P_{k,j}) - \lambda_k P_{k,j})\middle|\mathbb{G}_j\right]\mathrm{Pr}\left(\mathbb{G}_j\right). \quad (32)$$

*All the properties depicted in Section III can be easily applied to the case of average individual power constraints. In addition, if* $\sum_{k=1}^{K}p_k = \mathcal{P}_\mathrm{tot}$, *solution of Problem* (30) *is one of the feasible solutions of Problem* (10).

## IV. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed power allocation algorithms via numerical simulations. In our simulations, we consider a network in which there are $K$ sensors around the FC and distance from sensors to the FC are $d_1,\dots,d_K$ (in meters) with the corresponding operating frequency $f_1,\dots,f_K$ (in megahertz) in the range of 902-928 MHz (915 MHz band) based on IEEE 802.15.4 standard (2006/2011) [49].[2] The path-loss in free

[2]According to the IEEE 802.15.4 standard (2006/2011) [49], there are a few license-free frequency industrial, scientific, and medical radio (ISM) bands available for wireless sensor networks, such as: (1)868-868.6 MHz (868 MHz band) with data rate 20kbps used in Europe, (2) 902-928 MHz (915 MHz band) with data rate 40kbps used mainly in North America and Australia (3) 2400-2483.5 MHz (2.4 GHz band) with data rate 250kbps is used worldwide.
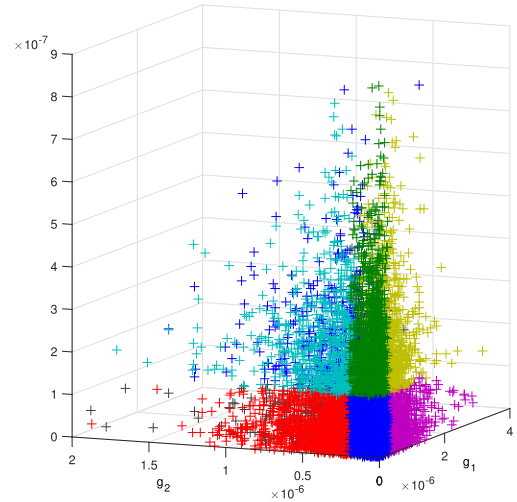


(a) SQPA
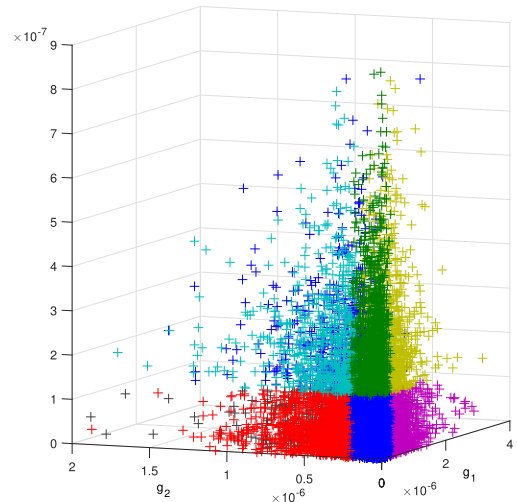


(b) GLA

Fig. 2. Quantization regions of three-bit feedback for a system with three sensors.

space of signal power at the FC from each sensor follows the free-space path-loss model [50],

$$PL_k = 20\log_{10}(d_k) + 20\log_{10}(f_k) - 27.55.$$

Thus, the channel gain of each sensor follows an exponential distribution with mean of $10^{\frac{-PL_k}{10}}$. The noise variance at the FC is $\sigma^2 = -70\,\mathrm{dBm}$ and the maximum transmitting power of each sensor is $P_{\max}^k = 2\,\mathrm{mW}$. The total power budget is less than $2K\,\mathrm{mW}$, otherwise each sensor uses its maximum transmitting power. In practice, many of the WSN applications, such as mission-critical surveillance, require a low false alarm rate (no more than 0.05) and a high detection probability (no lower than 0.5) [51]. Thus, we assume all sensors perform Neyman-Pearson detection with false alarm probability set to $P_F(k) = 0.04$ and the detection probability $P_D(k) \geq 0.5, \quad k = 1,\dots,K.$

TABLE I

ALLOCATED POWER AT EACH REGION FOR THE SQPA CASE AND THE GLA CASE

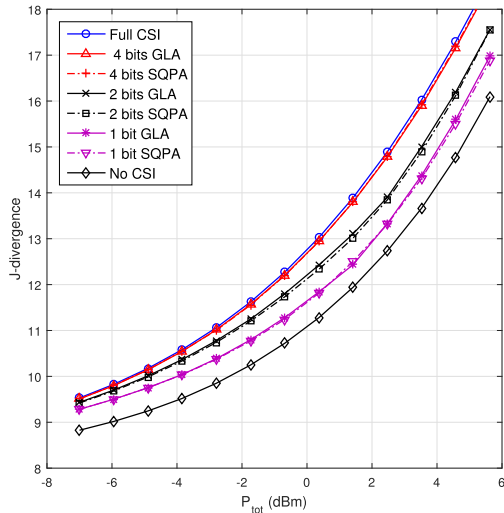| Quantization Regions | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| Sensor 1 | SQPA | 1.94 | 0 | 1.94 | 1.94 | 0 | 0 | 1.94 | 0 |
| | GLA | 1.93 | 0 | 1.95 | 1.95 | 0 | 0 | 1.92 | 0 |
| Sensor 2 | SQPA | 0 | 2 | 0 | 2 | 0 | 2 | 2 | 0 |
| | GLA | 0 | 2 | 0 | 2 | 0 | 2 | 2 | 0 |
| Sensor 3 | SQPA | 2 | 2 | 0 | 2 | 2 | 0 | 0 | 0 |
| | GLA | 2 | 2 | 0 | 2 | 2 | 0 | 0 | 0 |



Fig. 3. Detection performance comparison between GLA and SQPA with 1, 2, and 4 bits feedback in a 4 sensors system.
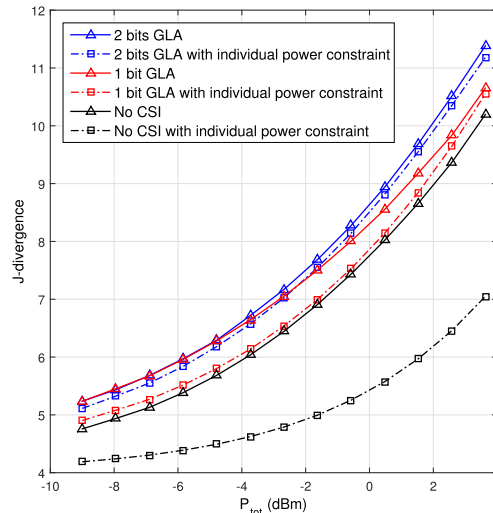


Fig. 4. Comparison between total power constraint and individual power constraint with 1, 2 bits feedback and no feedback in a 2 sensors system.

