Performance of Distributed Estimation Over Unknown Parallel Fading Channels

Habib Şenol, *Member, IEEE*, and Cihan Tepedelenlioğlu, *Member, IEEE*

***Abstract*—We consider distributed estimation of a source in additive Gaussian noise, observed by sensors that are connected to a fusion center with unknown orthogonal (parallel) flat Rayleigh fading channels. We adopt a two-phase approach of i) channel estimation with training and ii) source estimation given the channel estimates and transmitted sensor observations, where the total power is fixed. In the second phase we consider both an equal power scheduling among sensors and an optimized choice of powers. We also optimize the percentage of total power that should be allotted for training. We prove that 50% training is optimal for equal power scheduling and at least 50% is needed for optimized power scheduling. For both equal and optimized cases, a power penalty of at least 6 dB is incurred compared to the perfect channel case to get the same mean squared error performance for the source estimator. However, the diversity order is shown to be unchanged in the presence of channel estimation error. In addition, we show that, unlike the perfect channel case, increasing the number of sensors will lead to an eventual degradation in performance. We approximate the optimum number of sensors as a function of the total power and noise statistics. Simulations corroborate our analytical findings.**

***Index Terms*—Channel estimation, convex optimization, distributed estimation, estimation diversity, parallel (orthogonal) multiple access, sensor networks.**

I. Introduction

A wireless sensor network (WSN) consists of spatially distributed sensors that are capable of monitoring physical phenomena. Sensors typically have limited processing and communication capability because of their limited battery power. In most WSNs, a fusion center (FC) that has fewer limitations in terms of processing and communication receives transmissions from the sensors over the wireless channels so as to combine the received signals to make inferences on the observed phenomenon.

Especially over the past few years, research on distributed estimation has been evolving very rapidly [1]. Universal decentralized estimators of a source observed in additive noise have been considered in [2] and [3]. Much of the literature has focused on finite-rate transmissions of quantized sensor observations [4]–[10]. The observations of the sensors can be delivered to the FC by analog or digital transmission methods. Amplify-and-forward is one analog option, whereas in digital transmission, observations are quantized, encoded, and transmitted via digital modulation. The optimality of amplify-and-forward is described in [11]–[14]. A type-based approach to estimating the histogram of the sensor observations is considered in [15] and [16], and the sensitivity of this approach to system nonidealities is addressed in [17]. In [14], an amplify-and-forward approach is employed with an orthogonal multiple-access channel (MAC) and perfect channel knowledge at the sensor side. In [14], it is argued that increasing the number of sensors improves the performance, and the concept of estimation diversity is introduced and shown to be given by the number of sensors.

To the best of our knowledge, there is not much work in the literature on distributed estimation over unknown fading channels. In this paper, we assume amplify-and forward transmission over *unknown* parallel (orthogonal MAC) fading channels. We follow a two-phase procedure where, in the first phase, sensors transmit pilots and the FC estimates the fading channels. In the second phase, sensors transmit amplified noisy observations of the source, and the FC estimates the source using the channel estimates. We characterize the effect of channel estimation error (CEE) on mean square error (MSE) performance for an equal power scheduling scenario that requires no channel status information at the sensor (CSIS). We also consider the case of CSIS where the sensors use the estimated channel information to optimize their transmission power. We show that when the total power is fixed, increasing the number of sensors will eventually lead to a degradation in performance, which is due to the increased CEE. We find an approximate expression for the optimum number of sensors to achieve minimum MSE performance. We also characterize the power penalty for not knowing the channel to be a factor of four (6 dB) or more for both equal and optimized power strategies.

This paper is organized as follows. Section II gives the system model. In Sections III and IV, channel estimation and source estimation are considered, respectively. In Section V, the MSE performance of the equal power allocation case is analyzed in the presence of channel estimation errors and compared with the perfect channel case. The MSE performance of the optimal power allocation case is analyzed in Section VI. Numerical results are discussed in Section VII. Section VIII concludes this paper.

II. System Model

We assume the WSN has K sensors and the kth sensor observes an unknown zero-mean complex random source signal

Manuscript received October 19, 2007; revised July 17, 2008. Current version published November 19, 2008. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Danilo P. Mandic. The work of H. Şenol was supported by The Scientific and Technological Research Council of Turkey (TUBITAK) between February and September 2007. The work of C. Tepedelenlioglu was supported by the National Science Foundation under CAREER grant CCR-0133841.

H. Şenol is with the Department of Computer Engineering, School of Engineering, Kadir Has University, Cibali 34083, Istanbul, Turkey (e-mail: hsenol@khas.edu.tr).

C. Tepedelenlioğlu is with the Department of Electrical Engineering, Fulton School of Engineering, Arizona State University, Tempe, AZ 85287 USA (e-mail: cihan@asu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2008.2005090

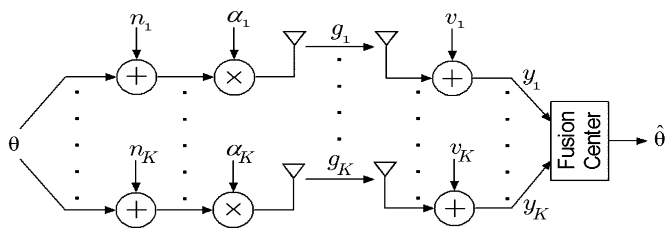


Fig. 1. Wireless sensor network with orthogonal MAC scheme.

θ with zero mean and variance σ_θ^2, corrupted by a zero-mean additive complex Gaussian noise $n_k \sim \mathcal{CN}(0, \sigma_{n_k}^2)$, as shown in Fig. 1. Since we assume the amplify-forward analog transmission scheme, the kth sensor amplifies its incoming analog signal $\theta + n_k$ by a factor of α_k and transmits it on the kth flat fading orthogonal channel to the FC. In Fig. 1, $g_k \sim \mathcal{CN}(0, \sigma_{g_k}^2)$ and $v_k \sim \mathcal{CN}(0, \sigma_{v_k}^2)$ are the flat fading gain and the noise on the kth channel path, respectively. The observation noise samples $\{n_k\}_{k=1}^K$ and the channel noise samples $\{v_k\}_{k=1}^K$ are assumed spatially uncorrelated. The amplification factor α_k might or might not depend on the fading coefficient g_k, depending on whether channel state information (CSI) is available at the sensor side. The data sample received at the FC over the kth channel is given by

$$y_k = g_k \alpha_k (\theta + n_k) + v_k, \quad k = 1, \ldots, K. \tag{1}$$

Based on this receive model, we will estimate the source signal θ. Even though it is possible to estimate the source signal without knowing the parallel channels $\{g_k\}_{k=1}^K$ by using the maximum likelihood (ML) estimator $\hat{\theta} = \max_\theta p(y_1, \ldots, y_K|\theta)$, it is rather complex analytically and computationally. Potentially tractable expectation maximization (EM) formulations are possible, but they are not guaranteed to converge to the true ML solution. Moreover, we would like to have an analytically tractable formulation to derive the MSE in closed form and optimize power. Therefore, instead of using an EM algorithm, we adopt a two-phase strategy to first estimate the parallel channels and then estimate the source signal given the channel estimates of the first (training) phase and the received signal of the second (data) phase. Assuming that the channels remain time-invariant over the course of estimation, we will use a linear minimum mean squared error (LMMSE) approach [18] for both phases.

In the first phase, the sensors send training symbols of total power $P_{\rm trn}$ to estimate the parallel channels $\{g_k\}_{k=1}^K$. In the second phase, the sensors transmit their amplified data that bear information about θ. The total power in this second phase is $P_{\rm tot} - P_{\rm trn}$. The fusion center uses the received signal in the second phase and the channel estimates from the first phase to estimate the source signal θ. The amplification factor used in the second phase might or might not be a function of the channel estimates, depending on if one is pursuing equal power allocation (Section V) or optimized power allocation (Section VI).

III. Fading Channel Estimation

To estimate the parallel fading channels $\{g_k\}_{k=1}^K$ in the training phase, we consider pilot-based channel estimation, where each sensor sends a pilot symbol to the FC over its own fading channel. The model for a pilot s_k transmitted over the kth channel is

$$x_k = g_k s_k + \nu_k \tag{2}$$

where x_k is the received signal and $\nu_k \sim \mathcal{CN}(0, \sigma_{v_k}^2)$ is the noise in the training phase. According to our observation model in (2), the LMMSE estimate $\hat{g}_k$ of the channel g_k is given as follows [18]:

$$\hat{g}_k = \frac{E_{\{g_k, x_k\}}[g_k x_k^*]}{E_{\{x_k\}}[|x_k|^2]} x_k = \frac{\sigma_{g_k}^2 s_k^*}{\sigma_{v_k}^2 + \sigma_{g_k}^2 |s_k|^2} x_k \tag{3}$$

where $(\cdot)^*$ denotes the complex conjugate and the channel estimation error variance δ_k^2 is given as

$$\delta_k^2 := \mathrm{var}(g_k - \hat{g}_k) = \left(\frac{1}{\sigma_{g_k}^2} + \frac{|s_k|^2}{\sigma_{v_k}^2}\right)^{-1} = \frac{\sigma_{v_k}^2 \sigma_{g_k}^2}{\sigma_{v_k}^2 + \sigma_{g_k}^2 |s_k|^2}. \tag{4}$$

Averaging (4) across sensors, we have

$$\frac{1}{K}\sum_{k=1}^K \delta_k^2 = \frac{1}{K}\sum_{k=1}^K \frac{\sigma_{v_k}^2 \sigma_{g_k}^2}{\sigma_{v_k}^2 + \sigma_{g_k}^2 |s_k|^2}. \tag{5}$$

It is possible to optimize the pilot symbols by minimizing (5) with respect to the pilot symbol powers $t_k := |s_k|^2$

$$\begin{aligned} \min_{t_1, \ldots, t_K} \quad & \frac{1}{K}\sum_{k=1}^K \frac{\sigma_{v_k}^2 \sigma_{g_k}^2}{\sigma_{v_k}^2 + \sigma_{g_k}^2 t_k} \\ \text{s.t.} \quad & \sum_{k=1}^K t_k \le P_{\rm trn} \\ & t_k \ge 0, \ \forall k. \end{aligned} \tag{6}$$

The Lagrangian function of the problem in (6) is obtained as

$$L(t_1, \ldots, t_K) = \frac{1}{K}\sum_{k=1}^K \frac{\sigma_{v_k}^2 \sigma_{g_k}^2}{\sigma_{v_k}^2 + \sigma_{g_k}^2 t_k} + \lambda\left(\sum_{k=1}^K t_k - P_{\rm trn}\right) - \sum_{k=1}^K \mu_k t_k.$$

With a straightforward application of the following Karush–Kuhn–Tucker (KKT) conditions [19] for this convex problem, we arrive at the following solution:

$$|s_k|^2 = \frac{\sqrt{\sigma_{v_k}^2}}{\sum_{i=1}^K \sqrt{\sigma_{v_i}^2}}\left(P_{\rm trn} + \sum_{j=1}^K \frac{\sigma_{v_j}^2}{\sigma_{g_j}^2}\right) - \frac{\sigma_{v_k}^2}{\sigma_{g_k}^2}. \tag{7}$$

Note that optimal power of the pilot symbols will be equal $|s_k|^2 = P_{\rm trn}/K \ \forall k$, for equal noise and channel variances ($\sigma_{v_k}^2 = \sigma_v^2$ and $\sigma_{g_k}^2 = \sigma_g^2 \ \forall k$). Substituting (7) in (4), the channel estimation error variance for the kth channel is obtained as

$$\delta_k^2 = \frac{\sqrt{\sigma_{v_k}^2}\sum_{i=1}^{K}\sqrt{\sigma_{v_i}^2}}{P_{\rm trn}+\sum_{j=1}^{K}\frac{\sigma_{v_j}^2}{\sigma_{g_j}^2}}. \quad (8)$$

Note that optimal training design requires a priori knowledge of channel and noise statistics and might not always be available. We will henceforth assume that the training power is uniformly allocated among sensors to keep our exposition simple.

