Distributed Estimation with Channel Estimation Error over Orthogonal Fading Channels

Habib Şenol
Department of Computer Engineering
Kadir Has University
34083, Cibali, Istanbul, TURKEY
email: hsenol@khas.edu.tr

Cihan Tepedelenlioğlu
Department of Electrical Engineering
Arizona State University
85287 Tempe, AZ, USA
email: cihan@asu.edu

***Abstract*— We study distributed estimation of a source corrupted by an additive Gaussian noise and observed by sensors which are connected to a fusion center with unknown orthogonal (parallel) flat Rayleigh fading channels. The fading communication channels are estimated with training. Subsequently, source estimation given the channel estimates and transmitted sensor observations is performed. We consider a setting where the estimated channels are fed-back to the sensors for optimal power allocation which leads to a threshold behavior of sensors with bad channels being unused (inactive). We also show that at least half of the total power should be used for training. Simulation results corroborate our analytical findings.**

I. Introduction

In a typical wireless sensor network(WSN), a large number of sensors that each one observes the physical phenomenon represented by a parameter θ are deployed randomly in a geological zone, and transmit their observations to the fusion center (FC) over the wireless channels. FC which has less limitations in terms of processing and communication, whereas sensors have limited processing and communication capability because of their limited battery power, receives transmissions from the sensors over the wireless channels so as to combine the received signals to make inferences on the observed phenomenon.

Over the past few years, research on distributed estimation has been evolving very rapidly [1]. Universal decentralized estimators of a source over additive noise have been considered in [2], [3]. Much of the literature has focused in finite-rate transmissions of quantized sensor observations [1]. 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 in several settings described in [4], [5]. In [5], amplify-and-forward over orthogonal parallel multiple access channels(MAC) with perfect channel knowledge at the FC is considered, where increasing the number of sensors is shown to improve performance.

H. Şenol is supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) during his visit as a post-doctoral researcher at Arizona State University between Feb. 2007 and Sept. 2007.

C. Tepedelenlioglu's work is supported by NSF CAREER grant No. CCR-0133841.

Unlike our work in [6] which considers no channel status information (CSI) at the sensor side with *equal* power allocation among sensors, here we consider power optimization in this estimated CSI setting, trading off feedback complexity with MSE performance. By doing this, we follow a two-step procedure to first estimate the unknown fading channel coefficients with pilots, and use channel estimates in constructing the estimator for the source signal with linear minimum mean square error (LMMSE) estimators. We characterize the effect of channel estimation error on performance for optimal power scheduling at the sensors, and imperfect estimated channels at the FC.

We show that increasing the number of sensors will eventually lead to a degradation in performance for a fixed total power. Hence, in the absence of channel information, deploying more sensors might not necessarily lead to better performance. We also characterize the penalty paid for estimating the channel to be factor of at least 4 (6 dB).

II. System Model and Channel Estimation

In our system model, as shown in Fig.1, we consider a wireless sensor network with K sensors whose k^{th} sensor observes an unknown zero-mean complex random source signal θ with variance σ_θ^2, corrupted by a zero-mean additive complex Gaussian noise $n_k \sim \mathcal{CN}(0, \sigma_n^2)$. Since we assume the amplify-forward analog transmission scheme, the k^{th} sensor amplifies its incoming analog signal $\theta + n_k$ by a factor of α_k and transmits it on the k^{th} flat fading orthogonal channel to the fusion center (FC). In Fig.1, $g_k \sim \mathcal{CN}(0, \sigma_g^2)$ and $v_k \sim \mathcal{CN}(0, \sigma_v^2)$ are the flat fading channel gain and the channel noise of the k^{th} channel path, respectively. The amplification factor α_k varies with respect to the CSI

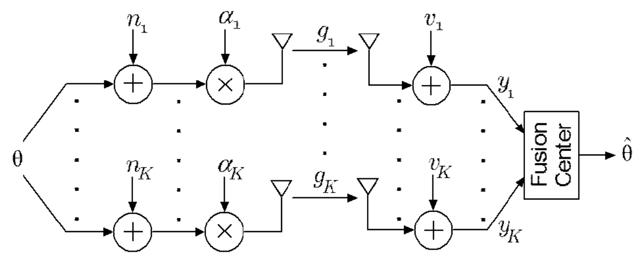

Fig. 1. Wireless Sensor Network with Orthogonal MAC Scheme

available at the sensor side. The k^{th} received signal at the FC is given as

$$y_k = g_k\, \alpha_k\, (\theta + n_k) + v_k \;, k = 1, \ldots, K \;. \qquad (1)$$