For a comparison, Fig. 2 plots the quantization regions obtained by both SQPA and GLA strategies with 3 bits feedback under a total power constraint of 2.2523 mW in a three-sensor ($K = 3$) network, where $d_1 = 44.26$ m, $f_1 = 909$ MHz; $d_2 = 58.57$ m, $f_2 = 916$ MHz; $d_3 = 72.66$ m, $f_3 = 923$ MHz, and $P_D(1) = 0.7$, $P_D(2) = 0.8$, $P_D(3) = 0.9$. The partition regions of the SQPA case is almost identical to that of the GLA case, where the allocated power at each region is in Table I. Fig. 2 and Table I clearly demonstrate that the suboptimal SQPA scheme with 3 bits of feedback can almost approach the optimal GLA performance, which thus confirms the effectiveness SQPA.

Fig. 3 depicts the performance of the J-divergence obtained by the SQPA algorithm versus $\mathcal{P}_{\text{tot}}$ with feedback bits $B = \{1, 2, 4\}$ in a four-sensor ($K = 4$) network, where $d_1 = 29.74$ m, $f_1 = 902$ MHz; $d_2 = 44.26$ m, $f_2 = 909$ MHz; $d_3 = 58.57$ m, $f_3 = 916$ MHz; $d_4 = 72.66$ m, $f_4 = 923$ MHz and $P_D(1) = 0.6$, $P_D(2) = 0.7$, $P_D(3) = 0.8$, $P_D(4) = 0.9$. We also plot the corresponding J-divergence performance obtained by the GLA. Even with only 1 bit feedback, the J-divergence performance of SQPA and the corresponding GLA are very close. Further, we plot the corresponding J-divergence performance of two extreme cases: i) no CSI at sensors in which sensors use the pre-specified power level to transmit for all channel conditions; and ii) full CSI

at sensors with the power allocation scheme given in (9). As shown in Fig. 3, when the number of feedback bits increases, the performance gap between the SQPA and the full CSI decreases considerably. With only 4 bits of feedback, SQPA can achieve almost the same performance as full CSI case.

For example, given $\mathcal{P}_{\text{tot}} = 2.8701$ mW (4.5789 dBm), the J-divergence loss due to imperfect CSI are approximately 14.52%, 10.29%, 6.77%, and 0.57% for no CSI case, 1 bit feedback, 2 bits feedback, and 4 bits feedback, respectively. A similar behavior also can be observed from Fig. 3 for the J-divergence obtained by GLA. These results are particularly useful as the GLA-based optimal scheme becomes computationally intensive for the case of large number of feedback bits (as mentioned in Section III A), whereas SQPA provides a performance that is extremely close to that of the GLA, while requiring much less computation time.

Fig. 4 and Fig. 5 show the J-divergence behavior of the GLA algorithm with the average and individual power constraints for two-sensor and four-sensor networks, respectively. For the sake of comparison and simplicity, we assume all sensors are under the same average power constraint and $\sum_{k=1}^{K} p_k = \mathcal{P}_{\text{tot}}$. For two-sensor network, parameters are $d_1 = 29.74$ m, $d_2 = 72.66$ m, $f_1 = 902$ MHz, $f_2 = 923$ MHz, $P_D(1) = 0.7$ and $P_D(2) = 0.9$. In general, we notice that the total
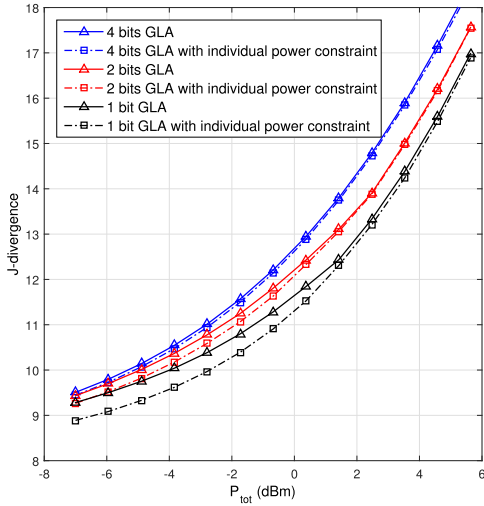
Fig. 5. Comparison between total power constraint and individual power constraint with 1, 2, and 4 bits feedback in a 4 sensors system.
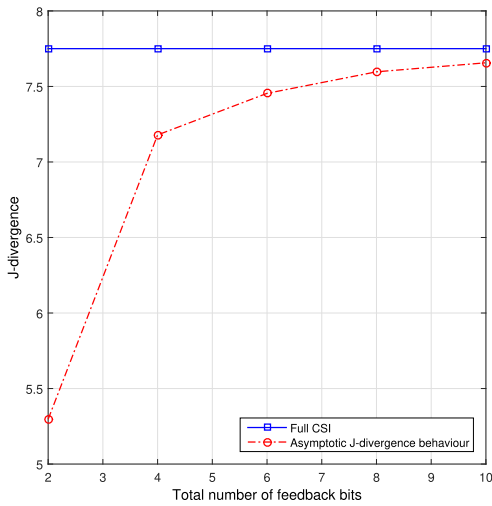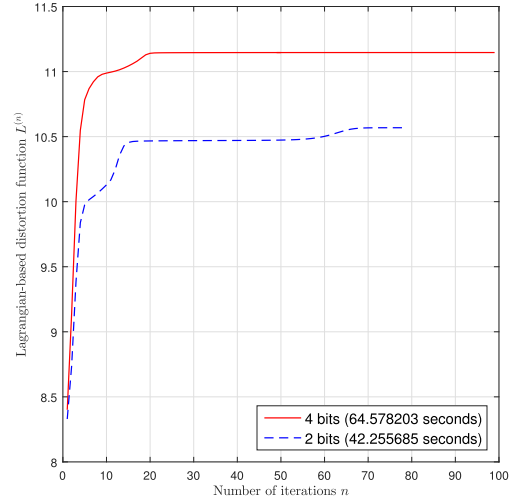