IV. Source Estimation

In this section, we describe the estimation of the source signal θ. We choose the LMMSE source estimator given the amplification factors $\{\alpha_k\}_{k=1}^K$, the channel estimates $\{\hat{g}_k\}_{k=1}^K$ in (3), and the received signal $y_1,\ldots,y_K$ in (1). By doing this, we obtain the source estimator $\hat{\theta}$ in the presence of CEE. Note that the LMMSE estimator of θ when the channels $\{g_k\}_{k=1}^K$ are perfectly known is given by [18]

$$\hat{\theta}_{\rm PERF}(\boldsymbol{g}) = \left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{|g_k|^2|\alpha_k|^2}{|g_k|^2|\alpha_k|^2\sigma_{n_k}^2+\sigma_{v_k}^2}\right)^{-1} \times \sum_{k=1}^{K}\frac{g_k^*\alpha_k^*y_k}{|g_k|^2|\alpha_k|^2\sigma_{n_k}^2+\sigma_{v_k}^2} \quad (9)$$

where $\boldsymbol{g} := [g_1,\ldots,g_K]^T$ and $(\cdot)^T$ denotes transpose. In what follows, we will derive the LMMSE estimator of θ when we have only $\hat{g}_k$ in (3) and not the perfect CSI. We begin by expressing the received signal in (1) in terms of the estimated channel $\hat{g}_k$

$$y_k = \hat{g}_k\alpha_k\theta + \underbrace{\hat{g}_k\alpha_k n_k + (g_k-\hat{g}_k)\alpha_k(\theta+n_k)+v_k}_{:=w_k}. \quad (10)$$

We will now argue that the first and second terms of w_k in (10) are orthogonal as follows:

$$\begin{aligned} &E_{\{\theta,n_k,g_k|\hat{g}_k\}}\left[(\hat{g}_k\alpha_k n_k)^*\left((g_k-\hat{g}_k)\alpha_k(\theta+n_k)\right)\right]\\ &= E_{\{n_k,g_k|\hat{g}_k\}}\left[\hat{g}_k^*(g_k-\hat{g}_k)|\alpha_k|^2|n_k|^2\right]\\ &= \hat{g}_k^*\left(E_{\{g_k|\hat{g}_k\}}[g_k]-\hat{g}_k\right)|\alpha_k|^2\sigma_{n_k}^2\\ &= \hat{g}_k^*(\hat{g}_k-\hat{g}_k)|\alpha_k|^2\sigma_{n_k}^2 = 0 \end{aligned} \quad (11)$$

where in the second equality we used the fact that α_k is a function of $\hat{g}_k$ and the third equality follows because $\hat{g}_k = E_{g_k|x_k}[g_k] = E_{g_k|\hat{g}_k}[g_k]$, which is because x_k is a multiple of $\hat{g}_k$ as seen from (3).

In order to estimate source signal θ at the FC, we need to rewrite (10) using vectors

$$\boldsymbol{y} = \boldsymbol{h}\theta + \boldsymbol{w} \quad (12)$$

where $\boldsymbol{y} := [y_1,\ldots,y_K]^T$, $\boldsymbol{h} := [\hat{g}_1\alpha_1,\ldots,\hat{g}_K\alpha_K]^T$, $\boldsymbol{w} := [w_1,\ldots,w_K]^T$. Using (11), it is straightforward to show that the covariance matrix $\boldsymbol{C}_{\boldsymbol{w}|\hat{\boldsymbol{g}}}$ of $\boldsymbol{w}$ given the channel estimates $\hat{\boldsymbol{g}} := [\hat{g}_1,\ldots,\hat{g}_K]^T$ is diagonal with its kth diagonal element

$$E_{\{w_k|\hat{g}_k\}}\left[|w_k|^2\right] = |\hat{g}_k|^2|\alpha_k|^2\sigma_{n_k}^2 + |\alpha_k|^2\left(\sigma_\theta^2+\sigma_{n_k}^2\right)\delta_k^2+\sigma_{v_k}^2. \quad (13)$$

Given the channel estimate vector $\hat{\boldsymbol{g}}$, the LMMSE estimate $\hat{\theta}$ of the source θ is given by [18]

$$\begin{aligned}\hat{\theta} &= E_{\{\theta,\boldsymbol{y}|\hat{\boldsymbol{g}}\}}[\theta\boldsymbol{y}^\dagger]\left(E_{\{\boldsymbol{y}|\hat{\boldsymbol{g}}\}}[\boldsymbol{y}\boldsymbol{y}^\dagger]\right)^{-1}\boldsymbol{y}\\ &= \left(\frac{1}{\sigma_\theta^2}+\boldsymbol{h}^\dagger\boldsymbol{C}_{\boldsymbol{w}|\hat{\boldsymbol{g}}}^{-1}\boldsymbol{h}\right)^{-1}\boldsymbol{h}^\dagger\boldsymbol{C}_{\boldsymbol{w}|\hat{\boldsymbol{g}}}^{-1}\boldsymbol{y}\\ &= \left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{|\hat{g}_k|^2|\alpha_k|^2}{|\hat{g}_k|^2|\alpha_k|^2\sigma_{n_k}^2+|\alpha_k|^2\left(\sigma_\theta^2+\sigma_{n_k}^2\right)\delta_k^2+\sigma_{v_k}^2}\right)^{-1}\\ &\quad\times\sum_{k=1}^{K}\frac{\hat{g}_k^*{\alpha_k}^*y_k}{|\hat{g}_k|^2|\alpha_k|^2\sigma_{n_k}^2+|\alpha_k|^2\left(\sigma_\theta^2+\sigma_{n_k}^2\right)\delta_k^2+\sigma_{v_k}^2}.\end{aligned} \quad (14)$$

The conditional (on the channel estimates) MSE, which we will henceforth refer to simply as the MSE, is given by

$$\begin{aligned}D &= \left(\frac{1}{\sigma_\theta^2}+\boldsymbol{h}^\dagger\boldsymbol{C}_{\boldsymbol{w}|\hat{\boldsymbol{g}}}^{-1}\boldsymbol{h}\right)^{-1}\\ &= \left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{|\hat{g}_k|^2|\alpha_k|^2}{|\hat{g}_k|^2|\alpha_k|^2\sigma_{n_k}^2+|\alpha_k|^2\left(\sigma_\theta^2+\sigma_{n_k}^2\right)\delta_k^2+\sigma_{v_k}^2}\right)^{-1}\end{aligned} \quad (15)$$

where $(\cdot)^\dagger$ denotes conjugate transpose.

A natural alternative to (14) is obtained by substituting the channel estimates $\hat{\boldsymbol{g}}$ into (9), which is the LMMSE estimator when the CSI is perfect. We note that the proposed estimator in (14) is different from this approach. In other words, $\hat{\theta} \neq \hat{\theta}_{\rm PERF}(\hat{\boldsymbol{g}})$. Since both of these estimators are linear, and the proposed estimator achieves the minimum MSE among linear estimators, it is immediate that the proposed estimator $\hat{\theta}$ outperforms $\hat{\theta}_{\rm PERF}(\hat{\boldsymbol{g}})$ in the mean-squared sense. This result is also confirmed in Fig. 2 in Section VII.

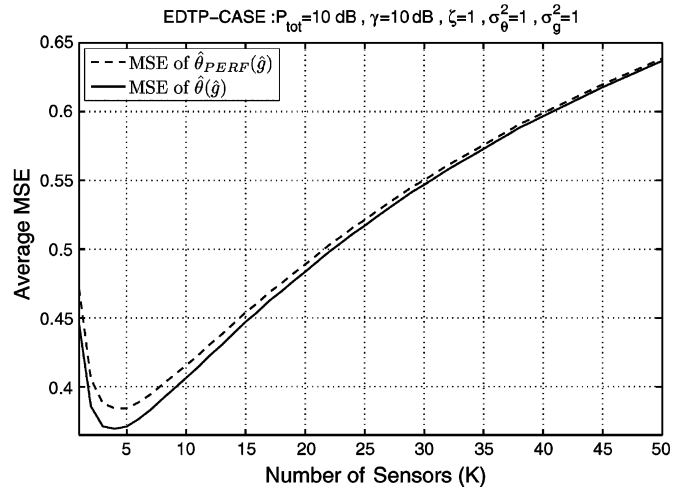


Fig. 2. MSE performance comparison of the estimators $\hat{\theta}$ and $\hat{\theta}_{\rm PERF}(\hat{\boldsymbol{g}})$ under equal power allocation.