We will consider this receive model to estimate the source signal θ. Our two-step strategy, as illustrated in Fig.2, is first to estimate parallel channels, and then estimate the source signal given the channel estimates. We will use a LMMSE approach [7] for both steps. In the first phase, the sensors send training symbols of total power P_{trn} to estimate the parallel channels $\{g_k\}_{k=1}^{K}$. In the second phase the sensors transmit their amplified data, which bear information about θ, with the optimally shared powers $\{P_k\}_{k=1}^{K}$ among the sensors with respect to CSI. Note that the total power in the two phases add to P_{tot}. 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. To estimate

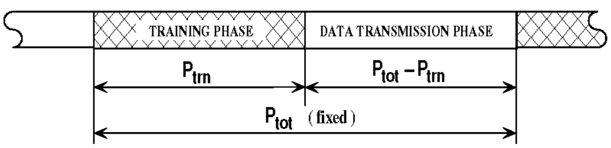

Fig. 2. Training and Data Transmission Phases

the parallel fading channels $\{g_k\}_{k=1}^{K}$ in the training phase, we consider pilot-based channel estimation as illustrated in Fig.3, where each sensor sends a pilot symbol to the FC over its own fading channel. The receive model for a pilot s transmitted over the k^{th} channel is

$$x_k = g_k\, s + \nu_k \;, \qquad (2)$$

where x_k is the received signal over k^{th} channel and ν_k is zero-mean additive complex Gaussian channel noise, $\nu_k \sim \mathcal{CN}(0, \sigma_v^2)$. Since the total transmitted training power is P_{trn}, we have $P_{trn} = K|s|^2$. According to our observation model

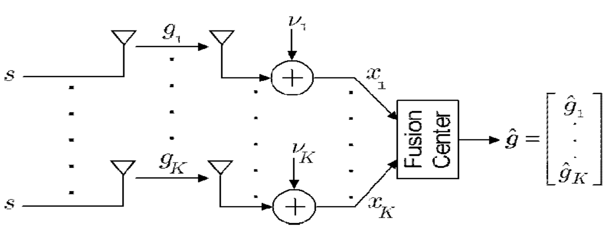

Fig. 3. Channel Estimation Scheme for Orthogonal MAC Channels

in (2), the linear minimum mean square error (LMMSE) estimate $\hat{g}_k$ of the channel g_k is given as follows [7]

$$\hat{g}_k = \frac{E_{\{g_k, x_k\}}[\, g_k\, x_k^* \,]}{E_{\{x_k\}}[\, |x_k|^2 \,]}\, x_k = \frac{\sigma_g^2\, s^*}{\sigma_v^2 + \sigma_g^2\, |s|^2}\, x_k \;, \qquad (3)$$

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

$$\delta^2 = \left(\frac{1}{\sigma_g^2} + \frac{|s|^2}{\sigma_v^2} \right)^{-1} = \frac{\sigma_v^2\, \sigma_g^2}{\sigma_v^2 + \sigma_g^2\, |s|^2} \;. \qquad (4)$$

III. Source Estimation and Power Allocation

After estimation of the unknown fading channels in the first phase, in the second phase, we estimate the source signal θ by choosing an LMMSE type source estimator given the channel estimates $\{\hat{g}_k\}_{k=1}^{K}$ in (3), and the received signal $y_1, \ldots, y_K$ in (1). In other words, we obtain the source estimator $\hat{\theta}$ in the presence of channel estimation error (CEE). The resulting MSE of source estimator will be our figure of merit. Exploiting the orthogonality principle of the LMMSE estimator, it is possible to give the minimum MSE in the presence of CEE as follows [8]

$$D = \sigma_\theta^2 \left(1 + \sum_{k=1}^{K} \frac{\gamma\, \hat{\eta}_k\, (\sigma_g^2 - \delta^2) P_k}{\left(\hat{\eta}_k\, (\sigma_g^2 - \delta^2) + \zeta \delta^2 \right) P_k + \sigma_g^2} \right)^{-1} \qquad (5)$$

with the following definitions:

Variance of $\hat{g}_k$	$\sigma_{\hat{g}}^2 = \sigma_g^2 - \delta^2$
Observation SNR	$\gamma := \sigma_\theta^2 / \sigma_n^2$
Total training power	$P_{trn} := K\, \vert s \vert^2$
k^{th} sensor data transmission power	$P_k := \frac{\vert\alpha_k\vert^2\, \sigma_\theta^2\, \gamma}{\gamma + 1}$
Channel SNR	$\zeta := \sigma_g^2 / \sigma_v^2$
k^{th} estimated channel power	$\hat{\eta}_k := \frac{\zeta\, \vert\hat{g}_k\vert^2}{(\sigma_g^2 - \delta^2)\, (\gamma+1)}$
k^{th} channel power	$\eta_k := \frac{\zeta\, \vert g_k \vert^2}{\sigma_g^2\, (\gamma+1)}$

Here, we express the channel estimator variance δ^2 using (4) and $P_{trn} = K|s|^2$ as

$$\delta^2 = (K \sigma_g^2) / (K + \zeta P_{trn}). \qquad (6)$$

Substituting (6) into (5), it is straightforward to verify that (5) is a convex function of $\{P_{trn}, P_1, \cdots, P_K\}$ by taking the second derivative. For the purposes of optimization of the MSE in (5) with respect to $\{P_{trn}, P_1, \cdots, P_K\}$, it suffices to work with

$$- \sum_{k=1}^{K} \frac{\gamma\, \hat{\eta}_k\, (\sigma_g^2 - \delta^2) P_k}{\left(\hat{\eta}_k\, (\sigma_g^2 - \delta^2) + \zeta \delta^2 \right) P_k + \sigma_g^2} \;. \qquad (7)$$

The above function is a general form of the convex objective functions considered in the sequel. We will work with special cases of (7) to obtain MSE expressions both in the presence and absence of CEE. Before we optimize the training power, we will briefly review the perfect CSI case.

A. Perfect CSI Case

With perfect CSI both at the FC and sensor side, the variance of the CEE is zero, $\delta^2 = 0$, and the normalized estimated channel powers are equal to the normalized channel powers $\hat{\eta}_k = \eta_k \; \forall k$. By substituting $\delta^2 = 0$ and $\hat{\eta}_k = \eta_k$ in (7), the optimization problem for the perfect CSI case is obtained as follows

$$\begin{aligned} \min_{P_1, \cdots, P_K} \quad & - \sum_{k=1}^{K} \frac{\gamma\, \eta_k P_k}{\eta_k P_k + 1} \\ \text{s.t.} \quad & \sum_{k=1}^{K} P_k \le P_{tot} \\ & P_k \ge 0 \quad , \quad k = 1, 2, \cdots, K \;, \end{aligned} \qquad (8)$$

where the optimization is with respect to the transmit powers at the sensors. This problem is considered in [5] for the best linear unbiased estimator (BLUE). Adapting to the LMMSE case, the optimum powers are given by

$$P_k^\star = \frac{\frac{1}{\sqrt{\eta_k}}}{\sum_{m\in A}\frac{1}{\sqrt{\eta_m}}}\left(P_{tot} + \sum_{m\in A}\frac{1}{\eta_m}\right) - \frac{1}{\eta_k}\ ,\ \eta_k > \tau^{(per)} \quad (9)$$

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

$$\tau^{(per)} = \left(\frac{\sum_{m\in A}\frac{1}{\sqrt{\eta_m}}}{P_{tot} + \sum_{m\in A}\frac{1}{\eta_m}}\right)^2 . \quad (10)$$

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

B. Estimated CSI Case

In the estimated CSI case, we assume that parallel channels $\{g_k\}_{k=1}^K$ are estimated at the FC and 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 $P_{tot} - P_{trn}$ is optimally shared among the sensors. Therefore, substituting (6) into (7) we get the objective function of the following convex optimization problem

$$\begin{aligned}
\min_{P_{trn},P_1,\cdots,P_K} \quad & -\sum_{k=1}^{K}\frac{\gamma\,\hat{\eta}_k\,\zeta P_{trn}P_k}{\hat{\eta}_k\,\zeta P_{trn}P_k + K\zeta P_k + \zeta P_{trn} + K}\\
\textit{s.t.} \quad & P_{trn} + \sum_{k=1}^{K} P_k \le P_{tot}\\
& P_{trn} \ge 0\\
& P_k \ge 0 \quad , \quad k = 1, 2, \cdots, K\ . \quad (11)
\end{aligned}$$

