PRIVACY-PRESERVING DISTRIBUTED PRECODER DESIGN FOR DECENTRALIZED ESTIMATION

Naveen K. D. Venkategowda and Stefan Werner

Department of Electronic Systems, Norwegian University of Science and Technology, Norway
Email: {naveen.dv, stefan.werner}@ntnu.no

ABSTRACT

We study privacy-preserving precoder design for decentralized estimation in wireless sensor networks where the sensor nodes want their local information such as the channel state information, observation matrices, and observation covariance matrices to be private. We propose a distributed algorithm with closed form expressions to design the precoders and fusion rule that minimize the estimation error by exchanging messages which do not reveal the local information. We derive the privacy limits offered by the proposed algorithm and prove that the algorithm is privacy-preserving. Simulation results illustrate the trade-off between privacy and estimation accuracy of the proposed algorithm.

1. INTRODUCTION

Wireless sensor networks (WSN) are widely employed for event detection, tracking, and estimation with applications in security, environmental monitoring, and smart infrastructure. In the decentralized estimation, sensor nodes transmit their observations of the source to the fusion center (FC), which estimates the parameter of interest. The observation noise and the noisy fading wireless channel between the sensors and FC degrades the estimation accuracy. However, transceivers can be designed to minimize these ill effects by utilizing the knowledge of channel state information (CSI) and sensor observation model information. Precoding enables us to exploit the multiple access channel (MAC) to coherently combine the transmissions from the sensor nodes over the channel, and thus leading to diversity and array gain, which enhance the estimation accuracy [1–5].

However, the local sensor information and the sensor measurements are sensitive and must be protected from leaking to unauthorized agents. Conventional cryptographic security solutions are prohibitively demanding in resources to employ them in WSNs. Hence, low complex physical layer security and privacy for WSNs has garnered significant attention recently. In [6] and [7], the authors studied secure distributed detection in presence of an malicious eavesdropper. Linear precoding was proposed in [8] to protect private hypothesis being inferred from a curious FC. In [9] and [10], precoding techniques were investigated for secure remote estimation in presence of an eavesdropper. Optimal encoding of parameter was considered in [11] to minimize the estimation error at the FC while ensuring the accuracy at eavesdropper to be greater than a threshold.

In many applications the sensors might be unwilling to share their observation models and CSI due to privacy and security concerns. For instance, in radar sensor networks the observation matrices contain sensitive information such as codes, timing, and location [12–14], which cannot be revealed to other entities. In such scenarios, it is imperative that the network designs the precoders and the fusion rule with information privacy requirements i.e. without sharing the CSI, observations or observation models directly.

To that end, we propose an iterative distributed algorithm to compute the precoders and fusion rule while protecting the privacy of the sensor nodes. As the minimum mean square error (MMSE) estimation framework results in a non-convex and non-separable objective function, we derive an upper bound on the optimal MMSE error, which is convex and separable across the network. This bound is used to optimize the WSN. To ensure information privacy, we employ alternating direction method of multipliers (ADMM) and privacy-preserving average consensus to solve the dual of the MSE minimization problem in a distributed manner.

In this approach, at each iteration, the sensor nodes update their precoder and shares the local perturbed dual variable to their neighboring nodes. We derive closed form expressions to update the precoders as well as the dual variables. In the proposed method, FC can determine the fusion rule using only a scalar quantity fed back from the sensors without acquiring any global information. We prove that the proposed algorithm is privacy-preserving and derive limits on privacy guaranteed for the sensor nodes. Simulation results are presented to demonstrate the estimation performance and the trade-off between privacy and estimation error of the proposed method.

Operators $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^\dagger$ and tr$(\cdot)$ denote transpose, Hermitian, pseudo-inverse, and trace of a matrix, respectively. $\Re(\cdot)$ represents the real part of a complex quantity and $\|\cdot\|$ represents the L_2 norm. $\mathbf{I}_n$ denotes an identity matrix of size n and $\otimes$ stands for the Kronecker product. $\mathsf{vec}(\mathbf{A})$ forms a column vector from a matrix $\mathbf{A}$ by stacking its column vectors. $\mathsf{diag}(\mathbf{a})$ denotes a diagonal matrix with elements of vector $\mathbf{a}$ on the principal diagonal.

2. SYSTEM MODEL

We consider a WSN with a FC employing r antennas and L sensor nodes, each employing t antennas. The WSN is modeled as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{C})$ where the set of vertices $\mathcal{V} = \{1, \ldots, L\}$ corresponds to the sensor nodes and the set $\mathcal{C}$ represents the communication links between the pair of nodes. Node $i \in \mathcal{V}$ can communicate with the nodes in its neighborhood $\mathcal{N}_i$. The adjacency matrix $\mathbf{E}$ of the WSN is defined as $[\mathbf{E}]_{ij} = 1$ if $(i, j) \in \mathcal{C}$, 0 otherwise.

The sensor nodes measure the parameter $\boldsymbol{\theta} \in \mathbb{C}^p$ having statistics $\boldsymbol{\theta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\theta)$. The observations $\mathbf{x}_i \in \mathbb{C}^m$ at the ith node can be modeled as

$$\mathbf{x}_i = \mathbf{A}_i\boldsymbol{\theta} + \mathbf{n}_i,$$

where $\mathbf{A}_i \in \mathbb{C}^{m\times p}$ is the observation matrix and $\mathbf{n}_i \in \mathbb{C}^m$ is the Gaussian observation noise with zero mean and covariance matrix $\boldsymbol{\Sigma}_i$.

