Fusion and Diversity Trade-offs in Cooperative Estimation over Fading Channels

Mehrzad Malmirchegini and Yasamin Mostofi
Cooperative Network Lab, Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, New Mexico 87131, USA
Email: {mehrzad,ymostofi}@ece.unm.edu

***Abstract*—In this paper we consider a network of distributed sensors that are trying to measure a parameter of interest cooperatively, by exchanging their acquired information repeatedly over fading channels. We consider different ways of using the available bandwidth, in terms of what each node can send to its neighbors. More specifically, we characterize the impact of local fusion and show how it is a suitable policy when graph connectivity is low. When poor link qualities are the main bottleneck, on the other hand, we show how a diversity approach can be more beneficial. The proposed framework highlights the underlying tradeoffs between fusion and diversity approaches in cooperative networks. It furthermore explores the impact of multiple sensing on the overall performance.**

***Index Terms*—Fading channels, Best Linear Unbiased Estimation (BLUE), Graph connectivity**

I. Introduction

In recent years, there has been significant interest in cooperative sensing, estimation/detection and control. Such problems arise in many different areas such as environmental monitoring, surveillance and security, smart homes and factories and target tracking. Distributed detection and estimation, in particular, is an integral part of such cooperative operations and has received considerable attention in recent years. Consensus problems arise when the agents need to reach an agreement on the value of a parameter and can help to improve the overall estimation/detection quality through repeated communications. *Detection Consensus* refers to the problems in which the parameter of interest takes values from a finite known set [1]. *Estimation consensus*, on the other hand, refers to the problems where the parameter of interest can take values over an infinite set or an unknown finite set [1].

Estimation consensus problems received considerable attention over the past few years. In particular, control community has applied tools from algebraic graph theory and advanced matrix analysis to characterize estimation consensus problems over graphs that are not fully connected [2]. [3] provides a comprehensive survey of the literature on such consensus problems. Recently, there has been considerable interest in estimation consensus over unreliable links from signal processing and communication community. In [4], authors considered the distributed average consensus problem and characterized the weight matrix to achieve the fastest possible convergence rate. In [5], authors extended the previous work to include modeling uncertainty. In [6], a simple distributed iterative scheme was proposed to compute the maximum-likelihood estimate of the parameter of interest. In [7], the authors developed distributed consensus-seeking algorithms for estimation of deterministic signals. The uncertainty in the exchanged information was also considered in [8] by using a Kalman filtering approach. In [9], the authors consider the average consensus problem under quantization constraints. In the context of detection consensus, [1], [10] and [13] considered a binary consensus problem (consensus over a parameter that can only take two values) over AWGN channels. In [11] we extended the binary consensus problem to the fading environments where the impact of link uncertainty on group consensus is characterized using a probabilistic approach.

In this paper, we consider the scenario where a number of sensors are trying to estimate a parameter of interest cooperatively by communicating over fading channels. We devise different iterative strategies in terms of local decision making and information processing. More specifically, we consider two strategies based on fusion and diversity and characterize their performance mathematically. We show the impact of graph connectivity and link quality on both approaches and highlight the underlying tradeoffs. Finally, our simulation results confirm the proposed framework.

II. Problem Formulation

Consider a cooperative network of M sensors that are trying to measure a parameter in their environment. Let ω_i represent the sensing noise of the ith node, which can be modeled as a zero-mean Gaussian variable with the variance $\sigma^2_{\omega_i}$. Let x_i represent the sensed value of the ith node. We have the following: $x_i = C + \omega_i$ for $1 \leq i \leq M$, where C represents the parameter of interest. Different nodes are assumed to make independent measurements, i.e. ω_i and ω_j are taken to be uncorrelated for $i \neq j$. Each agent sends its noisy measurement to those nodes that it can communicate with, over Rayleigh fading channels. Let $r_{j,i}$ represent the fading coefficient for the link from node j to node i. The receiver will learn $r_{j,i}$ and use it in the estimation process. We take the underlying communication network as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \cdots, M\}$ is the vertex set and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents the edge set. We assume that if $(i, j) \in \mathcal{E}$, then $(j, i) \in \mathcal{E}$, namely the graph is undirected (still $r_{j,i}$ and $r_{i,j}$ could be different). The graph $\mathcal{G}$ is connected if

This work was supported in part by the Defense Threat Reduction Agency through grant HDTRA1-07-1-0036.

there is a sequence of edges in $\mathcal{E}$ that can be traversed to go from any vertex to any other. In the following we denote the set of neighbors of node $i \in \mathcal{V}$ as $\mathcal{N}_i = \{j \in \mathcal{V} : (j,i) \in \mathcal{E}\} \cup \{i\}$, including itself.

Let $n_{j,i}$ represent the receiver noise in the transmission from the jth node to the ith one, where $j \in \mathcal{N}_i$ and $j \neq i$. $n_{j,i}$ is a zero-mean Gaussian variable with variance of $\sigma^2_{j,i}$. Let $y_{j,i}$ represent the reception of the ith node from the transmission of the jth one. Using the analog model for relay channels in amplify-and-forward mode [12] we have:

$$y_{j,i} = r_{j,i} x_j + n_{j,i} \quad \text{for} \quad j \in \mathcal{N}_i,\ j \neq i, \tag{1}$$

and $y_{i,i} = x_i$. We assume that each node has the corresponding knowledge of $\sigma_{j,i}$s and $r_{j,i}$s and uses it in the estimation process. Moreover, we assume that noise and fading coefficients are uncorrelated from link to link.