Fig. 7. Inner loop iteration convergence rate with given $\lambda = 2$ for proposed GLA-based locally optimal scheme in a 4 sensors system



Fig. 6. Asymptotic J-divergence behavior in a two-sensor network versus the number of quantization level, with $\mathcal{P}_{\mathrm{tot}} = -1.6316$dBm.
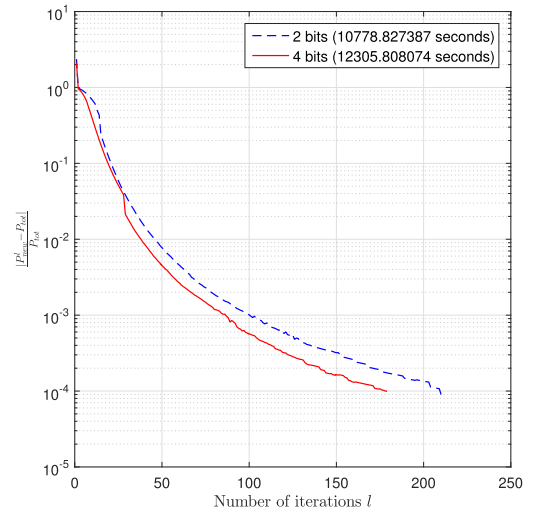


Fig. 8. Outer loop iteration convergence rate for proposed GLA-based locally optimal scheme in a 4 sensors system, given $\mathcal{P}_{\mathrm{tot}} = 1.6$ mW, initial value $\lambda = 1$ and step size $2/k$.

power constraint scenario outperforms the individual power constraint scenario, especially for no CSI case (with no CSI, under the individual power constraint, each sensor simply transmits by using all its power budget). The superiority of the total power constraint scenario is due to the fact that the solution of Problem (30) (individual power constraints) is always a feasible solution of Problem (10) (total power constraint), but it does not work the other way around. The total power constraint scenario therefore has a larger feasible set and achieves better resource efficiency. However, this performance gap gradually decreases when number of feedback bits increase.

A similar performance behavior can be observed in Fig.3 for the 4-sensor network case ( with same parameters). These observations confirm that the solution of Problem (30) is one of the feasible solutions of Problem (10).

Fig. 6 shows the asymptotic J-divergence behavior of SQPA versus the number of total feedback bits $B$ with the total transmission power budget $\mathcal{P}_{\mathrm{tot}} = -1.6316$dBm for

a two-sensor network. We compare the performance with full CSI. When the number of feedback bits increases, the J-divergence increases, and approaches the value corresponding to full CSI performance. This further confirms that a near-full CSI system performance can be achieved by SQPA with only a small number of feedback bits.

*Convergence Rate:* Fig. 7 and Fig. 8 illustrate the convergence rate of our proposed GLA-based locally optimal quantized power scheme (i.e., a two-step iteration scheme introduced in Section III A) in a four-sensor system (with system parameters same as Fig. 3) for 2 bits and 4 bits feedback, respectively. All schemes were implemented in MATLAB (R2015b) on an Intel(R) Core(TM) i7-5600 CPU with a clock speed of 2.60 GHz and a memory of 16 GB. For each inner loop simulation, $10^5$ randomly generated channel realizations are used for each sensor with fixed $\lambda = 2$. As we can observe from Fig. 7, the Lagrangian-based distortion function is non-decreasing at each iteration and converges as it is upper bounded due to the average total power constraint,

which agrees with Remark 2. Fig. 8 shows the convergence rate of the proposed scheme for the total transmission power budget $\mathcal{P}_{\text{tot}} = 1.6\,\text{mW}$. Starting with a randomly chosen initial value $\lambda = 1$ and step size $2/k$, the outer loop takes 10778.827387 seconds (i.e., $\approx$ 3 hours) and 12305.808074 seconds (i.e., $\approx$ 3.5 hours) to converge, respectively, for 2 bits and 4 bits feedback cases. We also compare the computational convergence speed of the suboptimal SQPA and optimal GLA algorithms. For a given total transmission power $\mathcal{P}_{\text{tot}} = 1.6$ mW with 2 bits feedback, SQPA takes only around 152.987741 seconds (i.e., $\approx$ 2.5 minutes) whereas the GLA takes 10778.827387 seconds (i.e., $\approx$ 3 hours) in order to achieve a comparable level of accuracy.

## V. CONCLUSION

In this paper, we consider a problem of transmit power allocation in order to maximize the detection performance subject to the total and individual power constraints on each sensor. Instead of full-CSI, we assume that only the quantized CSI is available at the sensors acquired via limited feedback from the FC. A locally optimal solution is derived by solving its dual Lagrangian optimization problem based on a Lagrangian based modified GLA. Due to the high complexity of the Modified GLA, we further propose an efficient and low-complexity scheme to determine the quantization regions. Although the problem is non-convex, we explicitly formulate the problem and solve for an unique solution, which provides a locally optimal solution. However, numerical results show that only a small amount of feedback bits can give almost the same detection performance as the full CSI case. Future work will involve extending the results to MIMO slow-fading channels.