In the decentralized estimation setting, the sensors transmit linearly-precoded observations to the FC over a wireless MAC. Let $\mathbf{B}_i \in \mathbb{C}^{t\times m}$ denote the precoding matrix at node i. The received

This work was partly supported by the ERCIM Alain Bensoussan Fellowship Programme and the Research Council of Noway.

data $\mathbf{y}_{\mathsf{FC}} \in \mathbb{C}^{r\times 1}$ at the FC can be expressed as

$$\mathbf{y}_{\mathsf{FC}} = \sum_{i=1}^{L} \mathbf{H}_i\mathbf{B}_i\mathbf{x}_i + \mathbf{v}_{\mathsf{FC}} = \sum_{i=1}^{L} \mathbf{H}_i\mathbf{B}_i(\mathbf{A}_i\boldsymbol{\theta} + \mathbf{n}_i) + \mathbf{v}_{\mathsf{FC}}, \quad (1)$$

where $\mathbf{H}_i \in \mathbb{C}^{r\times t}$ is the channel matrix between the ith node and the FC, $\mathbf{x}_i$ represents the observation at node i, and $\mathbf{v}_{\mathsf{FC}} \in \mathbb{C}^{r\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathsf{FC}})$ is the receiver noise at the FC. The transmit power at the ith sensor is limited by

$$\mathbb{E}\left[\|\mathbf{B}_i\mathbf{x}_i\|^2\right] = \mathrm{tr}(\mathbf{B}_i(\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H + \boldsymbol{\Sigma}_i)\mathbf{B}_i^H) \le P_i, \quad (2)$$

where P_i is the power available for data transmission.

The FC estimates the parameter $\boldsymbol{\theta}$ using the fusion rule $\mathbf{W} \in \mathbb{C}^{p\times r}$ as $\hat{\boldsymbol{\theta}} = \mathbf{W}^H\mathbf{y}_{\mathsf{FC}}$. Our objective is to compute $\{\mathbf{B}_i\}_{i=1}^L$ and $\mathbf{W}$ to maximize the estimation accuracy while guaranteeing privacy of the sensor information.

3. PRIVACY-PRESERVING PRECODER COMPUTATION

Substituting for $\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L)$ from (1), we can write the MMSE error $\mathcal{E}(\{\mathbf{B}_i\}, \mathbf{W}) = \mathbb{E}[\|\boldsymbol{\theta} - \mathbf{W}^H\mathbf{y}_{\mathsf{FC}}(\{\mathbf{B}_i\}_{i=1}^L)\|^2]$ as

$$\begin{aligned}\mathcal{E}(\{\mathbf{B}_i\}, \mathbf{W}) = \mathrm{tr}\Big(&\mathbf{R}_\theta + \mathbf{W}^H\mathbf{R}_{\mathsf{FC}}\mathbf{W} - \mathbf{R}_\theta\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)^H\mathbf{W}\\ &- \mathbf{W}^H\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)\mathbf{R}_\theta + \sum_{i=1}^{L}\mathbf{W}^H\mathbf{H}_i\mathbf{B}_i\boldsymbol{\Sigma}_i\mathbf{B}_i^H\mathbf{H}_i^H\mathbf{W}\\ &+ \mathbf{W}^H\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i\Big)\mathbf{R}_\theta\Big(\sum_{j=1}^{L}\mathbf{H}_j\mathbf{B}_j\mathbf{A}_j\Big)^H\mathbf{W}\Big). \quad (3)\end{aligned}$$

It can be seen that the objective function $\mathcal{E}(\{\mathbf{B}_i\}, \mathbf{W})$ is not separable and non-convex in $\{\mathbf{B}_i\}_{i=1}^L$ and $\mathbf{W}$. Hence, computing the optimal precoders and fusion rule that minimize the MMSE error with information privacy is intractable.

To overcome this, we constrain the precoders such that the response $\sum_{i=1}^L \mathbf{H}_i\mathbf{B}_i\mathbf{A}_i = \alpha\mathbf{D}$ and the fusion rule $\mathbf{W}^H = (1/\alpha)\mathbf{D}^\dagger$, where $\alpha \in \mathbb{R}$ is the gain that needs to be optimized and $\mathbf{D} \in \mathbb{C}^{r\times p}$ is a predefined matrix known to the FC and sensor nodes.

We can now upper bound the error $\mathcal{E}(\{\mathbf{B}_i\}_{i=1}^L, \mathbf{W})$ by

$$\mathcal{E}_{\mathsf{U}}(\{\mathbf{B}_i\}, \alpha) = \sum_{i=1}^{L}\mathrm{tr}(\mathbf{D}^\dagger\mathbf{H}_i\mathbf{B}_i\boldsymbol{\Sigma}_i\mathbf{B}_i^H\mathbf{H}_i^H(\mathbf{D}^\dagger)^H) + \frac{\sigma_{\mathsf{FC}}^2}{\alpha^2}, \quad (4)$$

which is convex and separable, and $\sigma_{\mathsf{FC}}^2 = \mathrm{tr}((\mathbf{D}^\dagger)^H\mathbf{R}_{\mathsf{FC}}\mathbf{D}^\dagger)$. The optimization problem to calculate $\{\mathbf{B}_i\}_{i=1}^L$ and α that minimize $\mathcal{E}_{\mathsf{U}}(\{\mathbf{B}_i\}, \alpha)$ can be formulated as

$$\begin{aligned}\min_{\{\mathbf{B}_i\},\alpha}\ & \mathcal{E}_{\mathsf{U}}(\{\mathbf{B}_i\}, \alpha)\\ \text{s. t.}\ & \sum\nolimits_{i=1}^{L}\mathbf{H}_i\mathbf{B}_i\mathbf{A}_i = \alpha\mathbf{D},\\ & \mathrm{tr}(\mathbf{B}_i(\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H + \boldsymbol{\Sigma}_i)\mathbf{B}_i^H) \le P_i,\ i = 1, \ldots, L.\end{aligned} \quad (5)$$