III. Scenario 1: Single Sensing and Multiple Communication (SSMC)

In this part we consider the case where each node senses the parameter of interest only once. It then keeps on communicating its value with its neighbors in order to improve its estimation quality. Let $\mathcal{O}^i = \{o^i_1, \cdots, o^i_{|\mathcal{N}_i|}\}$ denote the ordered set of the neighbors of node i. We define $\mathbf{y}_i = [y_{o^i_1,i}, \cdots, y_{o^i_{|\mathcal{N}_i|},i}]^T$, $R_i = \text{diag}(r_{o^i_1,i}, \cdots, r_{o^i_{|\mathcal{N}_i|},i})$, $\mathbf{x}_i = [x_{o^i_1}, \cdots, x_{o^i_{|\mathcal{N}_i|}}]^T$, $\mathbf{n}_i = [n_{o^i_1,i}, \cdots, n_{o^i_{|\mathcal{N}_i|},i}]^T$, $\boldsymbol{\omega}_i = [\omega_{o^i_1}, \cdots, \omega_{o^i_{|\mathcal{N}_i|}}]^T$ where $r_{i,i} = 1$, $n_{i,i} = 0$ and diag($\mathbf{x}$) is a diagonal matrix with the elements of vector $\mathbf{x}$ on its main diagonal. The sensing and communication equations for the ith node will be as follows:

$$\mathbf{x}_i = C\vec{1} + \boldsymbol{\omega}_i \text{ and } \mathbf{y}_i = R_i \mathbf{x}_i + \mathbf{n}_i, \tag{2}$$

where $\Omega_i = E\{\boldsymbol{\omega}_i \boldsymbol{\omega}_i^T\}$ and $N_i = E\{\mathbf{n}_i \mathbf{n}_i^T\}$. The ith node tries to estimate the common value by using the Best Linear Unbiased Estimation (BLUE) of the received information:

$$\hat{\phi}_i = \boldsymbol{\alpha}_i^T \mathbf{y}_i, \tag{3}$$

where $\hat{\phi}_i$ is the ith node's estimate of the parameter of interest and $\boldsymbol{\alpha}_i = [\alpha_{o^i_1,i}, \cdots, \alpha_{o^i_{|\mathcal{N}_i|},i}]^T$. The estimation error can be easily characterized as $E\{(\hat{\phi}_i - C)^2\} = \boldsymbol{\alpha}_i^T ({R_i}^2 \Omega_i + N_i)\boldsymbol{\alpha}_i$. To ensure an unbiased estimator, we should have $\boldsymbol{\alpha}_i^T R_i \vec{1} = 1$. We have

$$\begin{aligned} \boldsymbol{\alpha}_i^* = \arg\ \min\ & \boldsymbol{\alpha}_i^T ({R_i}^2 \Omega_i + N_i)\boldsymbol{\alpha}_i, \\ \text{subject to } & \boldsymbol{\alpha}_i^T R_i \vec{1} = 1. \end{aligned} \tag{4}$$

By noting that Eq. 4 is a convex function of $\boldsymbol{\alpha}_i$, we have

$$\alpha^*_{j,i} = \frac{1}{r_{j,i}} \frac{r^2_{j,i}(r^2_{j,i}\sigma^2_{\omega_j} + \sigma^2_{j,i})^{-1}}{\sum\limits_{l \in \mathcal{N}_i} r^2_{l,i}(r^2_{l,i}\sigma^2_{\omega_l} + \sigma^2_{l,i})^{-1}} \text{ for } j \in \mathcal{N}_i. \tag{5}$$

Let ϵ_i denote the estimation error of node i, we have:

$$\begin{aligned} \epsilon_i &= E\{(\hat{\phi}_i - C)^2\} \\ &= \frac{1}{\frac{1}{\sigma^2_{\omega_i}} + \sum\limits_{j \in \mathcal{N}_i, j \neq i} (\sigma^2_{\omega_j} + \frac{1}{\text{CNR}_{j,i}})^{-1}} < \sigma^2_{\omega_i}, \end{aligned} \tag{6}$$

where channel to noise ratio is defined as $\text{CNR}_{j,i} = \frac{r^2_{j,i}}{\sigma^2_{j,i}}$.

In this part we consider the scenario where the nodes sense the common value only one time but try to improve their estimate by communicating with their neighbors a number of times. We consider two strategies for this case: Fusion and Diversity. In the fusion strategy, each node fuses its received information, at every time step, and sends its fused value to its neighbors. We will show that, if graph connectivity is low, this approach can help propagate the information through the network. Another possible approach is for each node to use its transmission and repeat its original sensed value, without any fusion. This approach can provide a time diversity and can improve the overall reception quality [13]. If the graph connectivity is high enough but the link qualities are the main bottleneck, this approach can perform better than the fusion technique. It is the goal of this section to analyze both approaches and highlight the underlying tradeoffs.

A. SSMC - Fusion approach

Motivated by consensus strategies [14], in this part we consider the scenario where each node fuses the received information in every iteration and sends the fused value to its neighbors. This approach increases the flow of information in the network and is in particular suitable for cases where the graph connectivity is the main bottleneck. After the first iteration, each node's estimation, based on the best linear approach of Eq. 2 and Eq. 3, can be expressed by:

$$\begin{aligned} \hat{\phi}_i(1) &= \sum_{j \in \mathcal{N}_i} \alpha_{j,i}(0) y_{j,i}(0) \\ &= C + \sum_{j \in \mathcal{N}_i} \alpha_{j,i}(0)\big(r_{j,i}(0)\omega_j(0) + n_{j,i}(0)\big). \end{aligned} \tag{7}$$