## APPENDIX

### A. Proof of Lemma 1

Let $S(g_k, P'_k) = \xi(g_k, P'_k) - \lambda$, then by applying the 'Implicit Function Theorem', we have $\frac{\partial P'_k}{\partial g_k} = -\frac{\partial S(g_k, P'_k)/\partial g_k}{\partial S(g_k, P'_k)/\partial P'_k}$. As $\frac{\partial S(g_k, P'_k)}{\partial P'_k} = \frac{\partial \xi(g_k, P'_k)}{\partial P'_k} < 0$ [9], we know that Lemma 1 holds if $\frac{\partial S(g_k, P'_k)}{\partial g_k} > 0$ in $g_k \in [0, \hat{g}_k)$ and $\frac{\partial S(g_k, P'_k)}{\partial g_k} < 0$ over $g_k \in [\hat{g}_k, \infty)$, where $\hat{g}_k$ is the unique point of $\frac{\partial \xi(g_k, P'_k)}{\partial g_k} = 0$. Note that, we have,

$$\frac{\partial S(g_k, P'_k)}{\partial g_k} = \frac{\sigma^2\, Q'}{(\sigma^2 + \beta_F(k)g_k P'_k)^3 (\sigma^2 + \beta_D(k)g_k P'_k)^3},$$

with

$$Q' \triangleq Z_4(g_k P'_k)^4 + Z_3(g_k P'_k)^3 + Z_2(g_k P'_k)^2 + Z_1(g_k P'_k)^1 + Z_0$$

where,

$$Z_0 = \sigma^8 d_0 \overset{(a)}{>} 0,$$
$$Z_1 = \sigma^6(3d_1 - C_0)$$
$$\overset{(b)}{=} 2\sigma^6(P_D(k) - P_F(k))^2[P_D(k)^2 P_F(k) + P_D(k)P_F(k)^2$$
$$+ (2 - P_D(k) - P_F(k))(1 - P_D(k))(1 - P_F(k))] \overset{(c)}{>} 0,$$

$$Z_2 = 3\sigma^4(\beta_D(k)\alpha_F(k) - \beta_F(k)\alpha_D(k))(\beta_D(k) - \beta_F(k))$$
$$\overset{(d)}{=} 3\sigma^4(P_D(k) - P_F(k))^2(P_D(k) + P_F(k) - 1)^2$$
$$\times [P_D(k)(1 - P_F(k)) + P_F(k)(1 - P_D(k))] \overset{(e)}{>} 0,$$
$$Z_3 = (\alpha_F(k) - \beta_F(k))\sigma^2\beta_D(k)^2(\beta_D(k) - 3\beta_F(k))$$
$$+ (\alpha_D(k) - \beta_D(k))\sigma^6\beta_F(k)^2(\beta_F(k) - 3\beta_D(k)),$$
$$Z_4 = -\beta_D(k)\beta_F(k)d_2 \overset{(f)}{<} 0, \qquad (33)$$

with $d_0, d_1, d_2$ and $C_0$ defined as [9]:

$$d_1 = \alpha_F(k)\beta_D(k) + \alpha_D(k)\beta_F(k) - 2\beta_F(k)\beta_D(k),$$
$$d_2 = \alpha_F(k)\beta_D(k)^2 + \alpha_D(k)\beta_F(k)^2$$
$$- \beta_F(k)\beta_D(k)^2 - \beta_F(k)^2\beta_D(k),$$
$$C_0 = \beta_F(k)(\alpha_F(k) - \beta_F(k)) + \beta_D(k)(\alpha_D(k) - \beta_D(k)).$$

In (33), as shown in [9] that when $1 > P_D(k) > P_F(k) > 0$, $d_0, d_1, d_2, \beta_F(k), \beta_D(k)$ are positive, which can give us $(a), (f)$. We also can obtain $(b), (d)$ by applying the definition of $\alpha_F(k), \beta_F(k), \alpha_D(k), \beta_D(k)$, and $(c), (e)$ are followed from $1 > P_D(k) > P_F(k) > 0$. Therefore, we have $Z_0, Z_1, Z_2$ are positive, and $Z_4$ is negative, which gives us $Q'$ having one sign change (regardless the sign of $Z_3$). According to Descartes' rule of signs we can see there is only one positive solution $\hat{g}_k$ satisfying $\frac{\partial S(g_k, P_k)}{\partial g_k} = 0$; together with $\left.\frac{\partial S(g_k, P_k)}{\partial g_k}\right|_{g_k=0} = \frac{d_0}{\sigma^2} > 0$, we get $S(g_k, P_k)$ is first increasing in $g_k \in [0, \hat{g}_k)$ and then decreasing over $g_k \in [\hat{g}_k, \infty)$.

### B. Proof of Lemma 2

1) $\frac{\partial \psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1})}{\partial X_{k,l_k}} = 0$ *has unique solution:* After some mathematical manipulation, we can express

$$\frac{\partial \psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1})}{\partial X_{k,l_k}}$$
$$= \frac{Q}{(1+dX_{k,l_k})^2 (1+mX_{k,l_k})^2 (1+DX_{k,l_k})^2 (1+MX_{k,l_k})^2},$$

where

$$d = \frac{\beta_F(k)}{\sigma^2}P_{k,l_k}, \quad m = \frac{\beta_D(k)}{\sigma^2}P_{k,l_k},$$
$$D = \frac{\beta_F(k)}{\sigma^2}P_{k,l_k+1}, \quad M = \frac{\beta_D(k)}{\sigma^2}P_{k,l_k+1},$$
$$Q = T_6\, X_{k,l_k}^6 + T_5\, X_{k,l_k}^5 + T_4\, X_{k,l_k}^4 + T_3\, X_{k,l_k}^3$$
$$+ T_2\, X_{k,l_k}^2 + T_1\, X_{k,l_k}^1 + T_0$$

with

$$T_0 = \frac{d_0}{\sigma^2} \overset{(h)}{>} 0, \quad T_1 = \frac{2d_1}{\sigma^4}(P_{k,l_k+1} + P_{k,l_k}) \overset{(l)}{>} 0,$$
$$T_2 = \frac{d_2}{\sigma^6}(P_{k,l_k+1}^2 + P_{k,l_k}^2 + P_{k,l_k}P_{k,l_k+1})$$
$$+ \frac{P_{k,l_k+1}P_{k,l_k}}{\sigma^6}\left[\beta_F^2(k)(\alpha_F(k) - \beta_F(k))\right.$$
$$\left. + \beta_D^2(k)(\alpha_D(k) - \beta_D(k)) - 3d_2\right]$$

$$T_4 = \frac{P_{k,l_k+1}P_{k,l_k}}{\sigma^{10}} \left[ d_2 P_{k,l_k+1} P_{k,l_k} \left( \beta_F(k)^2 + 4\beta_F(k)\beta_D(k) + \beta_D(k)^2 \right) - 4d_1 \beta_F(k)\beta_D(k) P_{k,l_k+1} P_{k,l_k} \left( \beta_F(k) + \beta_D(k) \right) \right.$$
$$\left. - d_0 \beta_F(k)^2 \beta_D(k)^2 \left( P_{k,l_k+1}^2 + P_{k,l_k+1}P_{k,l_k} + P_{k,l_k}^2 \right) \right] \qquad (34)$$