3.1. Distributed Precoder Design

Defining $\mathbf{b}_i = \mathsf{vec}(\mathbf{B}_i)$ and exploiting the identity $\mathsf{vec}(\mathbf{AXB}) = \left(\mathbf{B}^T \otimes \mathbf{A}\right)\mathsf{vec}(\mathbf{X})$, the equality constraint in (5) is recast as $\sum_{i=1}^L \mathbf{G}_i\mathbf{b}_i = \alpha\mathbf{c}$, where $\mathbf{G}_i = \mathbf{A}_i^T \otimes \mathbf{H}_i$ and $\mathbf{c} = \mathsf{vec}(\mathbf{D})$. Define $\boldsymbol{\mu}$ and λ_i as the Lagrange multiplier for the complex equality constraint and the power constraint at the ith sensor node, respectively. Employing the relation $\mathrm{tr}\left(\mathbf{A}^H\mathbf{BCD}^H\right) = \mathsf{vec}(\mathbf{A})^H((\mathbf{D}^H)^T \otimes \mathbf{B})\mathsf{vec}(\mathbf{C})$ [15], the Lagrangian $\mathcal{L}(\{\mathbf{b}_i\}, \boldsymbol{\mu}, \{\lambda_i\})$ for (5) can be simplified as

$$\begin{aligned}\mathcal{L}(\{\mathbf{b}_i\}, \boldsymbol{\mu}, \{\lambda_i\}) = \frac{\sigma_{\mathsf{FC}}^2}{\alpha^2} + \sum_{i=1}^{L}\mathbf{b}_i^H\mathbf{R}_i\mathbf{b}_i + \Re(\boldsymbol{\mu}^H(\mathbf{G}_i\mathbf{b}_i - \alpha\mathbf{c}))\\ + \lambda_i(\mathbf{b}_i^H\mathbf{Q}_i\mathbf{b}_i - P_i), \quad (6)\end{aligned}$$

where $\mathbf{Q}_i \triangleq (\mathbf{A}_i\mathbf{R}_\theta\mathbf{A}_i^H + \boldsymbol{\Sigma}_i)^T \otimes \mathbf{I}_t$ and $\mathbf{R}_i \triangleq \boldsymbol{\Sigma}_i^T \otimes \mathbf{H}_i^H(\mathbf{D}^\dagger)^H\mathbf{D}^\dagger\mathbf{H}_i$.

From the Karush-Kuhn-Tucker (KKT) conditions $\nabla\mathcal{L}(\{\mathbf{b}_i^*\}, \boldsymbol{\mu}^*, \{\lambda_i^*\}) = \mathbf{0}$ [16], the optimal $\mathbf{b}_i^*$ is given by

$$\mathbf{b}_i^* = (\mathbf{R}_i + \lambda_i^*\mathbf{Q}_i)^{-1}\mathbf{G}_i^H\boldsymbol{\mu}^*. \quad (7)$$

Further, the optimal precoder $\mathbf{b}_i^*$ must satisfy the complementary slackness conditions $\lambda_i^*((\mathbf{b}_i^*)^H\mathbf{Q}_i\mathbf{b}_i^* - P_i) = 0$, $\forall i$. Therefore, substituting for $\mathbf{b}_i^*$ in slackness conditions, we compute the Lagrange multipliers λ_i^* by solving [17, App. A]

$$\left\|\mathbf{Q}_i^{\frac{1}{2}}(\mathbf{R}_i + \lambda_i\mathbf{Q}_i)^{-1}\mathbf{G}_i^H\boldsymbol{\mu}\right\|^2 = P_i. \quad (8)$$

Next, from the KKT conditions $\partial\mathcal{L}/\partial\alpha^* = 0$, we can derive that optimal gain α^* as

$$\alpha^* = \left(\sigma_{\mathsf{FC}}^2/\Re(\boldsymbol{\mu}^{*H}\mathbf{c})\right)^{1/3}. \quad (9)$$

From (7) and (9), it must be noted that the knowledge of $\boldsymbol{\mu}^*$ suffices to compute $\{\mathbf{b}_i^*\}_{i=1}^L$ at the sensor nodes. The optimal $\boldsymbol{\mu}^*$ can be found from the dual problem of (5) given by $\max_{\boldsymbol{\mu},\{\lambda_i\}\ge 0} g(\boldsymbol{\mu}, \{\lambda_i\})$, where $g(\boldsymbol{\mu}, \{\lambda_i\}) = \inf_{\{\mathbf{b}_i\}} \mathcal{L}(\{\mathbf{b}_i\}, \boldsymbol{\mu}, \{\lambda_i\})$ is the dual function. For the given optimal λ_i^*, substituting the solution $\mathbf{b}_i^*$ from (7) into (6) and determining $g(\boldsymbol{\mu}, \{\lambda_i\})$, the optimal dual can be obtained from

$$\min_{\boldsymbol{\mu}}\ \boldsymbol{\mu}^H\Big(\sum_{i=1}^{L}\mathbf{G}_i(\mathbf{R}_i + \lambda_i^*\mathbf{Q}_i)^{-1}\mathbf{G}_i^H\Big)\boldsymbol{\mu} - 2\alpha^*\Re(\boldsymbol{\mu}^H\mathbf{c}). \quad (10)$$