We have seen that our model allows for nonidentical channel and noise statistics. However, following [14] to simplify the exposition, we will henceforth assume that the measurement noise n_k, the channel noise v_k, and the channels g_k are independent identically distributed (i.i.d.). For convenience and future reference, we introduce the following parameters:

Observation SNR
$\gamma := \sigma_\theta^2/\sigma_n^2$
Data transmission power used by kth sensor
$P_k := |\alpha_k|^2\left(\sigma_\theta^2+\sigma_n^2\right) = |\alpha_k|^2\sigma_\theta^2(1+\gamma^{-1})$
Total training power
$P_{\mathrm{trn}} := K|s|^2$
Average channel SNR
$\zeta := \sigma_g^2/\sigma_v^2$
Normalized kth estimated channel power
$\hat{\eta}_k := \frac{\zeta|\hat{g}_k|^2}{\sigma_{\hat{g}}^2(\gamma+1)} = \frac{\zeta|\hat{g}_k|^2}{\left(\sigma_g^2-\delta^2\right)(\gamma+1)}$
Normalized kth channel power
$\eta_k := \frac{\zeta|g_k|^2}{\sigma_g^2(\gamma+1)}$
Channel estimation error (CEE) variance
$\delta^2 = \mathrm{var}(g_k - \hat{g}_k)$ [see (4) and (17)]

where $\sigma_{\hat{g}}^2$ is the common variance of the channel estimate for each k. It is easy to see that $\sigma_g^2 = \sigma_{\hat{g}}^2 + \delta^2$, since $g_k = \hat{g}_k + (g_k - \hat{g}_k)$, and the LMMSE channel estimate $\hat{g}_k$ and the estimation error $g_k - \hat{g}_k$ are orthogonal [18]. From the above parameter table, the normalized estimated and true channel powers $\hat{\eta}_k$ and η_k are exponentially distributed random variables with common mean $\zeta/(\gamma+1)$. Using the variables in the table above, we put the MSE in the following convenient form:

$$D(P_{\mathrm{trn}}, P_1, \ldots, P_K) = \frac{\sigma_\theta^2}{1+\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k\left(\sigma_g^2-\delta^2\right)P_k}{\left(\hat{\eta}_k\left(\sigma_g^2-\delta^2\right)+\zeta\delta^2\right)P_k+\sigma_g^2}} \tag{16}$$

and we express the estimation error variance δ^2 using (4) and $P_{\mathrm{trn}} = K|s|^2$ as

$$\delta^2 = \frac{K\sigma_g^2}{K+\zeta P_{\mathrm{trn}}}. \tag{17}$$

Substituting (17) into (16), it is straightforward to verify that (16) is a convex function of $\{P_{\mathrm{trn}}, P_1, \ldots, P_K\}$ by taking the second derivative. For the purposes of optimization of the MSE in (16) with respect to $\{P_{\mathrm{trn}}, P_1, \ldots, P_K\}$, it suffices to work with

$$-\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k\left(\sigma_g^2-\delta^2\right)P_k}{\left(\hat{\eta}_k\left(\sigma_g^2-\delta^2\right)+\zeta\delta^2\right)P_k+\sigma_g^2}. \tag{18}$$

The above function is a general form of the convex objective functions considered in the sequel. We will work with special cases of (18) to obtain MSE expressions for the equal or optimal power allocation cases, in both the presence and the absence of CEE. In both cases, the power constraint is given by

$$\sum_{k=1}^{K} P_k + P_{\mathrm{trn}} = P_{\mathrm{tot}}. \tag{19}$$

V. Equal Data Power Allocation

If there is no CSI feedback from the FC, the data transmission powers can be chosen to be equal, $P_1 = P_2 = \cdots = P_K$. In this section, we focus on this *equal data power allocation* scenario. We first begin with the perfect CSI benchmark.

A. Perfect CSI at the FC Only

In what follows, we adapt the best linear unbiased estimator (BLUE) of θ with perfect CSI at the fusion center in [14] to the LMMSE estimator of θ, since this will serve as a benchmark to the CEE case we derive later. With perfect CSI at the FC, the variance of the CEE is zero ($\delta^2 = 0$) and the normalized estimated channel powers are equal to the normalized channel powers $\hat{\eta}_k = \eta_k\ \forall k$. Moreover, in this case, we do not have an optimization problem, because the data transmission powers are equal $P_k = P_{\mathrm{tot}}/K\ \forall k$, due to our equal data power allocation consideration in this section. Therefore, by substituting $\delta^2 = 0$ and $\hat{\eta}_k = \eta_k$ in (16), the MSE expression for the equal data transmission power—perfect CSI (EDTP-PCSI) case is obtained as follows:

$$D^{(\mathrm{per})}(P_{\mathrm{tot}}, K) = \sigma_\theta^2\left(1+\sum_{k=1}^{K}\frac{\gamma\eta_k}{\eta_k+\frac{K}{P_{\mathrm{tot}}}}\right)^{-1}. \tag{20}$$

For a fixed K, (20) is lower bounded by

$$\lim_{P_{\mathrm{tot}}\to\infty} D^{(\mathrm{per})}(P_{\mathrm{tot}}, K) = \frac{\sigma_\theta^2}{1+K\gamma}. \tag{21}$$

To see the asymptotic behavior for large K, note that the sum in (20) can be written as $K^{-1}\sum_{k=1}^{K}(\gamma\eta_k K/(\eta_k + (K/P_{\mathrm{tot}})))$. Since the variance of the kth term is bounded, we can use the law of large numbers in [20, p. 277] to conclude that the sum in (20) converges in mean square, and therefore also in probability. Since (20) is continuous with respect to the sum and we also have established that the sum converges in probability, we have [21, Th. C.1]

$$\begin{aligned}\lim_{K\to\infty} D^{(\mathrm{per})}(P_{\mathrm{tot}}, K) &= \frac{\sigma_\theta^2}{1+\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}E\left[\frac{\gamma\eta_k K}{\eta_k+\frac{K}{P_{\mathrm{tot}}}}\right]} \\ &= \frac{\sigma_\theta^2}{1+\zeta P_{\mathrm{tot}}/(1+\gamma^{-1})}\end{aligned} \tag{22}$$

where we also used $E[\hat{\eta}_k] = \zeta/(\gamma+1)$.

By differentiating or taking the finite difference of (20) with respect to P_{tot} and K, respectively, it is straightforward to verify that the MSE is a monotonically decreasing function of both of these variables in the equal-power perfect CSI case. We will later see that when the channel is estimated, increasing the number of sensors will not always improve performance.

1) Outage and Diversity With Perfect CSI: In [14], for the BLUE estimator, the behavior of the outage variance

$$\mathcal{P}_{D_0}^{(\mathrm{per})}(P_{\mathrm{tot}}) := Pr\left\{D^{(\mathrm{per})}(P_{\mathrm{tot}}, K) > D_0\right\} \tag{23}$$

as a function of P_{tot} was studied. It was found that $\log(\mathcal{P}_{D_0}^{(\mathrm{per})}(P_{\mathrm{tot}})) \approx -K\log(P_{\mathrm{tot}})$, which shows that the outage exhibits an "estimation diversity" order of K. The derivation in [14] relied on the assumption that the number of sensors K is large.

The diversity order for a fixed K is formally defined as

$$d = -\lim_{P_{\rm tot}\to\infty} \frac{\log\left(\mathcal{P}_{D_0}^{(\rm per)}(P_{\rm tot})\right)}{\log P_{\rm tot}}. \tag{24}$$

The diversity order quantifies the rate with which the outage goes to zero with increased total power. In Section V-B-4), we will show that even when there is channel estimation error, the diversity order remains the same as the perfect channel case.

B. Estimated CSI at the FC Only

We now consider the case where the FC has the LMMSE estimates of the channel without feeding back the CSI to the sensors that transmit with equal power. We call this case equal data transmission power-estimated CSI (EDTP-ECSI).

1) Optimum Training Power: For the estimation of the parallel channels, training power $P_{\rm trn}$ is allocated, and the remaining power $P_{\rm tot} - P_{\rm trn}$ is equally shared among the sensors $P_k = (P_{\rm tot} - P_{\rm trn})/K\ \forall k$. Therefore, substituting δ^2 in (17) into (18) and setting $P_k = (P_{\rm tot} - P_{\rm trn})/K$, we get the following objective function:

$$\begin{aligned} \min_{P_{\rm trn}} \quad & -\sum_{k=1}^{K} \frac{\gamma\hat{\eta}_k\zeta(P_{\rm tot}-P_{\rm trn})P_{\rm trn}}{\hat{\eta}_k\zeta(P_{\rm tot}-P_{\rm trn})P_{\rm trn} + K\zeta P_{\rm tot} + K^2} \\ \text{s.t.} \quad & 0 \le P_{\rm trn} \le P_{\rm tot} \end{aligned} \tag{25}$$

which is a one-dimensional constrained optimization problem. It is clear that if the training power is too small, the resulting unreliable channel estimates will increase the MSE of the source estimates. On the other hand, if the training power $P_{\rm trn}$ is too close to $P_{\rm tot}$, then each sensor transmits with a small power $P_k = (P_{\rm tot} - P_{\rm trn})/K$ and the FC does not receive much information about θ in the training phase. In what follows, we quantify this optimum value.

Theorem 1: The solution to (25) is $P_{\rm trn}^{\star} = P_{\rm tot}/2$.

Proof: Please see Appendix I.

Note that the optimum total training power $P_{\rm trn}$ is always half of the total power, regardless of the number of sensors, the total power, or the noise statistics. Substituting these optimum training and data powers $P_{\rm trn}^{\star}$ and $P_k^{\star}$ into (16) together with (17), we reach the following MSE expression:

$$D^{(\rm est)}(P_{\rm tot}, K) = \sigma_\theta^2 \left(1 + \sum_{k=1}^{K} \frac{\gamma\hat{\eta}_k}{\hat{\eta}_k + \frac{4K}{P_{\rm tot}}\left(1+\frac{K}{\zeta P_{\rm tot}}\right)}\right)^{-1} \tag{26}$$

which is the imperfect CSI counterpart of the MSE in (20). The asymptotic value of the MSE as total power goes to infinity is

$$\lim_{P_{\rm tot}\to\infty} D^{(\rm est)}(P_{\rm tot}, K) = \frac{\sigma_\theta^2}{1+K\gamma} \tag{27}$$

which is the same as (21). This makes sense because, as $P_{\rm tot} \to \infty$, the fading channel estimation also becomes perfect causing the MSE to converge to the same value as the perfect channel case. However, unlike the perfect channel case in (20), the MSE in (26) is not monotonically decreasing with K. The sum in (26) converges to 0 due to the extra factor of K in the denominator. More rigorously, note that the sum is upper bounded by $\gamma\zeta P_{\rm tot}^2(4K)^{-1}[K^{-1}\sum_{k=1}^{K}\hat{\eta}_k]$, and since the term in the brackets converges in probability due to the law of large numbers, the upper bound converges to zero in probability. Using [21, Th. C.1] and the continuity of $D^{(\rm est)}(P_{\rm tot}, K)$ with respect to the sum, (26) converges to σ_θ^2 in probability as the number of sensor goes to infinity

$$\lim_{K\to\infty} D^{(\rm est)}(P_{\rm tot}, K) = \sigma_\theta^2. \tag{28}$$

Recalling that σ_θ^2 is the worst possible variance for $\hat{\theta}$, it is clear that increasing the number of sensors does not indefinitely improve performance, but rather degrades it after a certain point. This means that a finite optimum number of sensors minimizing the MSE exists in this equal power-estimated CSI case.