$$\frac{\partial \psi \left( X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1} \right)}{\partial P_{k,l_k}}$$

$$= \frac{-X_{k,l_k}^2 \sigma^2}{\left( \sigma^2 + \beta_F(k)X_{k,l_k}P_{k,l_k} \right)^2 \left( \sigma^2 + \beta_F(k)X_{k,l_k}P_{k,l_k+1} \right) \left( \sigma^2 + \beta_D(k)X_{k,l_k}P_{k,l_k} \right)^2 \left( \sigma^2 + \beta_D(k)X_{k,l_k}P_{k,l_k+1} \right)}$$
$$\times \left[ \sigma^6 C_0 + \beta_F(k)\beta_D(k) \left( 2\sigma^4 d_0 X_{k,l_k} P_{k,l_k} + \sigma^4 d_0 X_{k,l_k} P_{k,l_k+1} + \sigma^2 d_1 X_{k,l_k}^2 P_{k,l_k}^2 \right. \right.$$
$$\left. \left. + 2\sigma^2 d_1 X_{k,l_k}^2 P_{k,l_k} P_{k,l_k+1} + d_2 X_{k,l_k}^3 P_{k,l_k}^2 P_{k,l_k+1} \right) \right] \overset{(u)}{<} 0 \qquad (35)$$

---

$$\overset{(g)}{=} \frac{d_2}{\sigma^6} \left( P_{k,l_k+1}^2 + P_{k,l_k}^2 + P_{k,l_k} P_{k,l_k+1} \right)$$
$$+ \frac{2 P_{k,l_k} P_{k,l_k+1} \left( P_D(k) - P_F(k) \right)^2}{\sigma^6}$$
$$\times [(P_D(k) + P_F(k) - 1)^2 (2 - P_D(k) - P_F(k))(P_D(k)$$
$$+ P_F(k)) + 2(1 - P_D(k))(1 - P_F(k))$$
$$\times P_D(k) P_F(k)] \overset{(s)}{>} 0,$$

$$T_3 = \frac{2 Z_2 \left( P_{k,l_k+1} + P_{k,l_k} \right) P_{k,l_k+1} P_{k,l_k} (\beta_F(k) + \beta_D(k))}{3\sigma^{12}}$$
$$\overset{(r)}{>} 0,$$

$$T_5 = \frac{-2 d_1 \beta_D(k)^2 \beta_F(k)^2}{\sigma^{12}} \left( P_{k,l_k+1} + P_{k,l_k} \right) P_{k,l_k+1}^2 P_{k,l_k}^2$$
$$\overset{(t)}{<} 0,$$

$$T_6 = -\frac{d_2 \beta_D(k)^2 \beta_F(k)^2}{\sigma^{14}} P_{k,l_k+1}^3 P_{k,l_k}^3 \overset{(n)}{<} 0,$$

and $T_4$ is defined in (34) (shown in the top of this page), where $d_0, d_1, d_2$ and $Z_2$ are defined in Lemma 1. $(g)$ is obtained by applying the definition of $\alpha_F(k), \beta_F(k), \alpha_D(k), \beta_D(k)$; and $(h), (l), (s), (r), (t), (n)$ are obtained due to the fact that $d_0, d_1, d_2, Z_2$ are all positive and $1 > P_D(k) > P_F(k) > 0$.

Therefore, for $Q$, regardless the sign of $T_4$, the number of sign differences between consecutive coefficients is 1, i.e. $T_0, T_1, T_2, T_3$ are positive and $T_5, T_6$ are negative. Applying Descartes' rule of signs, we can obtain the only one solution $\hat{X}_{k,l_k}$ satisfying $\frac{\partial \psi \left( X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1} \right)}{\partial X_{k,l_k}} = 0$.

2) $\frac{\partial \psi \left( X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1} \right)}{\partial P_{k,l_k}} < 0$: Proof is given in (35) (shown in the top of this page), where $C_0$ is defined in Lemma 1. We obtain $(u)$ since $\beta_D(k), \beta_F(k), d_0, d_1, d_2$ are positive and $C_0$ is nonnegative. The proof of $\frac{\partial \psi \left( X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1} \right)}{\partial P_{k,l_k+1}} < 0$ follows a similar proof in (35).

### C. Proof of Lemma 3

Applying the mean value theorem (MVT) to (20), we can have $\xi(\bar{X}_{k,l_k}, P_{k,l_k}) - \lambda = 0, \forall l_k = 1, \ldots, L_k$, where $\bar{X}_{k,l_k} \in (X_{k,l_k-1}, X_{k,l_k})$. According to Lemma 1, when $\xi(g_k, P_k) - \lambda = 0$, $P_k$ is first increasing in $g_k \in [0, \hat{g}_k)$ (where $\hat{g}_k$ is the unique solution of $\frac{\partial \xi(g_k, P_k)}{\partial g_k} = 0$), and then decreasing over $g_k \in [\hat{g}_k, \infty)$. Without losing generality, assuming $\hat{g}_k \in [X_{k,j_k-1}, X_{k,j_k})$, $j_k$ is an integer and $j_k \in [1, L_k]$. If $P_{k,j_k-1} < P_{k,j_k} > P_{k,j_k+1}$, we must have

$P_{k,1} < \cdots < P_{k,j_k-1} < P_{k,j_k} > P_{k,j_k+1} > \cdots > P_{k,L_k}$. If $P_{k,j_k} < P_{k,j_k+1} > P_{k,j_k+2}$, we will have $P_{k,1} < \cdots < P_{k,j_k} < P_{k,j_k+1} > P_{k,j_k+2} > \cdots > P_{k,L_k}$. If $P_{k,j_k-2} < P_{k,j_k-1} > P_{k,j_k}$ we obtain $P_{k,1} < \cdots < P_{k,j_k-2} < P_{k,j_k-1} > P_{k,j_k} > \cdots > P_{k,L_k}$. Any of these cases can be written in the form of: $P_{k,1} < \cdots < P_{k,m_k-1} < P_{k,m_k} > P_{k,m_k+1} > \cdots > P_{k,L_k}$, where $m_k \in [1, L_k]$, and $P_{k,m_k}$ is the largest power level for Sensor $k$.