Now, the above problem can be reformulated as

$$\begin{aligned}\min_{\boldsymbol{\mu}_i,\boldsymbol{\mu}}\ & \sum_{i=1}^{L}\boldsymbol{\mu}_i^H\mathbf{G}_i(\mathbf{R}_i + \lambda_i^*\mathbf{Q}_i)^{-1}\mathbf{G}_i^H\boldsymbol{\mu}_i - \frac{2\alpha^*}{L}\Re(\boldsymbol{\mu}_i^H\mathbf{c})\\ \text{s. t.}\ & \boldsymbol{\mu}_i = \boldsymbol{\mu},\ \forall i,\end{aligned} \quad (11)$$

and it can be verified that both (10) and (11) have an identical solution. The constraint in (11) is the consensus constraint which forces the local information $\boldsymbol{\mu}_i$ at the ith node to be equal to the actual $\boldsymbol{\mu}$. As the objective function in (11) is separable, the ith sensor node can independently compute the optimal $\boldsymbol{\mu}_i$. For that purpose, we rely on the ADMM technique to solve (11) in a distributed manner [18].

The augmented Lagrangian for problem (11) with the quadratic penalty function for the constraint violations is formed as

$$\begin{aligned}\mathcal{L}_\rho(\boldsymbol{\mu}_i, \mathbf{y}_i, \boldsymbol{\mu}) = \sum_{i=1}^{L}\boldsymbol{\mu}_i^H\mathbf{G}_i(\mathbf{R}_i + \lambda_i\mathbf{Q}_i)^{-1}\mathbf{G}_i^H\boldsymbol{\mu}_i - \frac{2\alpha}{L}\Re(\boldsymbol{\mu}_i^H\mathbf{c})\\ + \Re(\mathbf{y}_i^H(\boldsymbol{\mu}_i - \boldsymbol{\mu})) + \frac{\rho}{2}\|\boldsymbol{\mu}_i - \boldsymbol{\mu}\|_2^2, \quad (12)\end{aligned}$$

where ρ is the penalty parameter. Using the ADMM technique [18], we obtain the following iterative steps at the ith node for (11)

$$\{\boldsymbol{\mu}_i^{(k+1)}\} = \arg\min_{\{\boldsymbol{\mu}_i\}} \mathcal{L}_\rho(\boldsymbol{\mu}_i, \mathbf{y}_i^{(k)}, \boldsymbol{\mu}^{(k)}), \quad (13)$$

$$\boldsymbol{\mu}^{(k+1)} = \arg\min_{\boldsymbol{\mu}} \mathcal{L}_\rho(\boldsymbol{\mu}_i^{(k+1)}, \mathbf{y}_i^{(k)}, \boldsymbol{\mu}), \quad (14)$$

$$\mathbf{y}_i^{(k+1)} = \mathbf{y}_i^{(k)} + \rho(\boldsymbol{\mu}_i^{(k+1)} - \boldsymbol{\mu}^{(k+1)}). \quad (15)$$

It is apparent that (13) and (14) are unconstrained convex quadratic minimization problems. Thus, by computing the gradient and equating it to zero [19, pg. 741], the optimal point at the ith node is obtained as

$$\boldsymbol{\mu}_i^{(k+1)} = \left(\boldsymbol{\Psi}_i^{(k)} + \frac{\rho}{2}\mathbf{I}_{pr}\right)^{-1}\left(\frac{\rho}{2}\boldsymbol{\mu}^{(k)} + \frac{\alpha^{(k)}}{L}\mathbf{c} - \frac{1}{2}\mathbf{y}_i^{(k)}\right), \quad (16)$$

where $\boldsymbol{\Psi}_i^{(k)} = \mathbf{G}_i(\mathbf{R}_i + \lambda_i^{(k)}\mathbf{Q}_i)^{-1}\mathbf{G}_i^H$ and similarly the solution to (14) is given by

$$\boldsymbol{\mu}^{(k+1)} = \frac{1}{L}\sum_{i=1}^{L}\left(\boldsymbol{\mu}_i^{(k+1)} + \frac{1}{\rho}\mathbf{y}_i^{(k)}\right). \quad (17)$$

Summing up the update step in (15) over all i and substituting for $\boldsymbol{\mu}^{(k+1)}$ from (17), we find that

$$\sum_{i=1}^{L}\mathbf{y}_i^{(k+1)} = \sum_{i=1}^{L}\mathbf{y}_i^{(k)} + \rho\sum_{i=1}^{L}\left(\boldsymbol{\mu}_i^{(k+1)} - \boldsymbol{\mu}^{(k+1)}\right) = \mathbf{0},$$

which shows that the sum of dual variables $\mathbf{y}_i$ is zero for $k \geq 1$. Therefore, updating the auxiliary variable simplifies to

$$\boldsymbol{\mu}^{(k+1)} = \frac{1}{L}\sum_{i=1}^{L}\boldsymbol{\mu}_i^{(k+1)}. \quad (18)$$

The computational requirement at each sensor node to evaluate (7) and (16) is $\mathcal{O}((mt)^3 + (pr)^3)$ as the nodes must compute inverse of matrices of dimension $mt \times mt$ and $pr \times pr$. However, it should be noted that if $r \gg t$, then one may apply matrix inversion lemma to (16) and reduce the complexity to $\mathcal{O}((mt)^3)$.