In the next time step, the problem would have the same form if we define an equivalent sensing noise. More specifically, we define the equivalent noise as follows based on Eq. 7: $\widetilde{\omega}_i(1) = \sum_{j \in \mathcal{N}_i} \alpha_{j,i}(0)\big(r_{j,i}(0)\omega_j(0) + n_{j,i}(0)\big)$. As it can been seen, the equivalent sensing noises are correlated, which makes the mathematical derivations more challenging. Therefore we need to change the architecture of the best linear unbiased estimation for the next iterations. For the ordered $\mathcal{O}^i$, we define $\hat{\boldsymbol{\phi}}_i(k) = [\hat{\phi}_{o^i_1}(k), \cdots, \hat{\phi}_{o^i_{|\mathcal{N}_i|}}(k)]^T$ where $\hat{\phi}_j(k)$ denotes node j's estimate after fusion at time k and $\widetilde{\boldsymbol{\omega}}_i(k) = [\widetilde{\omega}_{o^i_1}(k), \cdots, \widetilde{\omega}_{o^i_{|\mathcal{N}_i|}}(k)]^T$. We have:

$$\hat{\boldsymbol{\phi}}_i(k) = C\vec{1} + \widetilde{\boldsymbol{\omega}}_i(k). \tag{8}$$

Also we define $\boldsymbol{\alpha}_i(k) = [\alpha_{o^i_1,i}(k), \cdots, \alpha_{o^i_{|\mathcal{N}_i|},i}(k)]^T$ and $\boldsymbol{\beta}_i(k) = [\beta_{o^i_1,i}(k), \cdots, \beta_{o^i_{|\mathcal{N}_i|},i}(k)]^T$ where $\beta_{j,i}(k) = r_{j,i}(k)\alpha_{j,i}(k)$. The next estimate of node i can be expressed by:

$$\begin{aligned} \hat{\phi}_i(k+1) &= \boldsymbol{\alpha}_i^T(k)\big(R_i(k)\hat{\boldsymbol{\phi}}_i(k) + \mathbf{n}_i(k)\big) \\ &= \boldsymbol{\beta}_i^T(k)\hat{\boldsymbol{\phi}}_i(k) + \boldsymbol{\alpha}_i^T(k)\mathbf{n}_i(k) \\ &= \boldsymbol{\beta}_i^T(k)\vec{1}C + \boldsymbol{\beta}_i^T(k)\widetilde{\boldsymbol{\omega}}_i(k) + \boldsymbol{\alpha}_i^T(k)\mathbf{n}_i(k). \end{aligned} \tag{9}$$

We then have the following optimization problem:

$$\boldsymbol{\beta}_i^*(k) = \arg\min E\left\{\left(\boldsymbol{\beta}_i^T(k)\widetilde{\boldsymbol{\omega}}_i(k) + \boldsymbol{\alpha}_i^T(k)\mathbf{n}_i(k)\right)^2\right\}, \quad \text{subject to } \vec{1}^T\boldsymbol{\beta}_i(k) = 1. \tag{10}$$

We define $\boldsymbol{\gamma}_i(k) = [\frac{n_{o_1^i,i}(k)}{r_{o_1^i,i}(k)}, \cdots, \frac{n_{o^i_{|\mathcal{N}_i|},i}(k)}{r_{o^i_{|\mathcal{N}_i|},i}(k)}]^T$ and $\Gamma_i(k) = E\{\boldsymbol{\gamma}_i(k)\boldsymbol{\gamma}_i^T(k)\}$, which is a diagonal matrix (averaging is over noise). We have $\boldsymbol{\alpha}_i^T(k)\mathbf{n}_i(k) = \boldsymbol{\beta}_i^T(k)\boldsymbol{\gamma}_i(k)$, resulting in,

$$\boldsymbol{\beta}_i^*(k) = \arg\min \boldsymbol{\beta}_i^T(k)\left(S_i(k) + \Gamma_i(k)\right)\boldsymbol{\beta}_i(k), \quad \text{subject to } \vec{1}^T\boldsymbol{\beta}_i(k) = 1, \tag{11}$$

where $S_i(k) = E\{\widetilde{\boldsymbol{\omega}}_i(k)\widetilde{\boldsymbol{\omega}}_i^T(k)\}$. The solution can then be easily characterized as $\boldsymbol{\beta}_i^*(k) = \frac{Q_i^{-1}(k)\vec{1}}{\vec{1}^T Q_i^{-1}(k)\vec{1}}$, where $Q_i(k) = S_i(k) + \Gamma_i(k)$. Next we define $\mathcal{B}(k) \triangleq [\mathcal{B}_{m,n}(k)]$, where $\mathcal{B}_{m,n}(k) = \begin{cases} \beta_{n,m}(k), & n \in \mathcal{O}^m \\ 0, & n \notin \mathcal{O}^m \end{cases}$. The unbiased estimation guarantees that $\mathcal{B}(k)$ has one as an eigenvalue with vector $\vec{1}$ as its corresponding eigenvector. Let $\hat{\boldsymbol{\phi}}(k) = [\hat{\phi}_1(k), \cdots, \hat{\phi}_M(k)]^T$ and $\widetilde{\boldsymbol{\omega}}(k) = [\widetilde{\omega}_1(k), \cdots, \widetilde{\omega}_M(k)]^T$. We have:

$$\widetilde{\boldsymbol{\omega}}(k+1) = \mathcal{B}(k)\widetilde{\boldsymbol{\omega}}(k) + \begin{bmatrix} \vdots \\ \boldsymbol{\alpha}_i^T(k)\mathbf{n}_i(k) \\ \vdots \end{bmatrix}. \tag{12}$$

Therefore the overall error covariance matrix, $S(k) = E\{\widetilde{\boldsymbol{\omega}}(k)\widetilde{\boldsymbol{\omega}}^T(k)\}$, can be expressed by:

$$S(k+1) = \mathcal{B}(k)S(k)\mathcal{B}^T(k) + \text{diag}\left(\cdots, \boldsymbol{\beta}_i^T(k)\Gamma_i(k)\boldsymbol{\beta}_i(k), \cdots\right). \tag{13}$$