2) Optimum Number of Sensors: In what follows, we obtain an approximate value for the optimum number of sensors for the equal-power case. The optimum number of sensors $K^{\star}$ must be obtained by minimizing the expected value of the MSE since it is not desirable to have the number of sensors depend on instantaneous channel realizations. Since this expectation is not tractable, we follow a heuristic approach to find an approximate value of $K^{\star}$ by minimizing a lower bound on the MSE. We note that the MSE in (26) is convex with respect to the sum. Using Jensen's inequality, we obtain

$$\begin{aligned} E&\left[D^{(\rm est)}(P_{\rm tot}, K)\right] \\ &= E_{\{\hat{\eta}_1,\ldots,\hat{\eta}_K\}}\left[\frac{\sigma_\theta^2}{1+\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k}{\hat{\eta}_k+\frac{4K}{P_{\rm tot}}\left(1+\frac{K}{\zeta P_{\rm tot}}\right)}}\right] \\ &\ge \frac{\sigma_\theta^2}{1+E_{\{\hat{\eta}_1,\ldots,\hat{\eta}_K\}}\left[\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k}{\hat{\eta}_k+\frac{4K}{P_{\rm tot}}\left(1+\frac{K}{\zeta P_{\rm tot}}\right)}\right]} \\ &= \frac{\sigma_\theta^2}{1+E_{\{\hat{\eta}_1\}}\left[\frac{K\gamma\hat{\eta}_1}{\hat{\eta}_1+\frac{4K}{P_{\rm tot}}\left(1+\frac{K}{\zeta P_{\rm tot}}\right)}\right]} \end{aligned} \tag{29}$$

where the last equality is because $\hat{\eta}_1, \ldots, \hat{\eta}_K$ are i.i.d. To minimize with respect to K, we treat it as a continuous parameter and differentiate (29) with respect to K to get the following condition:

$$E_{\{\hat{\eta}_1\}}\left[\frac{\hat{\eta}_1^2 - \frac{4K^2}{\zeta P_{\rm tot}^2}\hat{\eta}_1}{\left(\hat{\eta}_1 + \frac{4K}{P_{\rm tot}}\left(1+\frac{K}{\zeta P_{\rm tot}}\right)\right)^2}\right]\Bigg|_{K=K^{\star}} = 0. \tag{30}$$

Since the expectation above is still intractable, to find an approximation, we treat the denominator as deterministic and carry out the required expectations. The optimum number of sensors is then approximated as

$$K^{\star} \approx \text{round}\left(\frac{\zeta P_{\rm tot}}{\sqrt{2(\gamma+1)}}\right) \tag{31}$$

where the round$(\cdot)$ operator rounds the result to the nearest integer. Treating the denominator of (30) as deterministic can be justified by seeing that the second term is dominant since the first term is rarely bigger than the second term: $P[\hat{\eta}_1 > (4K^\star/P_{\rm tot})(1+(K^*/\zeta P_{\rm tot}))] = \exp(-(2+\sqrt{8(\gamma+1)}))$, which is an exponentially decaying function of the sensing signal-to-noise ratio (SNR) and can easily be shown using the fact that $\hat{\eta}_1$ is exponentially distributed. The expression in (31) reflects the dependence of $K^\star$ on the system parameters, as we elaborate below.

When the total power $P_{\rm tot}$ or the channel SNR ζ are large, the optimum number of sensors increase. This is because when $P_{\rm tot}$ is large, $P_{\rm trn} = P_{\rm tot}/2$ will also be large, leading to almost perfect channel estimates. This is in agreement with the fact that in the perfect channel case, the optimum number of sensors is infinite since in this case the performance always improves with the number of sensors. From (31), we also see that if the sensor observation SNR γ is increased, then it is better to use a smaller number of sensors. The optimum value of K, therefore, strikes a balance between averaging more observation noise (quantified by the sensing SNR γ) and increased channel estimation error variance.

As a final remark on the optimum number of sensors, we note that even though (31) is an approximation, it is quite an accurate one, as shown in the simulations. Often, the MSE curve is rather flat over a wide range of number of sensors. This especially occurs when the total power is large. In these cases, it is sufficient to find an appropriate number of sensors that will yield an MSE that is close to the minimum value. Also, in thinking about the number of sensors to be deployed, one might want to penalize larger values of K over smaller values due to cost/complexity considerations even if larger K yields smaller MSEs. Hence (31) is meant only to be a useful starting point in determining the number of sensors to be deployed.

3) Comparison of Perfect and Imperfect CSI: In order to compare the MSE performances of the perfect and the imperfect CSI cases for a fixed number of sensors K, we first note that the MSE expressions in (20) and (26) are random variables. Hence it is appropriate to derive the conditions under which the distributions of MSEs in (20) and (26) are identical. We will do this by exploiting the fact that the random variables η_k and $\hat{\eta}_k$ have identical distributions [both are exponential with mean $\zeta/(\gamma+1)$] and allow the perfect CSI case and the imperfect CSI case to have different total transmit powers $P_{\rm tot}^{(\rm per)}$ and $P_{\rm tot}^{(\rm est)}$ to see how much more power one would need in the imperfect CSI case to get the same performance. The MSE expressions in (20) and (26) have identical distributions if and only if the deterministic terms in the denominator of the sums are equal

$$\frac{K}{P_{\rm tot}^{(\rm per)}} = \frac{4K}{P_{\rm tot}^{(\rm est)}}\left(1+\frac{K}{\zeta P_{\rm tot}^{(\rm est)}}\right). \tag{32}$$

Equation (32) can be re-expressed as

$$\frac{P_{\rm tot}^{(\rm est)}}{P_{\rm tot}^{(\rm per)}} = 4\left(1+\frac{K}{\zeta P_{\rm tot}^{(\rm est)}}\right) = 2+2\sqrt{1+\frac{K}{\zeta P_{\rm tot}^{(\rm per)}}} \tag{33}$$

where the first equality is immediate from (32) and the second is obtained by expressing the ratio in terms of $P_{\rm tot}^{(\rm per)}$. Both equations in (33) yield the condition for the distribution of MSEs to be the same, which also ensures that the expected MSE (averaged over the channel distribution) will be the same. We can interpret the ratio in (33) as a *power penalty ratio* for obtaining the same performance between the perfect and imperfect channel cases. From (33), we see that $P_{\rm tot}^{(\rm per)}/P_{\rm tot}^{(\rm est)} \le 1/4$, which is a penalty of at least 6 dB for having to estimate the channel. The inequality becomes equal to 6 dB for large total powers

$$\lim_{P_{\rm tot}^{(\rm est)}\to\infty} \frac{P_{\rm tot}^{(\rm per)}}{P_{\rm tot}^{(\rm est)}} = \frac{1}{4} \tag{34}$$

which is easily seen from (33). Recalling that half of the total power has to be spared for training, we can conclude that another 3 dB is lost due to the effect of estimation error at the FC when the total power is large. Instead, if the total power $P_{\rm tot}^{(\rm per)}$ or $P_{\rm tot}^{(\rm est)}$ is small, the power loss ratio in (33) can become arbitrarily large. This indicates that the power penalty for not knowing the channel to get the same performance as the perfect channel case gets worse when the total power is smaller.

4) Outage and Diversity With Channel Estimation Error: In what follows, we wish to calculate the outage in the presence of CEE $\mathcal{P}_{D_0}^{(\rm est)}(P_{\rm tot}^{(\rm est)})$, analogous to (23). We have shown that if the total powers for the perfect CSI case are related to those of the imperfect CSI case through (32), then the distributions of $D^{(\rm per)}(P_{\rm tot}^{(\rm per)}, K)$ and $D^{(\rm est)}(P_{\rm tot}^{(\rm est)}, K)$ will be identical. This means that the corresponding outage expressions must also be the same. That is to say, $\mathcal{P}_{D_0}^{(\rm per)}(P_{\rm tot}^{(\rm per)})$ in (23) is equal to $\mathcal{P}_{D_0}^{(\rm est)}(P_{\rm tot}^{(\rm est)})$ if (32) holds. Since we would like to calculate the diversity order in the presence of CEE, we define $d^{(\rm est)}$ as

$$\begin{aligned} d^{(\rm est)} &:= -\lim_{P_{\rm tot}^{(\rm est)}\to\infty} \frac{\log\left(\mathcal{P}_{D_0}^{(\rm est)}\left(P_{\rm tot}^{(\rm est)}\right)\right)}{\log P_{\rm tot}^{(\rm est)}} \\ &= -\lim_{P_{\rm tot}^{(\rm est)}\to\infty} \frac{\log\left(\mathcal{P}_{D_0}^{(\rm per)}\left(P_{\rm tot}^{(\rm per)}\right)\right)}{\log P_{\rm tot}^{(\rm est)}} \end{aligned} \tag{35}$$

where the last equality holds because (32) ensures that $\mathcal{P}_{D_0}^{(\rm per)}(P_{\rm tot}^{(\rm per)}) = \mathcal{P}_{D_0}^{(\rm est)}(P_{\rm tot}^{(\rm est)})$. Using (34), as $P_{\rm tot}^{(\rm est)} \to \infty$, so does $P_{\rm tot}^{(\rm per)}$. Moreover, we can substitute $P_{\rm tot}^{(\rm est)}$ in the denominator of (35) using (33) to obtain

$$d^{(\rm est)} = -\lim_{P_{\rm tot}^{(\rm per)}\to\infty} \frac{\log\left(\mathcal{P}_{D_0}^{(\rm per)}\left(P_{\rm tot}^{(\rm per)}\right)\right)}{\log P_{\rm tot}^{(\rm per)} + \log\left(2+2\sqrt{1+\frac{K}{\zeta P_{\rm tot}^{(\rm per)}}}\right)}. \tag{36}$$

It is clear that the second term in the denominator has no influence asymptotically. Recalling the definition of the diversity order for the perfect channel case in (24), we conclude $d^{(\rm est)} = d$. In other words, the presence of the channel estimation error does not change the diversity order.

While the diversity order gives a good idea about the behavior of the outage, it does not lead to the best way to compute it. In fact, the outage can be expressed as

$$\mathcal{P}_{D_0}^{(\rm est)}(P_{\rm tot}) = Pr\left\{\sum_{k=1}^{K} \frac{\hat{\eta}_k\varphi}{\hat{\eta}_k + K\varphi} < \frac{\varphi}{\gamma}\left(\frac{\sigma_\theta^2}{D_0} - 1\right)\right\} \tag{37}$$

where $\varphi := (4/P_{\rm tot})(1+(K/\zeta P_{\rm tot}))$. Since the upper bound

$$\sum_{k=1}^{K}\frac{\hat{\eta}_k\varphi}{\hat{\eta}_k+K\varphi}\leq\frac{1}{K}\sum_{k=1}^{K}\hat{\eta}_k \tag{38}$$

always holds, we have the following lower bound of outage probability:

$$\mathcal{P}_{D_0}^{(\rm est)}(P_{\rm tot})\geq Pr\left\{\frac{1}{K}\sum_{k=1}^{K}\hat{\eta}_k<c\right\} \tag{39}$$

where $c := (\varphi/\gamma)((\sigma_\theta^2/D_0)-1)$ depends on $P_{\rm tot}$ only through φ. Keeping in mind that $\hat{\eta}_k$ is exponentially distributed with mean $\zeta/(\gamma+1)$, it is clear that the lower bound to the outage in (39) can be obtained using the central χ^2 distribution with $2K$ degrees of freedom

$$Pr\left\{\frac{1}{K}\sum_{k=1}^{K}\hat{\eta}_k<c\right\}=1-e^{-\frac{Kc(\gamma+1)}{\zeta}}\sum_{q=0}^{K-1}\frac{\left(\frac{Kc(\gamma+1)}{\zeta}\right)^q}{q!}. \tag{40}$$

In the simulations, we will show that this lower bound of outage probability in (39) is remarkably very tight. This is because the bound in (38) is very tight since $\hat{\eta}_k \ll K\varphi$ with very high probability for reasonably large values of K.