From (16) and (18) it is evident that the sensor nodes need only $\boldsymbol{\mu}^{(k+1)}$, which is an average of $\boldsymbol{\mu}_i^{(k+1)}$ over the network, to solve the optimization problem (5). The average $\boldsymbol{\mu}^{(k+1)}$ can be computed in a distributed manner without privacy leakage.

3.2. Privacy-Preserving Consensus

Let $\boldsymbol{\gamma}_i(0) = \boldsymbol{\mu}_i^{(k+1)}$ denote the initial state at node i during the kth ADMM update. At each iteration, node i communicates with its neighbors and updates its local state as $\boldsymbol{\gamma}_i(n+1) = a_{ii}\boldsymbol{\gamma}_i(n) + \sum_{j\in\mathcal{N}_i} a_{ij}\boldsymbol{\gamma}_i(n)$, where $[\mathbf{A}]_{ij} = a_{ij}$ is the weight employed by node i for message received from node j. If the weighting matrix $\mathbf{A}$ is chosen such that its eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_L$ satisfy $\lambda_1 = 1$ and $\lambda_i < 1$ for all i in addition to $\mathbf{A}\mathbf{1} = \mathbf{1}$, then we have $\lim_{n\to\infty}\boldsymbol{\gamma}_i(n) = \frac{1}{L}\sum_{i=1}^{L}\boldsymbol{\mu}_i^{(k+1)}$.

However, mischievous nodes may estimate $\boldsymbol{\mu}_i^{(k+1)}$ from the messages shared by its neighbors and then infer the local information. To prevent this, we adapt the privacy-preserving average consensus algorithm proposed in [20] to compute $\boldsymbol{\mu}^{(k+1)}$ at each node in a secure manner. First, the fusion center determines the node connectivity such that $\mathcal{N}_i \cup \{i\} \nsubseteq \mathcal{N}_j \cup \{j\}$ for all $i \neq j$, $i, j = 1, \ldots, L$ and weighting coefficient matrix $\mathbf{A} = \frac{1}{L-1}(\mathbf{E} - \mathsf{diag}(\{L-1-|\mathcal{N}_i|\})_{i=1}^{L})$, and then feeds back these information to the respective nodes.

Next, at consensus iteration n, the ith node generates a random variable $\mathbf{v}_i(n)$ with normal distribution $\mathcal{CN}(\mathbf{0}, \eta\boldsymbol{\gamma}_i(n))$ and $\mathbb{E}[\mathbf{v}_i(l)\mathbf{v}_j(n)^H] = \mathbf{0}$ for $l \neq n$ and $i \neq j$, where η is the *privacy parameter*. Each node perturbs the message shared with its neighbor by adding a noise $\mathbf{w}_i(n)$ that is obtained as

$$\mathbf{w}_i(n) = \phi^n\mathbf{v}_i(n) - \phi^{n-1}\mathbf{v}_i(n-1) \quad (19)$$

Algorithm 1 Privacy-Preserving Distributed Precoder Design

At fusion center:
1: Determine neighborhood of the nodes such that $\mathcal{N}_i \cup \{i\} \nsubseteq \mathcal{N}_j \cup \{j\}$ for all $i \neq j$, $i = 1, \ldots, L$, and $j = 1, \ldots, L$.
2: Find weight matrix $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_L]$ and feedback $\mathbf{a}_i$ and neighborhood $\mathcal{N}_i$ to node i.
3: Broadcast $\mathbf{D}$ to sensors and initialize $0 < \phi < 1, \alpha^{(0)} = 1$, $\{\lambda_i^{(0)}\}_{i=1}^{L} = 0$, $\{\mathbf{y}_i^{(0)}\}_{i=1}^{L} = \mathbf{0}$, and $\boldsymbol{\mu}^{(0)} = \mathbf{0}$.
At node i:
4: **for** $k = 0, 1, 2, \ldots$ **do**
5: Determine $\boldsymbol{\mu}_i^{(k+1)}$ from (16) and $\lambda_i^{(k+1)}$ via (8).
6: Compute $\mathbf{b}_i{}^{(k+1)}$ from (7) and $\alpha^{(k+1)}$ from (9).
7: Set $\boldsymbol{\gamma}_i(0) = \boldsymbol{\mu}_i^{(k+1)}$ and $\mathbf{v}_i(-1) = \mathbf{0}$
8: **for** $n = 0, 1, 2, \ldots$ **do**
9: Generate $\mathbf{v}_i(n) \sim \eta\mathcal{CN}(\mathbf{0}, \eta\mathsf{diag}(\boldsymbol{\gamma}_i(n)))$
10: Obtain $\mathbf{w}_i(n) = \phi^n\mathbf{v}_i(n) - \phi^{n-1}\mathbf{v}_i(n-1)$
11: Perturb the message $\boldsymbol{\gamma}_i^+(n) = \boldsymbol{\gamma}_i(n) + \mathbf{w}_i(n)$
12: Receive $\boldsymbol{\gamma}_j^+(n)$ from neighbors
13: Update $\boldsymbol{\gamma}_i(n+1) = a_{ii}\boldsymbol{\gamma}_i^+(n) + \sum_{j\in\mathcal{N}_i} a_{ij}\boldsymbol{\gamma}_j^+(n)$
14: **end for**
15: Update $\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\gamma}_i(n+1)$ and update $\mathbf{y}_i^{(k+1)}$ using (15).
16: **end for**
17: Forward α to FC and obtain $\{\mathbf{B}_i\} \in \mathbb{C}^{t\times m}$ from $\{\mathbf{b}_i\} \in \mathbb{C}^{tm}$