As can be seen from the solution of Eq. 11, in order for the ith node to find $\boldsymbol{\beta}_i^*(k)$, it needs to calculate $S_i(k)$. Calculating $S_i(k)$, however, requires knowing current estimation error covariance (after fusion) of all the neighbors of node i as well as the cross correlation between any two estimation errors in its neighborhood. Since calculating the cross correlations would require node i to use information that may not be available locally (for instance use update Eq. 13), we also propose a sub-optimal version of the optimum solution as follows. Let $\epsilon_i(k) = E\{(\hat{\phi}_i(k) - C)^2\}$. From Eq. 13, each node can update its local error covariance as follows: $\epsilon_i(k+1) = \boldsymbol{\beta}_i^T(k)\big(S_i(k)+\Gamma_i(k)\big)\boldsymbol{\beta}_i(k) = \frac{1}{\vec{1}^T\left(S_i(k)+\Gamma_i(k)\right)^{-1}\vec{1}}$ and sends it to other nodes in its neighbor set. Then, we can approximate $S_i(k+1)$, using only the information available locally: $S_i(k+1) \approx \text{diag}(\epsilon_{o_1^i}(k+1), \cdots, \epsilon_{o^i_{|\mathcal{N}_i|}}(k+1))$. While this is a sub-optimal strategy, our simulation results show that the performance loss is considerably small, as we will see later in this section. Alternatively, one can consider strategies where each node learns its cross correlations with its neighbors.

Lemma 1: Let $S(k+1)$ be the error covariance matrix in iteration $k+1$. We have $\text{trace}(S(k+1)) = \sum_{i=1}^M \frac{1}{\vec{1}^T(S_i(k)+\Gamma_i(k))^{-1}\vec{1}}$, where $S_i(k) = E\{\widetilde{\boldsymbol{\omega}}_i(k)\widetilde{\boldsymbol{\omega}}_i^T(k)\}$.

Proof:

$$\begin{aligned}&\text{trace}(S(k+1)) = \\ &= \sum_{i=1}^M \left[\boldsymbol{\beta}_i^T(k)S_i(k)\boldsymbol{\beta}_i(k) + \boldsymbol{\beta}_i^T(k)\Gamma_i(k)\boldsymbol{\beta}_i(k)\right] \\ &= \sum_{i=1}^M \boldsymbol{\beta}_i^T(k)Q_i(k)\boldsymbol{\beta}_i(k) = \sum_{i=1}^M \frac{1}{\vec{1}^T Q_i^{-1}(k)\vec{1}} \\ &= \sum_{i=1}^M \frac{1}{\vec{1}^T(S_i(k)+\Gamma_i(k))^{-1}\vec{1}}.\end{aligned}$$

■

Theorem 1: For a fully-connected graph with AWGN links and same sensing and communication qualities, the trace of the error covariance matrix is a decreasing sequence with time.

Proof: Consider a fully-connected graph with $r_{j,i} = 1$, $\sigma_{j,i}^2 = \sigma^2$ and $\sigma_{\omega_i}^2 = \sigma_\omega^2$ for $1 \le i,j \le M$ and $j \ne i$. Then, we have the following decomposition of $S_i(k)$: $S_i(k) = a(k)\vec{1}\vec{1}^T + b(k)I$. Furthermore, $\Gamma_i(k) = \sigma^2 I^i$, where I is the identity matrix and I^i is the identity matrix with zero in (i,i) entry. Using the Matrix Inversion Lemma for a general $B = A + XRY$, we have: $B^{-1} = A^{-1} - A^{-1}X(R^{-1} + YA^{-1}X)^{-1}YA^{-1}$. Let $A = b(k)I + \sigma^2 I^i, X = \vec{1}, R = a(k)$ and $Y = \vec{1}^T$. We have:

$$\begin{aligned}&(S_i(k)+\Gamma_i(k))^{-1} = \\ &A^{-1} - A^{-1}\vec{1}(a^{-1}(k) + \vec{1}^T A^{-1}\vec{1})^{-1}\vec{1}^T A^{-1},\end{aligned}$$

and $\vec{1}^T(S_i(k)+\Gamma_i(k))^{-1}\vec{1} = \vec{1}^T A^{-1}\vec{1} - \frac{(\vec{1}^T A^{-1}\vec{1})^2}{a^{-1}(k)+\vec{1}^T A^{-1}\vec{1}}$, where A is a diagonal matrix and $\vec{1}^T A^{-1}\vec{1} = \frac{M-1}{b(k)+\sigma^2} + \frac{1}{b(k)}$. After some derivations, we have

$$\begin{aligned}\frac{1}{\vec{1}^T(S_i(k)+\Gamma_i(k))^{-1}\vec{1}} &= a(k) + \frac{1}{\frac{M-1}{b(k)+\sigma^2} + \frac{1}{b(k)}} \\ &< a(k) + b(k),\end{aligned}$$

and as a result

$$\text{trace}(S(k+1)) = \sum_{i=1}^M \frac{1}{\vec{1}^T(S_i(k)+\Gamma_i(k))^{-1}\vec{1}} < \text{trace}(S(k)).$$

■

Theorem 2: For a fully-connected graph with AWGN links and same sensing and communication qualities, the average steady state estimation error covariance of any node can be approximated as follows for large enough M:

$$\epsilon(\infty) \approx \frac{\sigma_\omega^2}{M} + \frac{\sigma^2}{M^2}\sum_{k=0}^{\infty} \frac{1}{\left(k+1+\frac{\sigma^2}{M\sigma_\omega^2}\right)^2}, \tag{14}$$

where σ^2 and σ_ω^2 are as defined in Theorem 1.