Now we solve the problem in (11). The Lagrangian function is given by

$$\begin{aligned}
L \;=\; & -\sum_{k=1}^{K}\frac{\gamma\,\hat{\eta}_k\,\zeta P_{trn}P_k}{\hat{\eta}_k\,\zeta P_{trn}P_k + K\zeta P_k + \zeta P_{trn} + K}\\
& -\lambda_1(P_{tot} - Pt - \sum_{k=1}^{K} P_k) - \lambda_2 P_{trn} - \sum_{k=1}^{K}\mu_k P_k \quad ,
\end{aligned}$$

and the following Karush-Kuhn-Tucker conditions are derived from the Lagrangian function [9]:

$$\lambda_1 - \lambda_2 - \sum_{k=1}^{K}\frac{\gamma\,K\hat{\eta}_k\,\zeta\,(\zeta P_k + 1)\,P_k}{(\hat{\eta}_k\,\zeta P_{trn}P_k + K\zeta P_k + \zeta P_{trn} + K)^2} \overset{(1)}{=} 0\ ,$$

$$\lambda_1 - \mu_k - \frac{\gamma\hat{\eta}_k\zeta\,(\zeta P_{trn} + K)\,P_{trn}}{(\hat{\eta}_k\,\zeta P_{trn}P_k + K\zeta P_k + \zeta P_{trn} + K)^2} \overset{(2)}{=} 0\ \ \forall k\ ,$$

$$\lambda_1(P_{tot} - P_{trn} - \sum_{k=1}^{K} P_k) \overset{(3)}{=} 0\ ,\quad \lambda_1 \overset{(4)}{\ge} 0\ ,\quad P_{trn} + \sum_{k=1}^{K} P_k \overset{(5)}{\le} P_{tot}$$

$$\lambda_2 P_{trn} \overset{(6)}{=} 0\ ,\quad \lambda_2 \overset{(7)}{\ge} 0\ ,\quad P_{trn} \overset{(8)}{\ge} 0\ ,$$

$$\mu_k P_k \overset{(9)}{=} 0\ \forall k\ ,\quad \mu_k \overset{(10)}{\ge} 0\ \forall k\ ,\quad P_k \overset{(11)}{\ge} 0\ \forall k\ . \quad (12)$$

where (12.1) and (12.2) are obtained by $\partial L/\partial P_{trn} = 0$ and $\partial L/\partial P_k = 0$, respectively. We will assume $0 < P_{trn} < P_{tot}$ which means $\lambda_2 = 0$ as seen from the condition (12.6). From conditions (12.9) and (12.11) 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. The condition (12.2) can be rewritten for active sensors (i.e., $P_k > 0$ and $\mu_k = 0$) as

$$P_k + \frac{1 + \frac{K}{\zeta P_{trn}}}{\hat{\eta}_k + \frac{K}{P_{trn}}} = \frac{\sqrt{\hat{\eta}_k\,(1 + \frac{K}{\zeta P_{trn}})}}{\hat{\eta}_k + \frac{K}{P_{trn}}}\sqrt{\frac{\gamma}{\lambda_1}}\ \ ,\ \ P_k > 0\ . \quad (13)$$

Using (13), it follows that for active sensors ($P_k > 0$) we have $\hat{\eta}_k > \frac{\lambda_1}{\gamma}\left(1 + \frac{K}{\zeta P_{trn}}\right)$. This means that $\hat{\eta}_k$ if exceeds the following threshold

$$\tau^{(est)} = \frac{\lambda_1}{\gamma}\left(1 + \frac{K}{\zeta P_{trn}}\right)\ , \quad (14)$$

the k^{th} sensor will be activated in the data transmission phase. In the equations (13) and (14), 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 $\sum_{\ell\in B} P_\ell = P_{tot} - P_{trn}$, we sum both sides of (13)

$$P_{tot} - P_{trn} + \sum_{\ell\in B}\frac{1 + \frac{K}{\zeta P_{trn}}}{\hat{\eta}_\ell + \frac{K}{P_{trn}}} = \sqrt{\frac{\gamma}{\lambda_1}}\sum_{\ell\in B}\frac{\sqrt{\hat{\eta}_\ell\,(1 + \frac{K}{\zeta P_{trn}})}}{\hat{\eta}_\ell + \frac{K}{P_{trn}}}\ . \quad (15)$$