and $0 < \phi < 1$ is a constant. Let us define the perturbed state as $\boldsymbol{\gamma}_i^+(n) = \boldsymbol{\gamma}_i(n) + \mathbf{w}_i(n)$. The new state is updated through average consensus protocol given by

$$\boldsymbol{\gamma}_i(n+1) = a_{ii}\boldsymbol{\gamma}_i^+(n) + \sum_{j\in\mathcal{N}_i} a_{ij}\boldsymbol{\gamma}_j^+(n), \quad (20)$$

and at each node we have $\lim_{n\to\infty}\boldsymbol{\gamma}_i(n+1) = \boldsymbol{\mu}^{(k+1)}$.

Finally, the nodes update their precoders according to (7) and the fusion rule according to (9). It should be noted that the FC only requires α^* from the sensor nodes to calculate the fusion rule $\mathbf{W}^H = (1/\alpha^*)\mathbf{D}^\dagger$.

The distributed algorithm described above is summarized in Algorithm 1. To prove that the proposed algorithm is privacy-preserving, without loss of generality, we consider node L as the eavesdropper trying to infer the state of other nodes. The eavesdropper computes a maximum likelihood estimate of $\boldsymbol{\mu}_i^{(k+1)}$, $i \neq L$, given the information $\mathcal{I}(n) \triangleq \{\boldsymbol{\gamma}_L(0), \boldsymbol{\gamma}_{j_1}^+(n), \ldots, \boldsymbol{\gamma}_{j_{|\mathcal{N}_L|}}^+(n)\}$ from its neighbors, where $\mathcal{N}_L = \{j_1, \ldots, j_{|\mathcal{N}_L|}\}$, and the corresponding estimation error covariance matrix is denoted by $\boldsymbol{\Xi}_i$.

Definition 1. *The privacy of node i is breached if the estimation error covariance $\boldsymbol{\Xi}_i = \mathbf{0}$ and the privacy measure of node i is $\epsilon_i \triangleq tr(\boldsymbol{\Xi}_i)$ [20].*

Theorem 1. *Algorithm 1 is privacy-preserving i.e., $tr(\boldsymbol{\Xi}_i) > 0$ for $i = 1, \ldots, L-1$ and it converges to the optimal solution in the mean square sense.*

Proof. Since $\mathcal{N}_i \cup \{i\} \nsubseteq \mathcal{N}_j \cup \{j\}$ for all $i \neq j$ and $i, j = 1, \ldots, L$ and $\lim_{n\to\infty}\mathbb{E}[\sum_{l_1=0}^{n}\sum_{l_2=0}^{n}\mathbf{w}_i(l_1)\mathbf{w}_i^H(l_2)] = \mathbf{0}$, proof follows from [20, Corollary 1] and [20, Theorem 5]. □

Define the reduced weighting matrix as $\tilde{\mathbf{A}} = \mathbf{A}(1 : L-1, 1 : L-1) \otimes \mathbf{I}_{rp}$, which is obtained after removing the Lth row and column of matrix $\mathbf{A}$, and denote $\boldsymbol{\mathcal{A}} = (\mathbf{I}_{rp(L-1)} - \tilde{\mathbf{A}})^{-1}$. Let

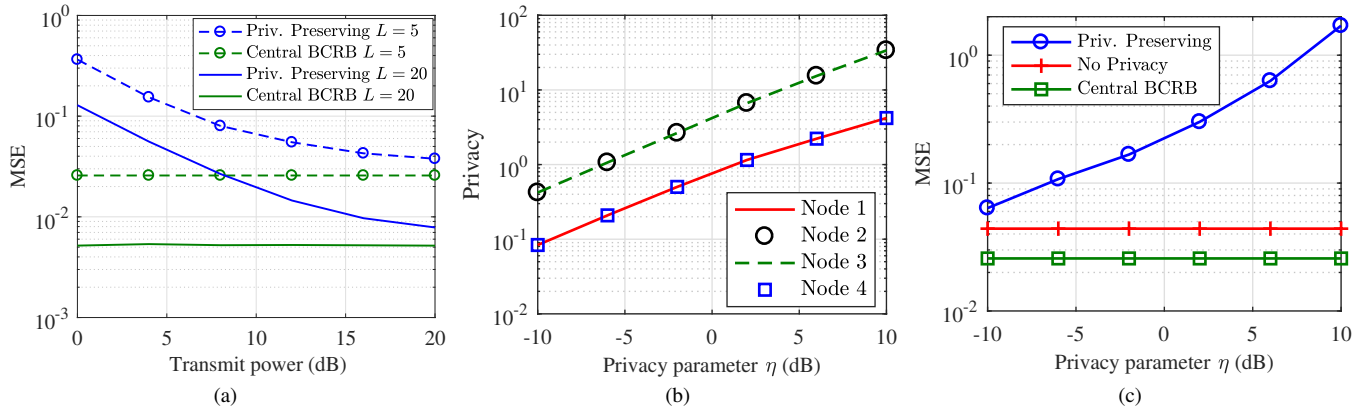

Fig. 1. (a) MSE vs. transmit power $P_i = P$, $\forall i$ and $\eta = 0.1$, (b) Privacy vs η for $P = 15$ dB and $L = 5$ sensors, (c) MSE as a function of privacy parameter η in the variables shared across nodes for $P = 15$ dB and $L = 5$ sensors.