Proof: For a fully-connected graph with AWGN links and same sensing and communication qualities, we can have the following decomposition: $\mathcal{B}(k) = p(k)\vec{1}\vec{1}^T + q(k)I$, where $p(k)$ and $q(k)$ are scalars. Therefore,

$$\begin{aligned}S(k+1) &= a(k+1)\vec{1}\vec{1}^T + b(k+1)I \\ &= \mathcal{B}(k)S(k)\mathcal{B}(k) + (M-1)p^2(k)\sigma^2 I.\end{aligned} \tag{15}$$

By substituting the decomposition of $S(k)$ and $\mathcal{B}(k)$, we have:

$$\begin{aligned} &a(k+1) \\ &= \big(Mp(k)+q(k)\big)a(k) + \big(p^2(k)M + 2p(k)q(k)\big)b(k) \\ &= a(k) + \frac{1-q^2(k)}{M}b(k), \end{aligned} \tag{16}$$

where $\mathcal{B}(k)\vec{1} = \vec{1}$ results in $Mp(k) + q(k) = 1$. Furthermore,

$$\begin{aligned} b(k+1) &= (M-1)p^2(k)\sigma^2 + q^2(k)b(k) \qquad (17) \\ &= (M-1)\big(\frac{1-q(k)}{M}\big)^2\sigma^2 + q^2(k)b(k). \end{aligned}$$

For a fully-connected graph, characterizing $\boldsymbol{\beta}_1(k)$ suffices for characterizing $\mathcal{B}(k)$. By utilizing Matrix Inversion Lemma, $\boldsymbol{\beta}_1(k)$ and $q(k)$ can be characterized as:

$$\boldsymbol{\beta}_1(k) = \frac{1}{\frac{1}{b(k)} + \frac{M-1}{b(k)+\sigma^2}} \begin{bmatrix} \frac{1}{b(k)} \\ \frac{1}{b(k)+\sigma^2} \\ \vdots \\ \frac{1}{b(k)+\sigma^2} \end{bmatrix}, \tag{18}$$

and

$$q(k) = \frac{1}{\frac{1}{b(k)} + \frac{M-1}{b(k)+\sigma^2}} \big[\frac{1}{b(k)} - \frac{1}{b(k)+\sigma^2}\big] = \frac{\sigma^2}{\sigma^2 + Mb(k)}. \tag{19}$$

Substituting Eq. 19 in Eq. 17 results in:

$$b(k+1) = \frac{(M-1)\sigma^2 b^2(k) + \sigma^4 b(k)}{\big(\sigma^2 + Mb(k)\big)^2}. \tag{20}$$

For large M, we can approximate $M-1$ with M, resulting in $b(k+1) \approx \frac{\sigma^2 b(k)}{\sigma^2 + Mb(k)}$. Considering the initial condition $b(0) = \sigma_\omega^2$, we have: $b(k) \approx \frac{\frac{\sigma^2}{M}}{k + \frac{\sigma^2}{M\sigma_\omega^2}}$. By noting that $a(0) = 0$, Eq. 16 will result in,

$$\begin{aligned} a(\infty) \approx &\frac{1}{M}\sum_{k=0}^{\infty}\Big[b(k) - \frac{b(k)}{\big(1+\frac{M}{\sigma^2}b(k)\big)^2}\Big] \qquad (21) \\ &= \frac{\sigma^2}{M^2}\sum_{k=0}^{\infty}\Big[\frac{1}{k+\frac{\sigma^2}{M\sigma_\omega^2}} - \frac{k+\frac{\sigma^2}{M\sigma_\omega^2}}{\big(k+1+\frac{\sigma^2}{M\sigma_\omega^2}\big)^2}\Big] \\ &= \frac{\sigma^2}{M^2}\sum_{k=0}^{\infty}\Big[\frac{1}{k+\frac{\sigma^2}{M\sigma_\omega^2}} - \frac{1}{k+1+\frac{\sigma^2}{M\sigma_\omega^2}}\Big] \\ &+ \frac{\sigma^2}{M^2}\sum_{k=0}^{\infty}\frac{1}{\big(k+1+\frac{\sigma^2}{M\sigma_\omega^2}\big)^2} \\ &= \frac{\sigma_\omega^2}{M} + \frac{\sigma^2}{M^2}\sum_{k=0}^{\infty}\frac{1}{\big(k+1+\frac{\sigma^2}{M\sigma_\omega^2}\big)^2}. \end{aligned}$$

As k goes to infinity, we have $b(\infty) = 0$ and $S(\infty) = a(\infty)\vec{1}\vec{1}^T$, which results in $\epsilon(\infty) = a(\infty)$. ■

B. SSMC - Diversity approach

As mentioned earlier, diversity techniques can be used to improve system's performance if link qualities are the main bottleneck. In this case, each node senses the common value one time. It then transmits the sensed value without fusion to all its neighbors a number of times. Upon receiving the estimates of other nodes, each node takes an average of the received information to overcome communication errors. It then uses the BLUE fusion approach to provide a better estimation of the common value. We have:

$$\begin{aligned} &x_j(k+1) = x_j(k), \\ &y_{j,i}(k) = r_{j,i}(k)x_j(k) + n_{j,i}(k). \end{aligned} \tag{22}$$