Solving for λ_1 in (15) and substituting into (13) and (14) the optimal data power $P_k^\star$ and the threshold level $\tau^{(est)}$ are obtained as

$$\begin{aligned}
P_k^\star \;=\; & \frac{\frac{\sqrt{\hat{\eta}_k}}{\hat{\eta}_k + \frac{K}{P_{trn}}}}{\sum_{\ell\in B}\frac{\sqrt{\hat{\eta}_\ell}}{\hat{\eta}_\ell + \frac{K}{P_{trn}}}}\left(P_{tot} - P_{trn} + \sum_{\ell\in B}\frac{1 + \frac{K}{\zeta P_{trn}}}{\hat{\eta}_\ell + \frac{K}{P_{trn}}}\right)\\
& - \frac{1 + \frac{K}{\zeta P_{trn}}}{\hat{\eta}_k + \frac{K}{P_{trn}}}\quad ,\quad \forall\,\hat{\eta}_k > \tau^{(est)} \quad (16)
\end{aligned}$$

and

$$\tau^{(est)} = \left(\frac{(1 + \frac{K}{\zeta P_{trn}})\sum_{\ell\in B}\frac{\sqrt{\hat{\eta}_\ell}}{\hat{\eta}_\ell + \frac{K}{P_{trn}}}}{P_{tot} - P_{trn} + (1 + \frac{K}{\zeta P_{trn}})\sum_{\ell\in B}\frac{1}{\hat{\eta}_\ell + \frac{K}{P_{trn}}}}\right)^2, \quad (17)$$

respectively. As seen from (16) and (17), the optimum data transmission power per sensor and the threshold depend on the training power P_{trn}. We now want to find the optimum training power $P_{trn}^\star$. Substituting (12.2) into (12.1), we get the following equation,

$$\frac{{P_{trn}^\star}^2}{K} + \frac{P_{trn}^\star}{\zeta} = \sum_{\ell\in B} P_\ell^2 + \frac{1}{\zeta}\sum_{\ell\in B} P_\ell\ , \quad (18)$$

and note that the total optimal training power $P^\star_{trn}$ depends on the power of active sensors, P_ℓ . Equations (16) and (18) show that $P^\star_k$ and $P^\star_{trn}$ 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,

$$\sum_{\ell \in B} P_\ell^2 \geq \frac{1}{|B|}\left(\sum_{\ell \in B} P_\ell\right)^2 \geq \frac{1}{K}\left(\sum_{\ell \in B} P_\ell\right)^2 , \qquad (19)$$

where $|B|$ is the cardinality of the set of active sensors. Substituting the above lower bound into (18), and using $\sum_{\ell \in B} P_\ell = P_{tot} - P^\star_{trn}$ on the right hand side, we obtain the following lower bound on the optimal training power $P^\star_{trn}$:

$$P^\star_{trn} \geq \frac{P_{tot}}{2} \quad , \qquad (20)$$

which establishes that the total training power should be chosen at least half of the total power.

C. Comparison of Perfect and Estimated CSI Cases

In subsection, we want to find the relationship between total power of the perfect and estimated CSI cases, $P^{(per)}_{tot}$ and $P^{(est)}_{tot}$, that would ensure that the MSE in the two cases would have the same distribution, for a finite number of sensors and large total power. More concretely, our goal is to determine the ratio of the total powers of the perfect and imperfect CSI cases for identical distributed MSEs while total power goes to infinity. Since $P^\star_{trn} \geq P_{tot}/2$ from (20), when $P_{tot} \to \infty$, the optimum training power $P^\star_{trn} \to \infty$ which means that channel estimation error variance goes to zero ($\delta^2 \to 0$) as seen from (6). Eventually, the estimated channel powers approach to the 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, for both perfect and estimated CSI cases because threshold levels in (10) and (17) go to zero as the total powers goes to infinity. Under these conditions, we wish to make the objective functions for perfect and estimated CSI cases in (8) and (11) equal which ensures that the resulting solutions will be the same. The objective functions in (8) and (11) are equal if and only if