VI. Optimal Data Power Allocation

With CSI feedback from the FC to sensors, each sensor can adjust its own data transmission power $\{P_k\}_{k=1}^K$ using its channel estimate. Recall that in the equal power allocation strategy, the CSI knowledge is used only for source estimation at the FC, but in this case, the CSI knowledge is used also by the sensors to adjust their data transmission powers. We now consider the perfect CSI case at both the FC and the sensors. This will be a benchmark for the estimated CSI case later.

A. Perfect CSI at the FC and Sensors

With perfect CSI, variance of the CEE $\delta^2=0$ and the normalized estimated channel powers are equal to the normalized channel powers $\hat{\eta}_k = \eta_k\ \forall k$. By substituting $\delta^2=0$ and $\hat{\eta}_k=\eta_k$ in (18), the optimization problem for the optimal data transmission power-perfect CSI (ODTP-PCSI) case is obtained as follows:

$$\begin{aligned}\min_{P_1,\ldots,P_K}\quad & -\sum_{k=1}^{K}\frac{\gamma\eta_k P_k}{\eta_k P_k+1}\\ \text{s.t.}\quad & \sum_{k=1}^{K}P_k\leq P_{\rm tot}\\ & P_k\geq 0,\quad k=1,2,\ldots,K\end{aligned} \tag{41}$$

where the optimization is with respect to the transmit powers at the sensors. This problem is considered in [14] for the BLUE source estimator. Adapting to the LMMSE case, the optimum powers are given by

$$P_k^\star=\frac{\frac{1}{\sqrt{\eta_k}}}{\sum_{m\in A}\frac{1}{\sqrt{\eta_m}}}\left(P_{\rm tot}+\sum_{m\in A}\frac{1}{\eta_m}\right)-\frac{1}{\eta_k},\quad \eta_k>\tau \tag{42}$$

where $A := \{m|\eta_m>\tau\}$ is the set of *active sensors* whose data transmission powers are positive (i.e. $P_k>0$ or equivalently $\eta_k>\tau$), and the threshold value τ is given by

$$\tau=\left(\frac{\sum_{m\in A}\frac{1}{\sqrt{\eta_m}}}{P_{\rm tot}+\sum_{m\in A}\frac{1}{\eta_m}}\right)^2. \tag{43}$$

The sensors whose normalized channel powers are below the threshold level are turned off in the data transmission phase.

B. Estimated CSI at the FC and Sensors

In the optimal data transmission power-estimated CSI (ODTP-ECSI) case, we assume that parallel channels $\{g_k\}_{k=1}^K$ are estimated as described in Section III. The channel estimates are fed back from the FC to sensors in order to perform the optimal data power allocation strategy. So, after training, the remaining power $P_{\rm tot}-P_{\rm trn}$ is optimally shared among the sensors. Therefore, substituting δ^2 in (17) into (18), we get the objective function of the following convex optimization problem:

$$\begin{aligned}\min_{P_{\rm trn},P_1,\ldots,P_K}\quad & -\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k\zeta P_{\rm trn}P_k}{\hat{\eta}_k\zeta P_{\rm trn}P_k+K\zeta P_k+\zeta P_{\rm trn}+K}\\ \text{s.t.}\quad & P_{\rm trn}+\sum_{k=1}^{K}P_k\leq P_{\rm tot}\\ & P_{\rm trn}\geq 0\\ & P_k\geq 0,\quad k=1,2,\ldots,K.\end{aligned} \tag{44}$$

The solution to this problem is given in two parts in the next theorem.

Theorem 2:

i) The P_k that solves (44) is nonzero only when the corresponding $\hat{\eta}_k$ is greater than a certain threshold τ

$$\begin{aligned}P_k^\star=&\frac{\frac{\sqrt{\hat{\eta}_k}}{\hat{\eta}_k+\frac{K}{P_{\rm trn}}}}{\sum_{\ell\in B}\frac{\sqrt{\hat{\eta}_\ell}}{\hat{\eta}_\ell+\frac{K}{P_{\rm trn}}}}\left(P_{\rm tot}-P_{\rm trn}+\sum_{\ell\in B}\frac{1+\frac{K}{\zeta P_{\rm trn}}}{\hat{\eta}_\ell+\frac{K}{P_{\rm trn}}}\right)\\&-\frac{1+\frac{K}{\zeta P_{\rm trn}}}{\hat{\eta}_k+\frac{K}{P_{\rm trn}}}\quad\forall\hat{\eta}_k>\tau\end{aligned} \tag{45}$$

$$\tau=\left(\frac{\left(1+\frac{K}{\zeta P_{\rm trn}}\right)\sum_{\ell\in B}\frac{\sqrt{\hat{\eta}_\ell}}{\hat{\eta}_\ell+\frac{K}{P_{\rm trn}}}}{P_{\rm tot}-P_{\rm trn}+\left(1+\frac{K}{\zeta P_{\rm trn}}\right)\sum_{\ell\in B}\frac{1}{\hat{\eta}_\ell+\frac{K}{P_{\rm trn}}}}\right)^2. \tag{46}$$

ii) Also, the optimum training power satisfies $P_{\rm trn}^\star\geq P_{\rm tot}/2$.

Proof: Please see Appendix II.

In what follows, we compare the perfect and imperfect CSI cases when the transmit powers are optimized. In Section V-B3 for the equal power case, we found the relationship between $P_{\rm tot}^{(\rm est)}$ and $P_{\rm tot}^{(\rm per)}$ that would ensure that the MSE in the two cases would have the same distribution, for a finite number of sensors and finite total power. In what follows, we will perform a similar comparison with the assumption that the total power is large. More concretely, we want to determine the ratio of the total powers of the ODTP-PCSI and the ODTP-ECSI cases for identical distributed MSEs while total power goes to infinity. Since $P_{\rm trn}^{\star} \geq P_{\rm tot}/2$ from (62), when $P_{\rm tot} \to \infty$, the optimum training power $P_{\rm trn}^{\star} \to \infty$, which means that channel estimation error variance goes to zero $(\delta^2 \to 0)$ as seen from (17). Eventually, the normalized estimated channel powers approach the normalized true channel powers $\hat{\eta}_k \to \eta_k \ \forall k$ since channel estimates approach to true channels $\hat{g}_k \to g_k \ \forall k$. Additionally, with large total powers, all the sensors become active, $|B| \to K$, for both perfect and imperfect CSI cases because threshold levels in (43) and (46) go to zero as the total powers goes to infinity. Under these conditions, we wish to make the objective functions for perfect and imperfect channel cases in (41) and (44) equal, which ensures that the resulting solutions will be the same. The objective functions in (41) and (44) are equal if and only if

$$\frac{K}{P_{\rm trn}} + \left(1 + \frac{K}{\zeta P_{\rm trn}}\right)\frac{1}{P_k^{\rm (est)}} = \frac{1}{P_k^{\rm (per)}} \qquad (47)$$

where $P_k^{\rm (per)}$ and $P_k^{\rm (est)}$ are the powers allocated to the kth sensor in the perfect and imperfect channel cases, respectively. Keeping in mind $\sum_{k=1}^{K} P_k^{\rm (per)} = P_{\rm tot}^{\rm (per)}$, multiplying both sides of (47) by $P_k^{\rm (per)}$ and summing, we can re-express (47) as

$$\frac{P_{\rm tot}^{\rm (per)}}{P_{\rm trn}} + \left(\frac{1}{K} + \frac{1}{\zeta P_{\rm trn}}\right)\sum_{k=1}^{K}\frac{P_k^{\rm (per)}}{P_k^{\rm (est)}} = 1. \qquad (48)$$

Multiplying both sides of (48) by $P_{\rm tot}^{\rm (est)}/P_{\rm tot}^{\rm (per)}$ and inverting both sides of the equation, we have the following expression for the power ratio $P_{\rm tot}^{\rm (per)}/P_{\rm tot}^{\rm (est)}$:

$$\frac{P_{\rm tot}^{\rm (per)}}{P_{\rm tot}^{\rm (est)}} = \left(\frac{P_{\rm tot}^{\rm (est)}}{P_{\rm trn}} + \left(\frac{1}{K} + \frac{\frac{P_{\rm tot}^{\rm (est)}}{P_{\rm trn}}}{\zeta P_{\rm tot}^{\rm (est)}}\right)\sum_{k=1}^{K}\frac{\frac{P_k^{\rm (per)}}{P_{\rm tot}^{\rm (per)}}}{\frac{P_k^{\rm (est)}}{P_{\rm tot}^{\rm (est)}}}\right)^{-1}. \qquad (49)$$

Recalling $P_{\rm tot}^{\rm (per)} \to \infty$ together with $P_{\rm tot}^{\rm (est)} \to \infty$, from (42) and (45), we obtain the limit of the kth summation term in (49) as follows:

$$\lim_{P_{\rm tot}^{\rm (est)}\to\infty}\frac{\frac{P_k^{\rm (per)}}{P_{\rm tot}^{\rm (per)}}}{\frac{P_k^{\rm (est)}}{P_{\rm tot}^{\rm (est)}}} = \frac{\frac{\frac{1}{\sqrt{\eta_k}}}{\sum_{k=1}^{K}\frac{1}{\sqrt{\eta_k}}}}{\left(1 - \lim_{P_{\rm tot}^{\rm (est)}\to\infty}\frac{P_{\rm trn}}{P_{\rm tot}^{\rm (est)}}\right)\frac{\frac{1}{\sqrt{\eta_k}}}{\sum_{k=1}^{K}\frac{1}{\sqrt{\eta_k}}}} = \frac{1}{1 - \lim_{P_{\rm tot}^{\rm (est)}\to\infty}\frac{P_{\rm trn}}{P_{\rm tot}^{\rm (est)}}} = \frac{1}{1-r} \qquad (50)$$

where we used $\hat{\eta}_k \to \eta_k$ and r is defined as the asymptotic ratio between the training and the total powers of the ODTP-ECSI case given as follows:

$$r := \lim_{P_{\rm tot}^{\rm (est)}\to\infty}\frac{P_{\rm trn}}{P_{\rm tot}^{\rm (est)}}. \qquad (51)$$

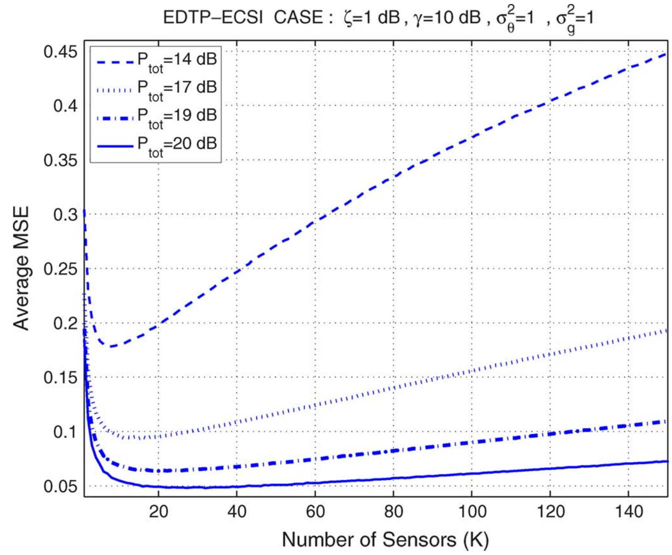


Fig. 3. Average MSE versus number of sensors for the EDTP-ECSI case.