Let $\hat{\phi}_{j,i}(k)$ denote the best linear estimation of node i of the noisy sensing of node $j \in \mathcal{N}_i \backslash \{i\}$ at time k. It can be easily confirmed that for $k \geqslant 0$ we have

$$\hat{\phi}_{j,i}(k) = \frac{\epsilon_{j,i}(k)}{\epsilon_{j,i}(k-1)}\hat{\phi}_{j,i}(k-1) + \epsilon_{j,i}(k)\text{CNR}_{j,i}(k)y_{j,i}(k), \tag{23}$$

where $\epsilon_{j,i}^{-1}(k) = \epsilon_{j,i}^{-1}(k-1) + \text{CNR}_{j,i}(k)$, and $\epsilon_{j,i}(k) = E\{(\hat{\phi}_{j,i}(k) - x_j(k))^2\}$. For $k = -1$, we take $\epsilon_{j,i}(-1) = \infty$. The ith node then fuses $\hat{\phi}_{j,i}$ for $j \in \mathcal{N}_i$. We have the following for the estimation of the ith node of C by using a BLUE estimator:

$$\hat{\phi}_i(k+1) = \sum_{j\in\mathcal{N}_i}\frac{(\epsilon_{j,i}(k)+\sigma_{\omega_j}^2)^{-1}}{\sum_{j\in\mathcal{N}_i}(\epsilon_{j,i}(k)+\sigma_{\omega_j}^2)^{-1}}\hat{\phi}_{j,i}(k), \tag{24}$$

where $\epsilon_{i,i}(k) = 0$ and $\hat{\phi}_{i,i}(k) = x_i(k)$. For $k \geqslant 0$, we have

$$\begin{aligned} \epsilon_i(k) &= E\{\big(\hat{\phi}_i(k) - C\big)^2\} \qquad (25) \\ &= \frac{1}{\frac{1}{\sigma_{\omega_i}^2} + \sum_{j\in\mathcal{N}_i, j\neq i}\big(\sigma_{\omega_j}^2 + \epsilon_{j,i}(k-1)\big)^{-1}}. \end{aligned}$$

As k goes to infinity, the effect of channel errors is reduced. The steady state behavior will then be similar to one time communication over ideal channels:

$$\lim_{k\to\infty}\epsilon_i(k) = \big(\sum_{j\in\mathcal{N}_i}\frac{1}{\sigma_{\omega_j}^2}\big)^{-1}. \tag{26}$$

For a fully-connected network with AWGN channels and same sensing and communication qualities, we have $\sigma_{\omega_j} = \sigma_\omega$, $r_{j,i}(k) = 1$ and $\sigma_{j,i} = \sigma$. Therefore, $\epsilon_{j,i}(k) = \frac{\sigma^2}{k+1}$ which results in $\epsilon_i(k) = \frac{\sigma_\omega^2}{M}\frac{k+\frac{\sigma^2}{\sigma_\omega^2}}{k+\frac{\sigma^2}{M\sigma_\omega^2}}$ for $\forall i$.

C. Comparison of fusion with diversity approaches

As indicated in the previous section, for a fully-connected graph, diversity approach can provide a better performance. On the other hand, fusion approach can outperform diversity if link qualities are the main bottleneck. Fig. 1 and Fig. 2 compare the performance of SSMC approach for both fusion and diversity cases over a star-shaped and fully-connected (FC) graph respectively. The figures show average of the estimation error covariance of all the nodes (averaged over fading as well). In all the simulations, we assume that $C = 10$ and $\sigma_\omega^2 = 5$. Link qualities are worse for Fig. 2 as compared to Fig. 1. As can be seen, fusion strategy outperforms the diversity one in Fig. 1 since the main bottleneck is connectivity. As the link qualities get worse in Fig. 2, however, it can be seen that the diversity approach provides a better overall performance by reducing link uncertainties. In practice, a combination of both approaches may be suitable to account for different scenarios. It can also be seen that the performance of the proposed sub-optimal fusion strategy, based on the diagonal approximation of $S_i(k)$, is considerably close to the optimum one.

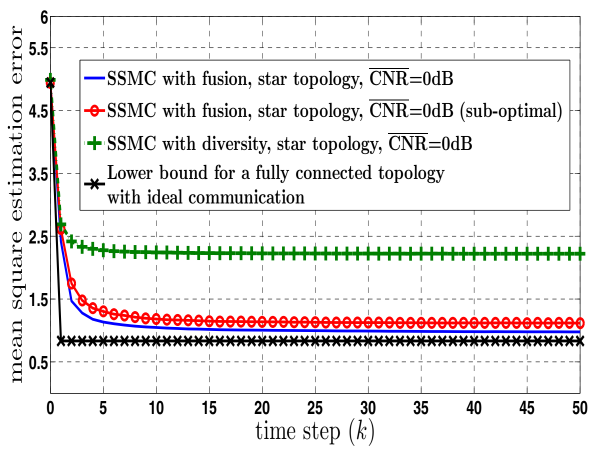

Fig. 1. A comparison of fusion and diversity techniques for the SSMC case with M=6, the graph is star-shaped with fading channels for non-ideal communication cases.

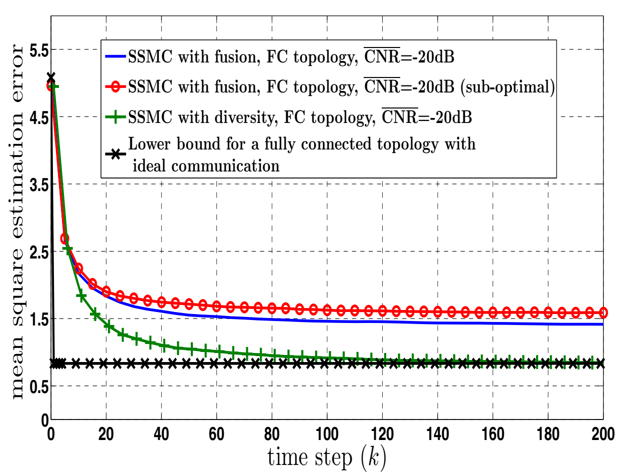

Fig. 2. A comparison of fusion and diversity techniques for the SSMC case with M=6, the graph is fully connected with fading channels for non-ideal communication cases.