### D. Proof of Lemma 4

A local solution of Problem (22) can be obtained by solving the Karush Kuhn Tucker (KKT) conditions, given as:

$$\left( P_{k,l_k+1} - P_{k,l_k} \right) \left[ \lambda - \psi \left( X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1} \right) \right] = 0.$$

1) $P_{k,l_k+1} > P_{k,l_k}$: As $\frac{\partial}{\partial X_{k,l_k}}(\lambda - \psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}))|_{X_{k,l_k}=0} = -T_0 < 0$, based on Lemma 3 we have $\lambda - \psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1})$ is first decreasing in $X_{k,l_k} \in [0, \hat{X}_{k,l_k})$, where $\hat{X}_{k,l_k}$ is the unique solution of $\frac{\partial \psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1})}{\partial X_{k,l_k}} = 0$, and then increasing over $X_{k,l_k} \in [\hat{X}_{k,l_k}, \infty)$. From Lemma 3 that $\psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1})$ is decreasing w.r.t $P_{k,l_k}$, we have $\psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}) > \xi(X_{k,l_k}, P_{k,l_k+1})$, together with $E\left[ \xi(g_k, P_{k,l_k+1}) - \lambda | \mathbb{G}_{(l_1, \ldots, l_{k+1}, \ldots, l_K)} \right] = 0$ (obtained from (20)), we can get $E[\psi(g_k, P_{k,l_k}, P_{k,l_k+1})| \mathbb{G}_{(l_1, \ldots, l_{k+1}, \ldots, l_K)}] > \lambda$; and from MVT we know that there exists $\tilde{X}_{k,l_k} \in (X_{k,l_k}, X_{k,l_k+1})$ such that

$$\psi \left( \tilde{X}_{k,l_k+1}, P_{k,l_k}, P_{k,l_k+1} \right)$$
$$= E\left[ \psi(g_k, P_{k,l_k}, P_{k,l_k+1}) | \mathbb{G}_{(l_1, \ldots, l_{k+1}, \ldots, l_K)} \right] > \lambda.$$

Thus, there are two solutions of $\lambda - \psi(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}) = 0$: one maximum point and one minimum point for Problem (22), and the maximum point can be obtained by (23).

Next, we will prove the resulting $\{X_{k,l_k}\}_{l_k=1}^{L_k-1}$ from Problem (22) satisfies $0 < X_{k,1} < \cdots < X_{k,L_k-1} < \infty$. Given $X'_{k,l_k-1} < X'_{k,l_k} < X'_{k,l_k+1} < X'_{k,l_k+2}$ from last iteration, and applying MVT to (20) we can obtain:

$$\lambda = E\left[ \xi(g_k, P_{k,l_k}) | \mathbb{G}_{(l_1, \ldots, l_k, \ldots, l_K)} \right] = \xi \left( \bar{X}_{k,l_k}, P_{k,l_k} \right)$$
$$= E\left[ \xi(g_k, P_{k,l_k+1}) | \mathbb{G}_{(l_1, \ldots, l_{k+1}, \ldots, l_K)} \right]$$
$$= \xi \left( \bar{X}_{k,l_k+1}, P_{k,l_k+1} \right)$$
$$= E\left[ \xi(g_k, P_{k,l_k+2}) | \mathbb{G}_{(l_1, \ldots, l_{k+2}, \ldots, l_K)} \right]$$
$$= \xi \left( \bar{X}_{k,l_k+2}, P_{k,l_k+2} \right),$$

where $\bar{X}_{k,l_k}$, $\bar{X}_{k,l_k+1}$ and $\bar{X}_{k,l_k+2}$ satisfy $X'_{k,l_k-1} < \bar{X}_{k,l_k} < X'_{k,l_k} < \bar{X}_{k,l_k+1} < X'_{k,l_k+1} < \bar{X}_{k,l_k+2} < X'_{k,l_k+2}$. Because $P_{k,l_k+1} > P_{k,l_k}$, from Lemma 3 we know $\psi\left(\bar{X}_{k,l_k+1}, P_{k,l_k}, P_{k,l_k+1}\right) > \xi\left(\bar{X}_{k,l_k+1}, P_{k,l_k+1}\right) = \lambda = \xi\left(\bar{X}_{k,l_k}, P_{k,l_k}\right) > \psi\left(\bar{X}_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)$. Let $X^*_{k,l_k}$ denote the solution of (23), from above analysis we can obtain $X^*_{k,l_k} < \hat{X}_{k,l_k}$, where $\hat{X}_{k,l_k}$ is the unique solution of $\frac{\partial \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)}{\partial X_{k,l_k}} = 0$. Since $X^*_{k,l_k}$ satisfies $\psi\left(\bar{X}_{k,l_k+1}, P_{k,l_k}, P_{k,l_k+1}\right) > \lambda = \psi\left(X^*_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right) > \psi\left(\bar{X}_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)$, and $\psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)$ is first increasing in $X_{k,l_k} \in [0, \hat{X}_{k,l_k})$ and then decreasing over $X_{k,l_k} \in [\hat{X}_{k,l_k}, \infty)$, we can obtain $\bar{X}_{k,l_k} < X^*_{k,l_k} < \bar{X}_{k,l_k+1}$. If $P_{k,l_k+2} > P_{k,l_k+1}$, by using the similar analysis we can have $\bar{X}_{k,l_k+1} < X^*_{k,l_k+1} < \bar{X}_{k,l_k+2}$, thus, combining all the results above, we can conclude $X^*_{k,l_k} < X^*_{k,l_k+1}$. If $P_{k,l_k+2} < P_{k,l_k+1}$, similarly, we have $\psi\left(\bar{X}_{k,l_k+2}, P_{k,l_k+1}, P_{k,l_k+2}\right) < \xi\left(\bar{X}_{k,l_k+2}, P_{k,l_k+2}\right) = \lambda = \xi\left(\bar{X}_{k,l_k+1}, P_{k,l_k+1}\right) < \psi\left(\bar{X}_{k,l_k+1}, P_{k,l_k+1}, P_{k,l_k+2}\right)$ and $X^*_{k,l_k+1} > \hat{X}_{k,l_k+1}$. As $X^*_{k,l_k+1}$ satisfies $\psi\left(\bar{X}_{k,l_k+2}, P_{k,l_k+1}, P_{k,l_k+2}\right) < \lambda = \psi\left(X^*_{k,l_k+1}, P_{k,l_k+1}, P_{k,l_k+2}\right) < \psi\left(\bar{X}_{k,l_k+1}, P_{k,l_k+1}, P_{k,l_k+2}\right)$, and $\psi\left(X_{k,l_k+1}, P_{k,l_k+1}, P_{k,l_k+2}\right)$ is first increasing in $X_{k,l_k+1} \in [0, \hat{X}_{k,l_k+1})$ and then decreasing over $X_{k,l_k+1} \in [\hat{X}_{k,l_k+1}, \infty)$, we know that $\bar{X}_{k,l_k+1} < X^*_{k,l_k+1} < \bar{X}_{k,l_k+2}$. Together with $\bar{X}_{k,l_k} < X^*_{k,l_k} < \bar{X}_{k,l_k+1}$, we can have $X^*_{k,l_k} < X^*_{k,l_k+1}$.