For a given r, substituting (50) and (51) into (49) and taking the limit, the asymptotic power penalty ratio between the total powers of perfect and estimated CSI cases that make the MSEs identical is obtained as

$$\lim_{P_{\rm tot}^{\rm (est)}\to\infty}\frac{P_{\rm tot}^{\rm (per)}}{P_{\rm tot}^{\rm (est)}} = \left(\frac{1}{r} + \frac{1}{1-r}\right)^{-1} = r(1-r). \qquad (52)$$

It is clear from (52) that the maximum ratio is obtained as

$$\lim_{P_{\rm tot}^{\rm (est)}\to\infty}\frac{P_{\rm tot}^{\rm (per)}}{P_{\rm tot}^{\rm (est)}} = \frac{1}{4} \qquad (53)$$

when $r = 1/2$ (50% training power). We can thus conclude that for large total transmit powers, the penalty paid for not knowing the channel is 6 dB, which is achieved when $P_{\rm trn}$ is half the total power. This is the same result obtained for equal power allocation scenario in Section V-B3. The only major difference is that in the equal-power case, the effect of CEE could be analyzed for any finite total power, whereas the analysis here is asymptotic in the total power.

VII. Numerical Results

Fig. 2 shows the average MSE of $\hat{\theta}$ is better than that of $\hat{\theta}_{PERF}(\hat{\boldsymbol{g}})$, as also argued in Section IV. In Fig. 3, we illustrate that in the equal data power estimated CSI (EDTP-ECSI) case there is an optimum number of sensors that minimize the MSE. We also observe that the number of sensors that minimize the MSE increases as the total power $P_{\rm tot}$ increases. This is in agreement with our theoretical results in (31). A similar trend with the number of sensors is observed in Fig. 4 for the ODTP-ECSI case with a 60% training power ratio. Both Figs. 3 and 4 confirm that the MSE performance in the estimated CSI case is exhibiting a degradation beyond an optimum number of sensors. Fig. 5 in which MSE performances of the EDTP-ECSI and ODTP-ECSI cases are compared for a fixed total power show that the optimal case outperforms the equal case. Moreover, the sensitivity of the optimal power allocation case to increasing the number of sensors is also less.

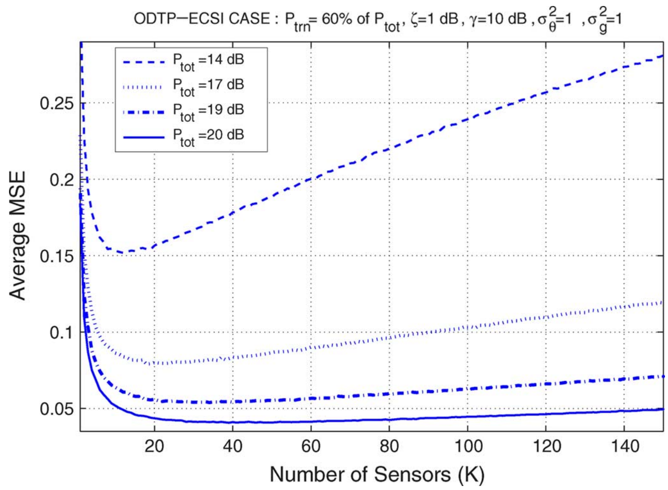


Fig. 4. Average MSE versus number of sensors for the ODTP-ECSI case.

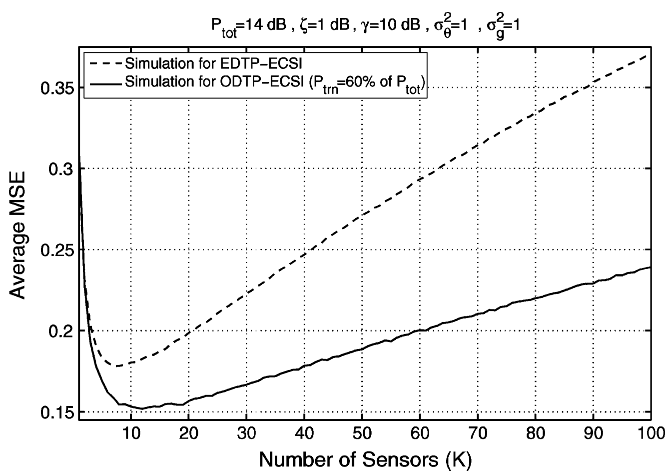


Fig. 5. MSE performance comparison of the EDTP-ECSI and ODTP-ECSI cases.

Figs. 6 and 7 show the average MSE for the equal and optimized power cases, respectively. In the curves labeled "Simulation," average MSE curves are generated by implementing the LMMSE estimators (3) and (14), where all the random variables involved in the system model are generated. These are compared with the MSE expressions obtained by averaging (16) with respect to the channel distribution. The perfect agreement in the respective curves indicates that the simulations match our MSE expressions.

In Fig. 8 for the EDTP-ECSI case, the power penalty ratios on the horizontal axis can be read off when the average MSEs are equal (the y-axis is one). We observe that the curve plotted for $P_{\rm tot}^{(\rm est)} = 20$ dB (solid curve) is crossing the horizontal line $E[D^{(\rm per)}]/E[D^{(\rm est)}] = 1$ at about $P_{\rm tot}^{(\rm per)}/P_{\rm tot}^{(\rm est)} = 0.23 < 1/4$, but the curves plotted for higher total powers ($P_{\rm tot}^{(\rm est)} = 25, 27, \text{and } 30$ dB) are crossing in the vicinity of $P_{\rm tot}^{(\rm per)}/P_{\rm tot}^{(\rm est)} = 0.25 = 1/4$ as given in (34) for the EDTP-ECSI case. In Fig. 9, the power penalty ratio $P_{\rm tot}^{(\rm per)}/P_{\rm tot}^{(\rm est)}$ is seen to be about 0.24 for the MSE performances of perfect and imperfect channel cases to be equal with $r = P_{\rm trn}/P_{\rm tot}^{(\rm est)} = 60\%$ when the sensor

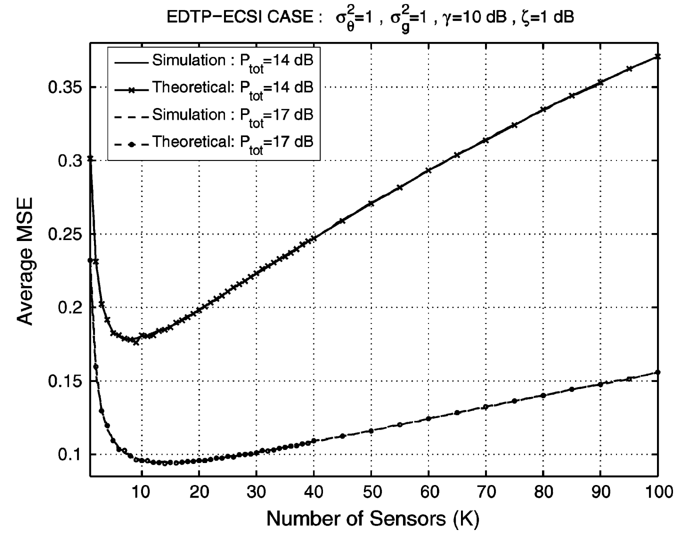


Fig. 6. Comparison of the theoretical and the simulation results for the EDTP-ECSI case.

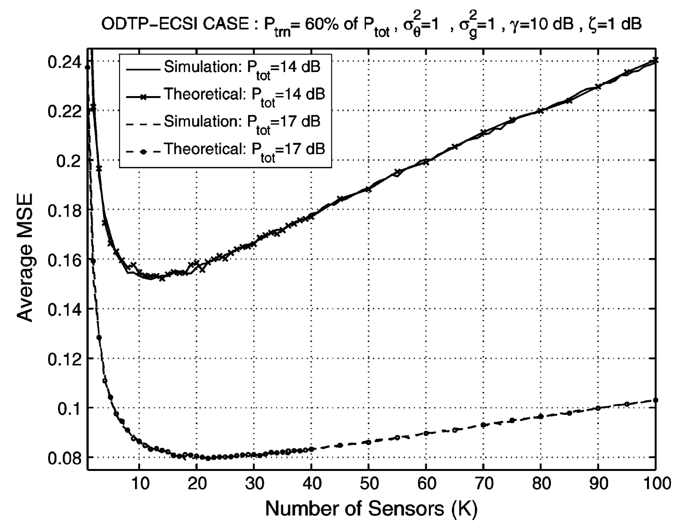


Fig. 7. Comparison of the theoretical and the simulation results for the ODTP-ECSI case.

powers are optimized. The curves in Fig. 10 are plotted for various training power ratios for the ODTP-ECSI case. In Fig. 10, we observe that the asymptotic ratios of total powers are roughly $P_{\rm tot}^{(\rm per)}/P_{\rm tot}^{(\rm est)} = 0.25, 0.24, 0.21$, and 0.16 for $r = P_{\rm trn}/P_{\rm tot}^{(\rm est)} = 0.5, 0.6, 0.7$, and 0.8, respectively, which is predicted by (52).

In Fig. 11, the simulation results for the EDTP-ECSI case indicate the accuracy of the optimum number of sensors $K^\star$ calculated from (31). It is clear that the optimum number of sensors increases with an increasing MSE performance, while accuracy of $K^\star$ decreases. However, the MSE value at $K^\star$ is close to the minimum MSE, as seen from Fig. 11.