$$\frac{K}{P_{trn}} + \left(1 + \frac{K}{\zeta P_{trn}}\right)\frac{1}{P_k{}^{(est)}} = \frac{1}{P_k{}^{(per)}} , \qquad (21)$$

where $P_k{}^{(per)}$ and $P_k{}^{(est)}$ are the powers allocated to the k^{th} sensor in the perfect and imperfect channel cases, respectively. Keeping in mind $\sum_{k=1}^{K} P_k{}^{(per)} = P^{(per)}_{tot}$, multiplying both sides of (21) by $P_k{}^{(per)}$ and summing, we can reexpress (21) as

$$\frac{P^{(per)}_{tot}}{P_{trn}} + \left(\frac{1}{K} + \frac{1}{\zeta P_{trn}}\right)\sum_{k=1}^{K}\frac{P_k{}^{(per)}}{P_k{}^{(est)}} = 1 . \qquad (22)$$

Multiplying both sides of (22) by $P^{(est)}_{tot}/P^{(per)}_{tot}$ and inverting both sides of the equation, we have the following expression for the power ratio $P^{(per)}_{tot}/P^{(est)}_{tot}$

$$\frac{P^{(per)}_{tot}}{P^{(est)}_{tot}} = \left(\frac{P^{(est)}_{tot}}{P_{trn}} + \left(\frac{1}{K} + \frac{\frac{P^{(est)}_{tot}}{P_{trn}}}{\zeta P^{(est)}_{tot}}\right)\sum_{k=1}^{K}\frac{\frac{P_k{}^{(per)}}{P^{(per)}_{tot}}}{\frac{P_k{}^{(est)}}{P^{(est)}_{tot}}}\right)^{-1} . \qquad (23)$$

Recalling $P^{(per)}_{tot} \to \infty$ together with $P^{(est)}_{tot} \to \infty$, from (9) and (16) we obtain the limit of the k^{th} summation term in (23) as follows

$$\lim_{P^{(est)}_{tot} \to \infty} \frac{\frac{P_k{}^{(per)}}{P^{(per)}_{tot}}}{\frac{P_k{}^{(est)}}{P^{(est)}_{tot}}} = \frac{1}{1 - \lim_{P^{(est)}_{tot} \to \infty} \frac{P_{trn}}{P^{(est)}_{tot}}} = \frac{1}{1-r} , \qquad (24)$$

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 estimated CSI case as follows

$$r := \lim_{P^{(est)}_{tot} \to \infty} \frac{P_{trn}}{P^{(est)}_{tot}} . \qquad (25)$$

For a given r, substituting (24) and (25) into (23) 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

$$\lim_{P^{(est)}_{tot} \to \infty} \frac{P^{(per)}_{tot}}{P^{(est)}_{tot}} = \left(\frac{1}{r} + \frac{1}{1-r}\right)^{-1} = r(1-r) \quad , \qquad (26)$$

It is clear from (26) that the maximum ratio is obtained as

$$\lim_{P^{(est)}_{tot} \to \infty} \frac{P^{(per)}_{tot}}{P^{(est)}_{tot}} = \frac{1}{4} \qquad (27)$$

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_{trn} is half the total power.

IV. Numerical Results

In Fig.4, we illustrate that 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_{tot} increases for the estimated CSI case with a 60% training power ratio. Fig.4 confirms that the MSE performance in the estimated CSI case is exhibiting a degradation beyond an optimum number of sensors.

In Fig.5, the power penalty ratios on the horizontal axis can be read off when the average MSEs are equal (the y-axis is one), and the power penalty ratio $P^{(per)}_{tot}/P^{(est)}_{tot}$ is seen to be about 0.24 for the MSE performances of perfect and imperfect channel cases to be equal with $r = P_{trn}/P^{(est)}_{tot} = 60\%$ when the sensor powers are optimized.