*2) $P_{k,l_k+1} < P_{k,l_k}$:* The results can be obtained similar to *1)* case.

Therefore, based on the above analysis, we can conclude that there exists only one maximum point in Problem (22), thus only one optimal solution to Problem (22). And this optimal $X^*_{k,l_k}$ follows $0 < X^*_{k,1} < \cdots < X^*_{k,L_k-1} < \infty$ and can be obtained by solving the following:

$$\begin{cases} \lambda - \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right) = 0; \\ (P_{k,l_k+1} - P_{k,l_k}) \frac{\partial \psi\left(X_{k,l_k}, P_{k,l_k}, P_{k,l_k+1}\right)}{\partial X_{k,l_k}} < 0. \end{cases}$$

## REFERENCES

[1] Y. Chen *et al.*, "Energy-autonomous wireless communication for millimeter-scale Internet-of-Things sensor nodes," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3962–3977, Dec. 2016.

[2] Z. Chu, F. Zhou, Z. Zhu, R. Q. Hu, and P. Xiao, "Wireless powered sensor networks for Internet of Things: Maximum throughput and optimal power allocation," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 310–321, Feb. 2018.

[3] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[4] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Robust and locally-optimum decentralized detection with censoring sensors," in *Proc. Int. Conf. Inf. Fusion*, Annapolis, MD, USA, vol. 1, Jul. 2002, pp. 56–63.

[5] R. Viswanathan, S. C. A. Thomopoulos, and R. Tumuluri, "Optimal serial distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. TAES-24, no. 4, pp. 366–376, Jul. 1988.

[6] Z. B. Tang, K. R. Pattipati, and D. L. Kleinman, "Optimization of detection networks. I. Tandem structures," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 5, pp. 1044–1059, Sep./Oct. 1991.

[7] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[8] R. Niu, B. Chen, and P. K. Varshney, "Decision fusion rules in wireless sensor networks using fading channel statistics," in *Proc. Conf. Inf. Sci. Syst.*, Mar. 2003, pp. 1–6.

[9] X. Zhang, H. V. Poor, and M. Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4124–4140, Sep. 2008.

[10] H.-S. Kim and N. A. Goodman, "Power control strategy for distributed multiple-hypothesis detection," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3751–3764, Jul. 2010.

[11] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1232–1241, Apr. 2011.

[12] Z. Zhang, X. Wen, H. Xu, and L. Yuan, "Sensing nodes selective fusion scheme of spectrum sensing in spectrum-heterogeneous cognitive wireless sensor networks," *IEEE Sensors J.*, vol. 18, no. 1, pp. 436–445, Jan. 2018.

[13] D. Zhang *et al.*, "Energy-harvesting-aided spectrum sensing and data transmission in heterogeneous cognitive radio sensor network," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 831–843, Jan. 2017.

[14] A. Patel, H. Ram, A. K. Jagannatham, and P. K. Varshney, "Robust cooperative spectrum sensing for MIMO cognitive radio networks under CSI uncertainty," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 18–33, Jan. 2018.

[15] W. Ejaz and M. Ibnkahla, "Multiband spectrum sensing and resource allocation for IoT in cognitive 5G networks," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 150–163, Feb. 2018.

[16] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2668–2673, Nov. 2005.

[17] R. Jiang, S. Misra, B. Chen, and A. Swami, "Robust suboptimal decision fusion in wireless sensor networks," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, vol. 4, Oct. 2005, pp. 2107–2113.

[18] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1018–1027, Mar. 2006.

[19] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Decentralized detection with censoring sensors," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1362–1373, Apr. 2008.

[20] B. Liu and B. Chen, "Decentralized detection in wireless sensor networks with channel fading statistics," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, 2006, Art. no. 062915.

[21] H. R. Ahmadi and A. Vosoughi, "Impact of channel estimation error on decentralized detection in bandwidth constrained wireless sensor networks," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, San Diego, CA, USA, Nov. 2008, pp. 1–7.

[22] J. Matamoros and C. Antón-Haro, "Opportunistic power allocation and sensor selection schemes for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 534–539, Feb. 2010.

[23] H. R. Ahmadi and A. Vosoughi, "Channel aware sensor selection in distributed detection systems," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Jun. 2009, pp. 71–75.

[24] N. K. D. Venkategowda, B. B. Narayana, and A. K. Jagannatham, "Precoding for robust decentralized estimation in coherent-MAC-based wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 24, no. 2, pp. 240–244, Feb. 2017.

[25] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

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

[27] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors I. Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[28] D. A. Pados, K. W. Halford, D. Kazakos, and P. Papantoni-Kazakos, "Distributed binary hypothesis testing with feedback," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, no. 1, pp. 21–42, Jan. 1995.

[29] S. Alhakeem and P. K. Varshney, "Decentralized Bayesian detection with feedback," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 26, no. 4, pp. 503–513, Jul. 1996.

[30] W. P. Tay and J. N. Tsitsiklis, "The value of feedback for decentralized detection in large sensor networks," in *Proc. Int. Symp. Wireless Pervasive Comput. (ISWPC)*, Feb. 2011, pp. 1–6.