IV. Scenario 2: Multiple Sensing and Multiple Communication (MSMC)

A. MSMC - no fusion

In this part, we consider the scenario that the nodes have the capability of sensing the parameter of interest multiple times. Similar to the previous section we consider two cases of fusion and no fusion. In each iteration, every node measures the common value and sends its measurement to its neighbors over the fading channels. In order to take into account new sensing, we modify the definition of $\mathbf{x}_i(k)$ and $\mathbf{y}_i(k)$ as follows:

$$\mathbf{y}'_i(k+1) = R_i(k)\mathbf{x}'_i(k) + \mathbf{n}_i(k), \tag{27}$$

where $[\mathbf{x}'_i(k)]_m = \begin{cases} [\mathbf{x}_i(k)]_m, & o^i_m \neq i \\ x_i(k+1), & o^i_m = i \end{cases}$ and $[.]_m$ denotes the mth element of the argument vector. In the time step $k+1$, node i estimates the common value based on its estimation at time k, new sensing at time $k+1$ and received information from its neighbors at time k. Therefore, $\hat{\phi}_i(k+1) = f\big(\hat{\phi}_i(k), \mathbf{y}'_i(k+1)\big)$ where $f(.)$ is a linear function that minimizes the following cost:

$$E\Big\{\Big(f\big(\hat{\phi}_i(k), \mathbf{y}'_i(k+1)\big) - C\Big)^2\Big\}. \tag{28}$$

By using a best linear estimation, we have:

$$\hat{\phi}_i(k+1) = \hat{\phi}_i(k) + \epsilon_i(k+1)\vec{1}^T R_i(k)\big(R_i^2(k)\Omega_i + N_i\big)^{-1} \times \big(\mathbf{y}'_i(k+1) - R_i(k)\vec{1}\hat{\phi}_i(k)\big) \quad \text{and}$$

$$\epsilon_i^{-1}(k+1) = \epsilon_i^{-1}(k) + \vec{1}^T R_i(k)\big(R_i^2(k)\Omega_i + N_i\big)^{-1} R_i(k)\vec{1}.$$

where $\hat{\phi}_i(0) = x_i(0)$ and $\epsilon_i(0) = E\{\omega_i^2\}$. After some straightforward simplifications, we can characterize the best linear estimator as:

$$\hat{\phi}_i(k+1) = \frac{\epsilon_i(k+1)}{\epsilon_i(k)}\hat{\phi}_i(k) + \epsilon_i(k+1)\vec{1}^T R_i(k)\big(R_i^2(k)\Omega_i + N_i\big)^{-1}\mathbf{y}'_i(k+1), \tag{29}$$

where

$$\epsilon_i^{-1}(k+1) = \epsilon_i^{-1}(k) + \vec{1}^T R_i^2(k)\big(R_i^2(k)\Omega_i + N_i\big)^{-1}\vec{1}. \tag{30}$$

For $k \geq 0$ we have:

$$\epsilon_i(k+1) = \frac{1}{\frac{1}{\epsilon_i(k)} + \frac{1}{\sigma^2_{\omega_i}} + \sum_{j\in\mathcal{N}_i, j\neq i}\big(\sigma^2_{\omega_j} + \text{CNR}^{-1}_{j,i}(k)\big)^{-1}}. \tag{31}$$

For a fully-connected graph with AWGN links and the same sensing and communication qualities $\big(\sigma_{\omega_i} = \sigma_\omega, r_{j,i}(k) = 1$ and $\sigma_{j,i} = \sigma\big)$, Eq. 31 results in:

$$\epsilon_i(k) = \frac{1}{(\frac{1}{\sigma_\omega^2} + \frac{M-1}{\sigma_\omega^2+\sigma^2})k + \frac{1}{\sigma_\omega^2}}, \ \forall i. \tag{32}$$

B. MSMC - with fusion

As discussed in the previous section, if the graph was fully connected, each node in each time had all the information of other nodes (albeit noisy). For a not fully connected graph, however, information of some of the nodes will never reach some other nodes directly. In order to increase the flow of information in the network, in this part we consider the case where each node fuses its current sensing with its received information and sends it to its neighboring nodes.

As an example, consider a star-shaped network. In each iteration, the leaf nodes send their current sensed values to the center node. At time k, the center node has all the information received from the end nodes at time $k-1$, its sensed value at time k as well as its previous estimation, based on which it can update its estimate. Without loss of generality, assume that node 1 is the center node. Let $f_{1,i}(k)$ represent the fused value that the center sends to node $i \neq 1$ at time k. We have:

$$f_{1,i}(k) = \sum_{j=2, j\neq i}^{M} \alpha_{j,1}(k) y_{j,1}(k-1) + \alpha_{1,1}(k) x_1(k), \quad (33)$$

where $x_1(k)$ and $y_{j,1}(k)$ denote the sensed value of the center node and its reception from node j at time k, respectively. $\alpha_{j,1}(k)$ can be easily calculated based on the BLUE approach. As can be seen from Eq. 33, we excluded the information received from node j from the update that will be sent to it, in order to facilitate mathematical derivations. In general, however, calculating the optimum coefficients for more general graphs requires considering correlation of different information flow paths and is a subject of further extensions.

Fig. 3 shows the impact of link qualities and graph topology on the performance of MSMC for the case of no fusion. Fig. 4 shows the impact of fusion on MSMC performance over a star-shaped network. It can be seen that fusion can improve the performance since graph connectivity is low. Its impact is more drastic when link qualities are better, as expected.