The outage probability results are illustrated by the simulations in Fig. 12 and 13 for the EDTP case. Fig. 12 justifies that the estimation diversity order is K in the EDTP-PCSI case and remains the same despite channel estimation, as proved in Section V-B4. Fig. 13 shows that the outage probability

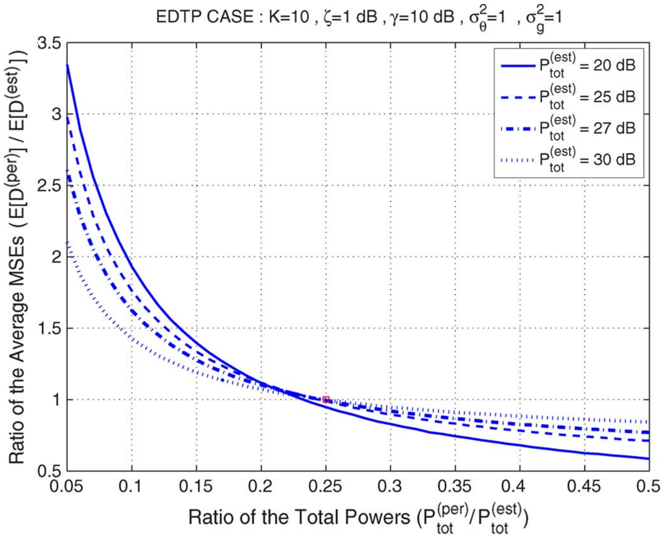


Fig. 8. Ratio of the average MSEs versus ratio of the total powers for the EDTP case.

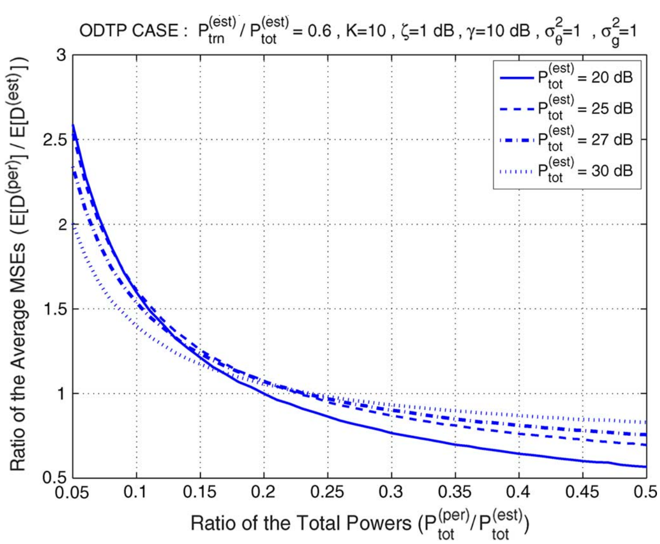


Fig. 9. Ratio of the average MSEs versus ratio of the total powers for the ODTP case.

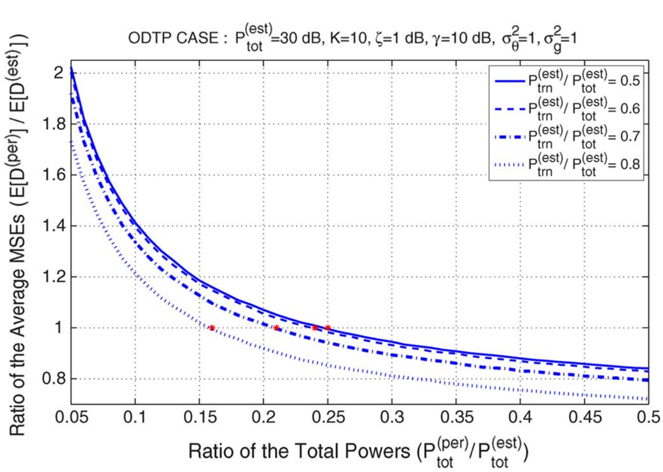


Fig. 10. Ratio of the average MSEs versus ratio of the total powers for the ODTP case for different $r = P_{\rm trn}/P_{\rm tot}^{\rm (est)}$ ratios.

lower bound obtained by using the χ^2 distribution derived in Section V-B-4) yields a very tight lower bound.

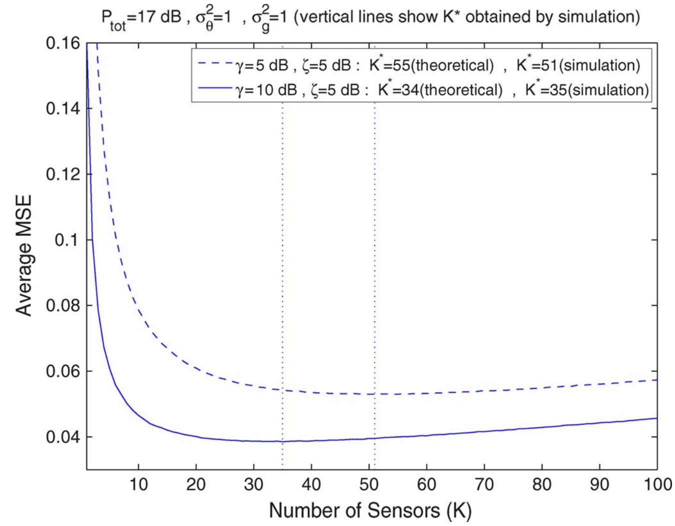


Fig. 11. Accuracy of the optimum number of sensors.

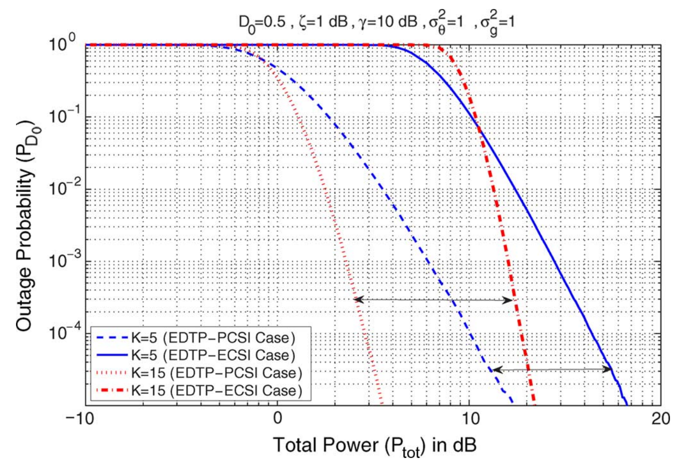


Fig. 12. Outage probability versus total power for the EDTP-ECSI and the EDTP-PCSI cases.

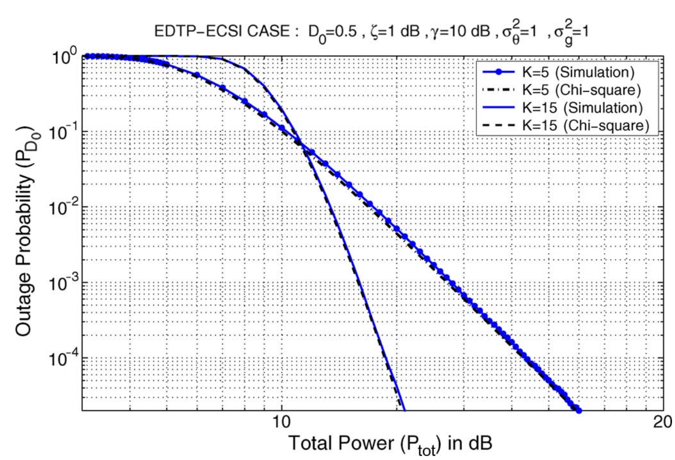


Fig. 13. Tightness of the outage probability bounds for the EDTP-ECSI case.

VIII. Conclusion

This paper studies the effect of fading channel estimation error on the performance of distributed estimators of a source θ. A two-phase approach was employed where, in the first phase, the fading coefficients are estimated and, in the second phase, these estimates and the received signal are used to estimate the source θ. Both equal power and optimized power were used for power allocation to sensors in the second phase. For the equal power case, it was found that exactly half of the total power should be used for training. The power penalty ratio in (33) to get the same performance as the perfect CSI case can be arbitrarily large but goes to 6 dB for large total powers. The diversity order, which quantifies the rate with which the variance outage goes to zero with the total power, was found to remain the same as the perfect channel case. However, unlike the perfect CSI case, for a fixed total power, increasing the number of sensors eventually degrades the MSE suggesting an optimum number. We approximated this optimum number sensors, which was shown to increase with the total power $P_{\rm tot}$ but decrease with the sensing SNR γ. The simulations verified that the approximation to the optimum number of sensors is accurate enough to yield MSEs that are close to the minimum.

Similar results were found when optimized power is used for the second phase. The optimum training power in this setting was shown to be greater than half the total power. In assessing the loss in total power due to channel estimation in this optimized sensor power setting, we used an asymptotic analysis where the total transmit power was large. It was found that the power penalty ratio between perfect and imperfect CSI cases was about 6 dB.

APPENDIX I

The Lagrangian function of the problem in (25) is given by

$$L(P_{\rm trn};\lambda_1,\lambda_2)=-\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k\zeta(P_{\rm tot}-P_{\rm trn})P_{\rm trn}}{\hat{\eta}_k\zeta(P_{\rm tot}-P_{\rm trn})P_{\rm trn}+K\zeta P_{\rm tot}+K^2} -\lambda_1(P_{\rm tot}-P_{\rm trn})-\lambda_2 P_{\rm trn}$$

with the following KKT conditions [19], [22]:

$$\begin{aligned}&-\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k\zeta(P_{\rm tot}-2P_{\rm trn})(K\zeta P_{\rm tot}+K^2)}{(\hat{\eta}_k\zeta(P_{\rm tot}-P_{\rm trn})P_{\rm trn}+K\zeta P_{\rm tot}+K^2)^2}+\lambda_1-\lambda_2\overset{(1)}{=}0\\ &\lambda_1(P_{\rm tot}-P_{\rm trn})\overset{(2)}{=}0,\quad \lambda_1\overset{(3)}{\geq}0,\quad P_{\rm trn}\overset{(4)}{\leq}P_{\rm tot}\\ &\lambda_2 P_{\rm trn}\overset{(5)}{=}0,\quad \lambda_2\overset{(6)}{\geq}0,\quad P_{\rm trn}\overset{(7)}{\geq}0.\end{aligned} \tag{54}$$

It is easy to see that $P_{\rm trn}$ cannot be equal to zero since $P_{\rm trn}=0$ implies $\lambda_1=0$ from (54.2), $\lambda_2\geq 0$ from (54.6), and $\lambda_1-\lambda_2>0$ from (54.1), which are not compatible. Similarly, $P_{\rm trn}=P_{\rm tot}$ implies $\lambda_1\geq 0$ and $\lambda_2=0$, and we have another contradiction $\lambda_1<0$ from (54.1). Therefore, $P_{\rm trn}$ must satisfy $0<P_{\rm trn}<P_{\rm tot}$, implying $\lambda_1=0$ and $\lambda_2=0$. In this case, the sum in (54.1) must be zero, in other words, $P_{\rm tot}-2P_{\rm trn}=0$. Therefore, the optimum value of the training power $P_{\rm trn}^{\star}$ is a unique solution of the system in (54)

$$P_{\rm trn}^{\star}=\frac{P_{\rm tot}}{2} \tag{55}$$

and $P_k^{\star}=(P_{\rm tot}-P_{\rm trn}^{\star})/K=P_{\rm tot}/2K\ \forall k$, since we consider equal data power case.