$\mathbf{C} \triangleq [\mathbf{e}_{j_1} \otimes \mathbf{I}_{rp}, \ldots, \mathbf{e}_{j_{|\mathcal{N}_L|}} \otimes \mathbf{I}_{rp}]^T \in \mathbb{R}^{rp|\mathcal{N}_L| \times rp(L-1)}$, where $\mathbf{e}_j \in \mathbb{R}^{L-1}$ is a vector whose jth entry equals unity with the rest of the entries set to zero.

Define covariance matrix $\mathbf{\Gamma}_n \triangleq \eta\,\mathsf{diag}([\boldsymbol{\gamma}_1^T(n), \ldots, \boldsymbol{\gamma}_L^T(n)]^T)$, $\mathbf{U}_n \triangleq \mathbf{C}^T(\mathbf{C}\mathbf{\Gamma}_n\mathbf{C}^T)^{-1}\mathbf{C}$, $\mathbf{V}_n \triangleq \mathbf{I}_{rp(L-1)} - \mathbf{\Gamma}_n^{1/2}\mathbf{U}_n\mathbf{\Gamma}_n^{1/2}$, and $\mathbf{S}_n \triangleq \boldsymbol{\mathcal{A}}\mathbf{U}_n\boldsymbol{\mathcal{A}}$. The eigenvalue decomposition of matrix $\mathbf{S}_n$ is given by

$$\mathbf{S}_n = [\mathbf{F}_{n,1}, \mathbf{F}_{n,2}]\mathsf{diag}(\mathbf{\Lambda}_n, \mathbf{0})[\mathbf{F}_{n,1}, \mathbf{F}_{n,2}]^H \in \mathbb{R}^{rp(L-1)\times rp(L-1)},$$

where $\mathbf{F}_{n,2} \in \mathbb{C}^{rp(L-1)\times rp(L-|\mathcal{N}_L|-1)}$ is the matrix of eigenvectors corresponding to eigenvalue zero. The next result derives the privacy guarantees offered by Algorithm 1.

Theorem 2. *The privacy measure for node i is given by $\epsilon_i = tr(\mathbf{\Xi}_i) = tr((\mathbf{e}_i^T \otimes \mathbf{I}_{rp})\boldsymbol{\mathcal{P}}(\mathbf{e}_i \otimes \mathbf{I}_{rp}))$, where $\mathbf{e}_i \in \mathbb{R}^{L-1}$ and $\boldsymbol{\mathcal{P}} = \lim_{n\to\infty}\boldsymbol{\mathcal{P}}_n$ is determined from the recursive equations*

$$\boldsymbol{\mathcal{P}}_n = \mathbf{F}_{n,2}\left[\mathbf{F}_{n,2}^H\boldsymbol{\mathcal{A}}\mathbf{Y}_n\boldsymbol{\mathcal{A}}\mathbf{F}_{n,2}\right]^{-1}\mathbf{F}_{n,2}^H, \tag{21}$$

with $\mathbf{Y}_{n+1} = \mathbf{Y}_0 + \phi^{-2}\tilde{\mathbf{A}}\left[\mathbf{Y}_n^+ - \mathbf{Y}_n^+\left(\phi^2\mathbf{I} + \mathbf{Y}_n^+\right)^{-1}\mathbf{Y}_n^+\right]\tilde{\mathbf{A}}$, $\mathbf{Y}_n^+ = \mathbf{\Gamma}_n^{1/2}\mathbf{V}_n\mathbf{Y}_n\mathbf{V}_n\mathbf{\Gamma}_n^{1/2}$, $\boldsymbol{\mathcal{A}} = (\mathbf{I}_{rp(L-1)} - \tilde{\mathbf{A}})^{-1}$, *and* $\mathbf{Y}_0 = \tilde{\mathbf{A}}\mathbf{U}_n\tilde{\mathbf{A}}$.

Proof. Follows from generalizing [20, Lemma 6] for a noise with distribution $\mathbf{v}_i(n) \sim \mathcal{CN}(\mathbf{0}, \eta\mathsf{diag}(\boldsymbol{\gamma}_i(n)))$ with $\boldsymbol{\gamma}_i(n)$ converging to a constant. □

4. SIMULATION RESULTS

The simulation setup consists of a WSN with L sensors in ring topology with each node having two neighbors. The sensor nodes employ $t = 2$ antennas and the FC employs $r = 2$ antennas. We assume $r \times t$ Rayleigh fading MIMO channel between the sensors and FC. The FC noise covariance matrix is set as $\mathbf{R}_{\mathsf{FC}} = \mathbf{I}_r$. The parameter is assumed to be Gaussian with $\boldsymbol{\theta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_p)$ and $p = 2$. The sensor nodes acquire observations of dimension $m = 2$. The elements of the observation matrix $\{\mathbf{A}_i\}$ are generated as i.i.d. standard normal random variables. The observation noise covariance matrix at the sensor nodes $\{\mathbf{\Sigma}_i\}$ is set to $0.1\mathbf{I}_m$. The response matrix $\mathbf{D}$ is chosen as $P\mathbf{I}_p$ and ADMM penalty parameter is set to $\rho = 4$. The number iterations for ADMM and average consensus is set to $k_{\max} = 30$ and $n_{\max} = 40$, respectively. The constant ϕ is set to 0.9.