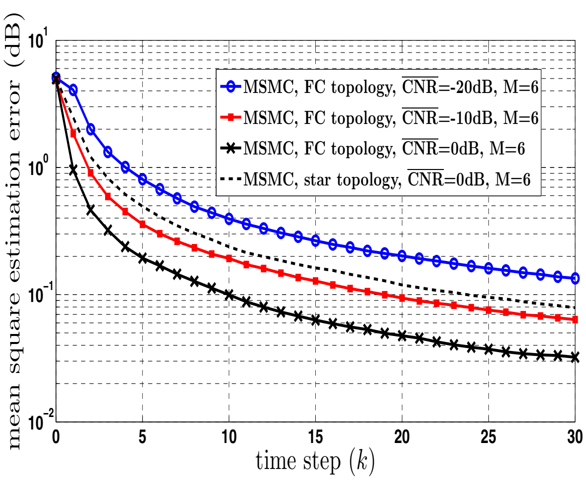

Fig. 3. Performance of MSMC for the case of no fusion over both a fully connected and a star-shaped graph for different link qualities.

V. Conclusions

In this paper we considered a distributed and cooperative estimation problem over fading channels. We considered different ways of using the available bandwidth, in terms of what each node can send to its neighbors. More specifically, we mathematically characterized the impact of both fusion and diversity strategies on the overall performance and highlighted the underlying tradeoffs. We showed that the fusion strategy is more suitable when the graph connectivity is low while diversity technique is a better candidate if poor link quality is the main bottleneck. We furthermore explored the impact of multiple sensing on the overall performance.

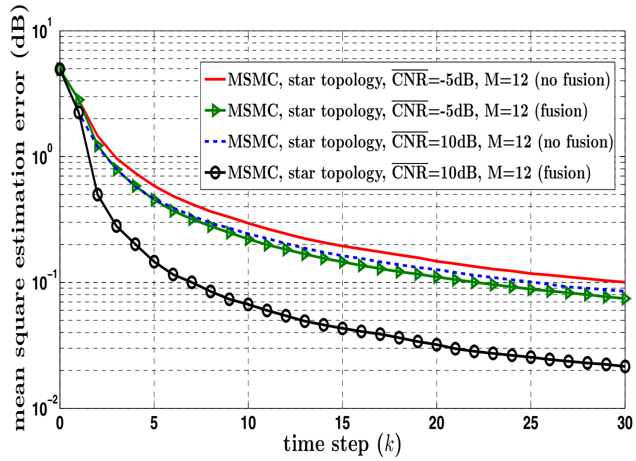

Fig. 4. Impact of fusion on the performance of MSMC over a star-shaped graph.

References

[1] Y. Mostofi, “Binary Consensus with Gaussian Communication Noise: A Probabilistic Approach,” Proceedings of the 46th IEEE Conference on Decision and Control (CDC), Dec. 2007.

[2] D. P. Spanos, R. Olfati-Saber and R. M. Murray, “Approximate distributed Kalman filtering in sensor networks with quantifiable performance,” in the proceedings of 4th international conference on Information Processing in Sensor Networks (IPSN), pp. 133-139, 2005.

[3] W. Ren, R. Beard and E. Atkins, “A Survey of Consensus Problems in Multi-agent Coordination,” Proceedings of IEEE ACC, 2005.

[4] Lin Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” Proceedings of the 42nd IEEE Conference on Decision and Control, vol. 5, pp. 4997-5002, 9-12 Dec. 2003.

[5] L. Xiao, S. Boyd and S. Kim, “ Distributed average consensus with least-mean-square deviation,” J. Parallel Distrib. Comput. vol. 67, no. 1, pp. 33-46, Jan. 2007.

[6] L. Xiao, S. Boyd and S. Lall, “A scheme for robust distributed sensor fusion based on average consensus,” Fourth International Symposium on Information Processing in Sensor Networks, pp. 63-70, April 2005.

[7] I. D. Schizas, A. Ribeiro and G. B. Gianakis, “Consensus in Ad Hoc WSNs with Noisy Links - Part I: Distributed Estimation of Deterministic Signals,” IEEE Transactions on Signal Processing, vol. 56, no. 1, pp. 350-364, January 2008.

[8] W. Ren, R. Beard and D. Kingston, “Multi-agent Kalman consensus with relative uncertainty,” Proceedings of IEEE American Control Conference, 2005.

[9] M. E. Yildiz and A. Scaglione, “The Limiting Rate Behavior and Rate Allocation Strategies for Average Consensus Problems with Bounded Convergence,” Proceedings of IEEE ICASSP, Las Vegas, NV, April 2008.

[10] Y. Ruan and Y. Mostofi, “Binary Consensus with Soft Information Processing in Cooperative Networks,” Proceedings of the 47th IEEE Conference on Decision and Control (CDC), Dec. 2008.

[11] M. Malmirchegini, Y. Ruan and Y. Mostofi, “Binary Consensus Over Fading Channels: A Best Affine Estimation Approach,” IEEE Globecom, 2008.

[12] M. Kaneko, K. Hayashi, P. Popovski, K. Ikeda, H. Sakai and R. Prasad, “Amplify-and-forward cooperative diversity schemes for multi-carrier systems,” IEEE Transactions on Wireless Communications, vol.7, no.5, pp. 1845-1850, May 2008.

[13] Y. Mostofi and Y. Ruan, “Binary Consensus over AWGN Channels,” under revision, IEEE Transactions on Automatic Control, Dec. 2008.

[14] R. Olfati-Saber and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” IEEE Transactions on Automatic Control, vol. 49, no. 9, pp. 1520-1533, 2004.