APPENDIX II

We now prove Theorem 2. The Lagrangian function is given by

$$\begin{aligned}&L(P_{\rm trn},P_1,\ldots,P_K;\lambda_1,\lambda_2,\mu_1,\ldots,\mu_K)\\ &\quad=-\sum_{k=1}^{K}\mu_k P_k-\sum_{k=1}^{K}\frac{\gamma\hat{\eta}_k\zeta P_{\rm trn}P_k}{\hat{\eta}_k\zeta P_{\rm trn}P_k+K\zeta P_k+\zeta P_{\rm trn}+K}\\ &\qquad-\lambda_1\left(P_{\rm tot}-P_{\rm trn}-\sum_{k=1}^{K}P_k\right)-\lambda_2 P_{\rm trn}\end{aligned}$$

and the following KKT conditions are derived from the Lagrangian function:

$$\begin{aligned}&-\sum_{k=1}^{K}\frac{\gamma K\hat{\eta}_k\zeta(\zeta P_k+1)P_k}{(\hat{\eta}_k\zeta P_{\rm trn}P_k+K\zeta P_k+\zeta P_{\rm trn}+K)^2}+\lambda_1-\lambda_2\overset{(1)}{=}0\\ &-\frac{\gamma\hat{\eta}_k\zeta(\zeta P_{\rm trn}+K)P_{\rm trn}}{(\hat{\eta}_k\zeta P_{\rm trn}P_k+K\zeta P_k+\zeta P_{\rm trn}+K)^2}+\lambda_1-\mu_k\overset{(2)}{=}0\ \forall k,\\ &\lambda_1\left(P_{\rm tot}-P_{\rm trn}-\sum_{k=1}^{K}P_k\right)\overset{(3)}{=}0,\ \lambda_1\overset{(4)}{\geq}0,\ P_{\rm trn}+\sum_{k=1}^{K}P_k\overset{(5)}{\leq}P_{\rm tot},\\ &\lambda_2 P_{\rm trn}\overset{(6)}{=}0,\quad \lambda_2\overset{(7)}{\geq}0,\quad P_{\rm trn}\overset{(8)}{\geq}0\\ &\mu_k P_k\overset{(9)}{=}0\quad\forall k,\quad \mu_k\overset{(10)}{\geq}0\quad\forall k,\quad P_k\overset{(11)}{\geq}0\quad\forall k.\end{aligned} \tag{56}$$

We will assume $0<P_{\rm trn}<P_{\rm tot}$, which means $\lambda_2=0$ as seen from (56.6). From (56.9) and (56.11), active sensors with $P_k>0$ have corresponding Lagrange multipliers $\mu_k=0$. We now want to determine how much optimum data transmission power has to be allocated for each active sensor. Condition (56.2) can be rewritten for active sensors (i.e., $P_k>0$ and $\mu_k=0$) as

$$P_k+\frac{1+\frac{K}{\zeta P_{\rm trn}}}{\hat{\eta}_k+\frac{K}{P_{\rm trn}}}=\frac{\sqrt{\hat{\eta}_k\left(1+\frac{K}{\zeta P_{\rm trn}}\right)}}{\hat{\eta}_k+\frac{K}{P_{\rm trn}}}\sqrt{\frac{\gamma}{\lambda_1}},\quad P_k>0. \tag{57}$$

Using (57), it follows that for active sensors $(P_k>0)$, we have $\hat{\eta}_k>(\lambda_1/\gamma)(1+(K/\zeta P_{\rm trn}))$. This means that $\hat{\eta}_k$ if it exceeds the following threshold:

$$\tau=\frac{\lambda_1}{\gamma}\left(1+\frac{K}{\zeta P_{\rm trn}}\right) \tag{58}$$

the kth sensor will be activated in the data transmission phase. In (57) and (58), the Lagrange multiplier λ_1 still needs to be determined. Let the active sensor set be defined as $B:=\{\ell|\hat{\eta}_\ell>\tau\}$ for the estimated CSI case. Recalling $\sum_{\ell\in B}P_\ell=P_{\rm tot}-P_{\rm trn}$, we sum both sides of (57) and use the power constraint in (19)

$$P_{\rm tot}-P_{\rm trn}+\sum_{\ell\in B}\frac{1+\frac{K}{\zeta P_{\rm trn}}}{\hat{\eta}_\ell+\frac{K}{P_{\rm trn}}}=\sqrt{\frac{\gamma}{\lambda_1}}\sum_{\ell\in B}\frac{\sqrt{\hat{\eta}_\ell\left(1+\frac{K}{\zeta P_{\rm trn}}\right)}}{\hat{\eta}_\ell+\frac{K}{P_{\rm trn}}}. \tag{59}$$

Solving for λ_1 in (59) and substituting into (57) and (58), the optimal data power $P_k^{\star}$ and the threshold level τ are obtained as (45) and (46), which establishes the first part of Theorem 2.

For the second part, we first note that from (45) and (46), the optimum data transmission power per sensor and the threshold depend on the training power $P_{\rm trn}$. We now want to find the optimum training power $P^{\star}_{\mathrm{trn}}$. Bearing in mind $\lambda_2 = 0$ and $\mu_\ell = 0$ for $\ell \in B$ for active sensors, and substituting the denominator term in (56.2) into (56.1), we get the following equation:

$$\frac{{P^{\star}_{\mathrm{trn}}}^{2}}{K} + \frac{P^{\star}_{\mathrm{trn}}}{\zeta} = \sum_{\ell \in B} P_\ell^2 + \frac{1}{\zeta}\sum_{\ell \in B} P_\ell. \tag{60}$$

Accordingly, note that the total optimal training power $P^{\star}_{\mathrm{trn}}$ depends on the power of active sensors P_ℓ. Equations (45) and (60) show that $P^{\star}_{k}$ and $P^{\star}_{\mathrm{trn}}$ depend on each other and the channel realizations. Since the total training power P_{trn} must be selected without knowing the channel realizations, we would like to bound it with a value that is not channel dependent. Toward this goal, we use Cauchy–Schwartz inequality:

$$\sum_{\ell \in B} P_\ell^2 \geq \frac{1}{|B|}\left(\sum_{\ell \in B} P_\ell\right)^2 \geq \frac{1}{K}\left(\sum_{\ell \in B} P_\ell\right)^2 \tag{61}$$

where $|B|$ is the cardinality of the set of active sensors. Substituting the above lower bound into (60) and using $\sum_{\ell \in B} P_\ell = P_{\mathrm{tot}} - P^{\star}_{\mathrm{trn}}$ on the right-hand side, we obtain the following lower bound on the optimal training power $P^{\star}_{\mathrm{trn}}$:

$$P^{\star}_{\mathrm{trn}} \geq \frac{P_{\mathrm{tot}}}{2} \tag{62}$$

which establishes the second part of Theorem 2.

References

[1] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, pp. 27–41, Jul. 2006.

[2] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Inf. Theory*, vol. 51, pp. 2210–2219, Jun. 2005.

[3] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413–422, Feb. 2006.

[4] J.-J. Xiao, Z.-Q. Luo, and G. B. Giannakis, "Performance bounds for the rate-constrained universal decentralized estimators," *IEEE Signal Process. Lett.*, vol. 14, pp. 47–50, 2007.

[5] J. Li and G. Al Regib, "Rate-constrained distributed estimation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pt. 1, pp. 1634–1643, May 2007.

[6] J. Karlsson, N. Wernersson, and M. Skoglund, "Distributed scalar quantizers for noisy channels," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP) 2007*, Honolulu, HI, Apr. 15–20, 2007.

[7] T. C. Aysal and K. E. Barner, "Constrained decentralized estimation over noisy channels for sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1398–1410, Apr. 2008.

[8] T. C. Aysal and K. E. Barner, "Blind decentralized estimation for bandwidth constrained wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1466–1471, May 2008.

[9] H. Li and J. Fang, "Distributed adaptive quantization and estimation for wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 14, pp. 669–672, Oct. 2007.

[10] S. Marano, V. Matta, and P. Willett, "Quantizer precision for distributed estimation in a large sensor network," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 4073–4078, Oct. 2006.

[11] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1147–1158, May 2003.

[12] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. 11, pp. 558–567, Oct. 1965.

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

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

[15] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.

[16] K. Liu and A. Sayeed, "Optimal distributed detection strategies for wireless sensor networks," presented at the 43rd Conf. Commun., Contr. Comput., Monticello, IL, Oct. 2004.

[17] P. Gao and C. Tepedelenlioglu, "Practical issues in parameter estimation over fading channels with TBMA wireless sensor networks," presented at the IEEE GLOBECOM 2006, San Francisco, CA, Nov. 2006.

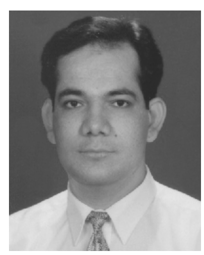
[18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[19] S. Boyd and L. Vanderberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[20] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2002.

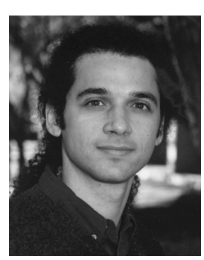
[21] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[22] H. Hindi, "A tutorial on convex optimization," in *Proc. 2004 Amer. Contr. Conf.*, Boston, MA, Jun. 30–Jul. 2 2004, pp. 3252–3265.



Habib Şenol (S'04–M'07) was born in Nazilli, Turkey, in 1971. He received the B.S. and M.S. degrees from Istanbul University, Istanbul, Turkey, in 1993 and 1999, respectively, and the Ph.D. degree from Işık University, Istanbul, in 2006, all in electronics engineering.

From 1996 to 1999, he was a Research Assistant with Istanbul University. He was a Postdoctoral Researcher in the Department of Electrical Engineering, Arizona State University, Tempe, in 2007. He is currently with the Department of Computer Engineering, Kadir Has University, Istanbul, Turkey. His research interests cover statistical signal processing, estimation and equalization algorithms for wireless communications, multicarrier (orthogonal frequency-division multiplexing) systems, distributed detection and estimation, and particle filters.



Cihan Tepedelenlioğlu (S'97–M'01) was born in Ankara, Turkey, in 1973. He received the B.S. degree (with highest honors) from Florida Institute of Technology, Melbourne, in 1995 and the M.S. degree from the University of Virginia, Charlottesville, in 1998, both in electrical engineering, and the Ph.D. degree in electrical and computer engineering from the University of Minnesota, Minneapolis.

From January 1999 to May 2001, he was a Research Assistant at the University of Minnesota. He is currently an Associate Professor of electrical engineering at Arizona State University, Tempe. He received the National Science Foundation (early) Career grant in 2001. His research interests include statistical signal processing, system identification, wireless communications, estimation and equalization algorithms for wireless systems, multiple-antenna communications, filter banks and multirate systems, OFDM, ultra-wide-band systems, distributed detection, and estimation.

Dr. Tepedelenlioğlu has served as an Associate Editor for several IEEE TRANSACTIONS, including IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE SIGNAL PROCESSING LETTERS.