The curves in Fig.6 are plotted for various training power ratios for the estmated CSI case. In this figure, we observe that the asymptotic ratios of total powers are roughly $P_{tot}^{(per)}/P_{tot}^{(est)} = 0.25, 0.24, 0.21$ and 0.16 for $r = P_{trn}/P_{tot}^{(est)} = 0.5, 0.6, 0.7$ and 0.8, respectively, which is predicted by (26).

V. Conclusions

In this work, we study 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. 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.

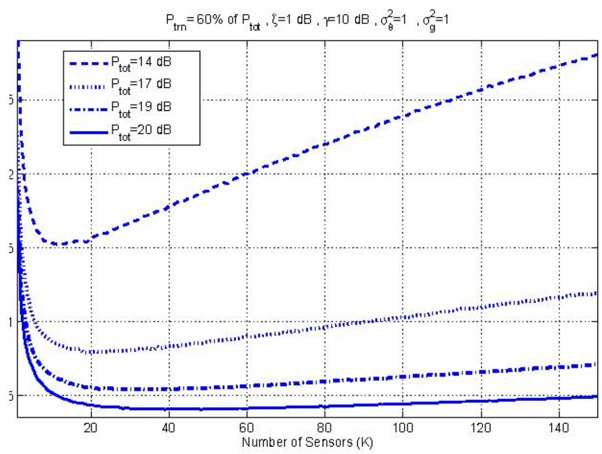

Fig. 4. Average MSE vs. number of sensors for the estimated CSI case

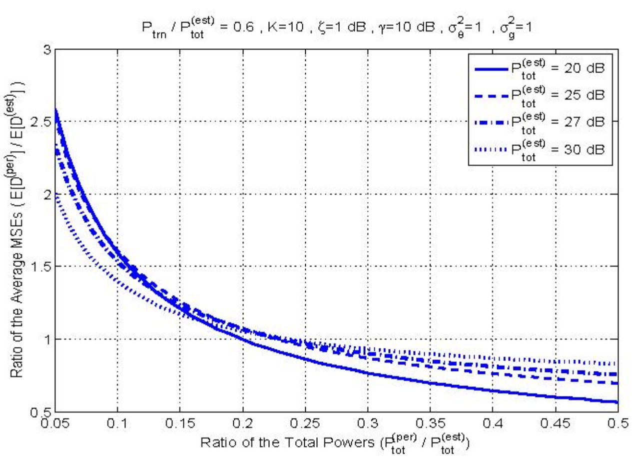

Fig. 5. Ratio of the average MSEs vs. ratio of the total powers

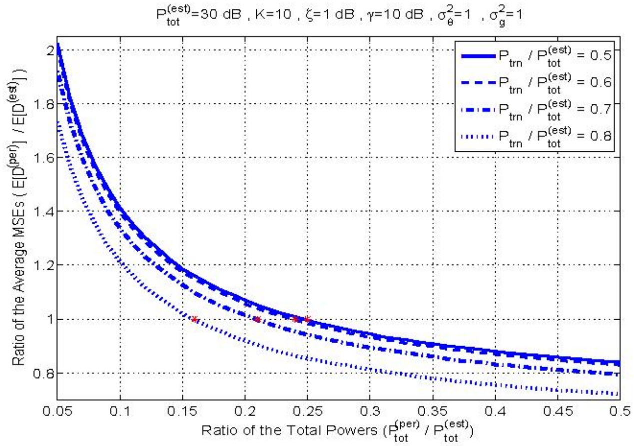

Fig. 6. Ratio of the average MSEs vs. ratio of the total powers for different $r = P_{trn}/P_{tot}^{(est)}$ ratios

References

[1] Jin-Jun Xiao, A. Ribeiro, Zhi-Quan Luo, and G.B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27–41, July 2006.

[2] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2210–2219, June 2005.

[3] Jin-Jun Xiao, Shuguang Cui, Zhi-Quan Luo, and A.J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, February 2006.

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

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

[6] H. Senol and C. Tepedelenlioglu, "Distributed estimation over parallel fading channels with channel estimation errors," in *ICASSP 2008*, Las Vegas, NV, (under review).

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

[8] H. Senol and C. Tepedelenlioglu, "Distributed estimation over unknown parallel fading channels," *IEEE Transactions on Signal Processing*, (under review).

[9] S. Boyd and L. Vanderberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2003.