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

[32] H. T. Nguyen, I. Balasingham, T. A. Ramstad, and P. A. Floor, "Limited feedback transmission scheme for wireless sensor networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2010, pp. 1–6.

[33] A. G. Marques, X. Wang, and G. B. Giannakis, "Minimizing transmit-power for coherent communications in wireless sensor networks using quantized channel state information," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Honolulu, HI, USA, vol. 3, Apr. 2007, pp. 529–532.

[34] A. G. Marques, X. Wang, and G. B. Giannakis, "Minimizing transmit power for coherent communications in wireless sensor networks with finite-rate feedback," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4446–4457, Sep. 2008.

[35] C.-H. Wang and S. Dey, "Distortion outage minimization in Nakagami fading using limited feedback," *EURASIP J. Adv. Signal Process.*, vol. 2011, Dec. 2011, Art. no. 92.

[36] M. Fanaei, M. C. Valenti, and N. A. Schmid, "Power allocation for distributed BLUE estimation with full and limited feedback of CSI," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2013, pp. 418–423.

[37] M. Fanaei, M. C. Valenti, and N. A. Schmid, "Limited-feedback-based channel-aware power allocation for linear distributed estimation," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2013, pp. 547–551.

[38] Y. He and S. Dey, "Power allocation in spectrum sharing cognitive radio networks with quantized channel information," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1644–1656, Jun. 2011.

[39] M. Lazaro, M. Sanchez-Fernandez, and A. Artes-Rodriguez, "Optimal sensor selection in binary heterogeneous sensor networks," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1577–1587, Apr. 2009.

[40] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1853–1863, May 2006.

[41] Y. He and S. Dey, "Power allocation for secondary outage minimization in spectrum sharing networks with limited feedback," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2648–2663, Jul. 2013.

[42] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[43] M. Sabin and R. Gray, "Global convergence and empirical consistency of the generalized Lloyd algorithm," *IEEE Trans. Inf. Theory*, vol. TIT-32, no. 2, pp. 148–155, Mar. 1986.

[44] Y. He and S. Dey, "Throughput maximization in Poisson fading channels with limited feedback," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4343–4356, Oct. 2013.

[45] R. Cendrillon, W. Yu, M. Moonen, J. Verliden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 922–933, May 2006.

[46] Y. He and S. Dey, "Throughput maximization in cognitive radio under peak interference constraints with limited feedback," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1287–1305, Mar. 2012.

[47] S. Lang, *A First Course in Calculus*. New York, NY, USA: Springer, 1986.

[48] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[49] *IEEE Standard for Local and Metropolitan Area Networks—Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)*, IEEE Standard 802.15.4-2011 (Rev.IEEE Standard 802.15.4-2006), 2011, pp. 1–314.

[50] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[51] G. Xing, R. Tan, B. Liu, J. Wang, X. Jia, and C.-W. Yi, "Data fusion improves the coverage of wireless sensor networks," in *Proc. Int. Conf. Mobile Comput. Netw. (MobiCom)*, Sep. 2009, pp. 157–168.

**Yuanyuan He** (S'08–M'12) received the Ph.D. degree from the Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, Australia, in 2011 (conferred in 2012). In 2017, she was a Visiting Researcher with the Department of Electrical and Computer Engineering, University of Alberta, Canada. She is currently a Research Fellow with the Department of Electrical and Electronic Engineering, The University of Melbourne. Her current research interests lie in the area of wireless communications including resource allocation, spatial modulation, and full duplex communications.

**Saman Atapattu** (M'14) received the B.Sc. degree in electrical and electronics engineering from the University of Peradeniya, Sri Lanka, the M.Eng. degree in telecommunications from the Asian Institute of Technology, Thailand, and the Ph.D. degree in electrical engineering from the University of Alberta, Canada.

He is currently a Research Fellow with the Department of Electrical and Electronic Engineering, The University of Melbourne, Australia. His general research interests include wireless communications and signal processing.

**Subhrakanti Dey** (SM'06) received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, IIT Kharagpur, Kharagpur, 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, in 1996.

From 1995 to 1997 and from 1998 to 2000, he was a Post-Doctoral Research Fellow with the Department of Systems Engineering, Australian National University. From 1997 to 1998, he was a Post-Doctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. He was a Professor with the Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, Australia, from 2000 to 2013. He has been a Professor with the Department of Engineering Sciences, Uppsala University, Sweden, since 2013. He is currently a Professor with the Institute of Telecommunications Research, University of South Australia, Adelaide, Australia. His current research interests include wireless communications and networks, signal processing for sensor networks, networked control systems, and molecular communication systems.

Dr. Dey currently serves on the Editorial Board of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS, and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was also an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2007 to 2010 and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL from 2004 to 2007, and also an Associate Editor for Elsevier *Systems and Control Letters* from 2003 to 2013.

**Xiaoxi Guo** (S'15) was born in Zhumadian, China, in 1988. He received the B.E. degree (Hons.) in telecommunications engineering from the University of Wollongong, Australia, in 2010, and the M.E. degree in telecommunications engineering and Ph.D. degree in electrical engineering from the University of Melbourne, Parkville, Australia, in 2012 and 2016, respectively.

His doctoral research focused on energy efficiency and security of the wireless sensor network. His current interests include statistical signal processing, machine learning, blockchain, and IoT.

**Jamie S. Evans** (S'93–M'98–SM'17) was born in Newcastle, Australia, in 1970. He received the B.S. degree in physics and the B.E. degree in computer engineering from the University of Newcastle in 1992 and 1993, respectively, and the M.S. and Ph.D. degrees in electrical engineering from The University of Melbourne, Australia, in 1996 and 1998, respectively. From 1998 to 1999, he was a Visiting Researcher with the Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley. Since returning to Australia in 1999, he has been holding academic positions at the University of Sydney, The University of Melbourne, and Monash University. He is currently a Professor and the Deputy Dean of the Melbourne School of Engineering, The University of Melbourne. His research interests are in communications theory, information theory, and statistical signal processing with a focus on wireless communications networks. He received the University Medal upon graduation from the University of Newcastle. He was awarded the Chancellor's Prize for excellence for his Ph.D. Thesis.