We compare the mean square error (MSE) performance of the proposed approach with the Bayesian Cramer-Rao bound (BCRB) on parameter estimation in the ideal scenario when the observations $\mathbf{x} = [\mathbf{x}_1^T, \ldots, \mathbf{x}_L^T]^T$ are available perfectly at the FC. The BCRB is given by [21]

$$\mathsf{BCRB} = \mathrm{tr}((\mathbf{R}_\theta^{-1} + \mathbf{A}^H\mathbf{\Sigma}^{-1}\mathbf{A})^{-1}), \tag{22}$$

where $\mathbf{\Sigma} \triangleq \mathsf{diag}(\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_L)$ denotes a block diagonal matrix with matrices $\mathbf{\Sigma}_i$ on the ith diagonal and $\mathbf{A} \triangleq [\mathbf{A}_1^T, \mathbf{A}_2^T, \ldots, \mathbf{A}_L^T]^T$.

Fig. 1a illustrates the MSE of the estimate at the FC for varying transmit power at the sensor nodes $P_i = P$, $\forall i$ and privacy parameter $\eta = 0.1$. We can observe that the distributed algorithm approaches the BCRB as the transmit power increases, which validates that the approximation in (4) is tight for high transmit power. For moderate to high transmit power, it shows that the proposed method yields good accuracy in addition to privacy guarantees.

Fig. 1b shows the privacy guarantees offered by Algorithm 1 for a WSN with $L = 5$ sensors when the variance of perturbation noise is controlled through *privacy parameter* η. The plot illustrates that $\mathrm{tr}(\mathbf{\Xi}_i) > 0$, for all i, and hence no breach of privacy. More importantly, it can been seen that as the number of hops between the eavesdropper (node 5) and the target node increases, larger privacy guarantees can be assured. The privacy value at node 1 is equal to node 4 and at node 2 is equal to node 3 since they are reachable from the eavesdropper with the same number of hops.

In Fig. 1c we plot the MSE achieved by the proposed method as a function of privacy parameter η. It can be observed that higher privacy results in poor estimation performance. The figure captures the trade-off between security against eavesdropper and MSE when the proposed method is employed for decentralized estimation.

5. CONCLUSION

We proposed a privacy-preserving precoding scheme for decentralized estimation where the sensor nodes are not willing to share the CSI and observation models with other entities in the network due to privacy concerns. The proposed distributed algorithm computes the precoders and fusion rule in a secure manner by exchanging perturbed dual variables to other nodes. Privacy limits offered by the proposed algorithm are derived and the trade-off between privacy and estimation accuracy is shown through numerical simulations.

6. REFERENCES

[1] A. Behbahani, A. Eltawil, and H. Jafarkhani, "Linear decentralized estimation of correlated data for power constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 16, no. 11, pp. 6003–6016, Nov. 2012.

[2] A. Shirazinia, S. Dey, D. Ciuonzo, and P. Salvo Rossi, "Massive MIMO for decentralized estimation of a correlated source," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2499–2512, May 2016.

[3] S. Liu, S. Kar, M. Fardad, and P. K. Varshney, "Optimized sensor collaboration for estimation of temporally correlated parameters," *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6613–6626, Dec. 2016.

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

[5] J. Akhtar and K. Rajawat, "Distributed sequential estimation in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 86–100, Jan. 2018.

[6] Z. Li and T. J. Oechtering, "Privacy-constrained parallel distributed Neyman-Pearson test," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 77–90, Mar. 2017.

[7] J. Guo, U. Rogers, X. Li, and H. Chen, "Secrecy constrained distributed detection in sensor networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 2, pp. 378–391, June 2018.

[8] X. He, W. P. Tay, and M. Sun, "Privacy-aware decentralized detection using linear precoding," in *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jul. 2016, pp. 1–5.

[9] X. Guo, A. S. Leong, and S. Dey, "Estimation in wireless sensor networks with security constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 2, pp. 544–561, Apr. 2017.

[10] X. Guo, A. S. Leong, and S. Dey, "Distortion outage minimization in distributed estimation with estimation secrecy outage constraints," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 12–28, Mar. 2017.

[11] C. Goken and S. Gezici, "ECRB-based optimal parameter encoding under secrecy constraints," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3556–3570, Jul. 2018.

[12] A. De Maio, S. De Nicola, Y. Huang, S. Zhang, and A. Farina, "Code design to optimize radar detection performance under accuracy and similarity constraints," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5618–5629, Nov. 2008.

[13] Y. Huang, A. De Maio, and S. Zhang, *Semidefinite programming, matrix decomposition, and radar code design*, pp. 192–228, Cambridge University Press, 2009.

[14] C. Y. Chen and P. P. Vaidyanathan, "MIMO radar waveform optimization with prior information of the extended target and clutter," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3533–3544, Sep. 2009.

[15] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Nov. 2012.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, NY, USA, 2004.

[17] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 214–226, Jan. 2004.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[19] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, New Jersey, 2000.

[20] Y. Mo and R. M. Murray, "Privacy preserving average consensus," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 753–765, Feb. 2017.

[21] H. L. Van Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley-IEEE Press, 2007.