U.S. ARMY RDECOM®

# Distributed Maximum Likelihood using Dynamic Average Consensus

ARL

Jemin George, jemin.george.civ@mail.mil, U.S. Army Research Laboratory, Adelphi, MD 20783, USA

## Problem Formulation

- Consider a sensor network
  - Connected undirected network of $n$ nodes $\mathcal{G}(\mathcal{V},\mathcal{E})$
  - Adjacency matrix: $\mathcal{A} \triangleq [a_{ij}] \in \{0,1\}^{n\times n}$
  - Degree matrix: $\Delta \triangleq \text{diag}(\mathcal{A}\mathbf{1}_n)$
  - Incidence matrix: $\mathcal{B} = [b_{ij}] \in \{-1,0,1\}^{n\times\ell}$
  - Laplacian: $\mathcal{L} = \mathcal{D} - \mathcal{A} = \mathcal{B}\mathcal{B}^T$
- Sensor measurements

$$\boldsymbol{z}_i = \mathbf{h}_i(\boldsymbol{\theta}) + \mathbf{w}_i$$

where $\boldsymbol{\theta} \in \mathbb{R}^r$, $\mathbf{w}_i \sim \mathrm{N}(\mathbf{0}, R_i)$, and $\mathbf{h}_i(\cdot): \mathbb{R}^r \mapsto \mathbb{R}^m$ is local.

- Maximum likelihood solution

$$\hat{\boldsymbol{\theta}}_{\text{ML}} \triangleq \min_{\boldsymbol{\theta}} \sum_{i=1}^{n} f_i(\boldsymbol{\theta})$$

where $f_i(\boldsymbol{\theta}) \triangleq \frac{1}{2}(\mathbf{z}_i - \mathbf{h}_i(\boldsymbol{\theta}))^T R_i^{-1} (\mathbf{z}_i - \mathbf{h}_i(\boldsymbol{\theta}))$.

- Objective is to solve the above problem using only local interactions such that each agent recovers the global minimizer.

## Existing Algorithms

- Distributed Subgradient Method
- Distributed Alternating Direction Method of Multipliers
- Zero Gradient Sum (ZGS) algorithms
- Exact first-order algorithm (EXTRA)
- Distributed dual subgradient methods

## Main Results

Centralized solution

$$\hat{\boldsymbol{\theta}}^{k+1} = \hat{\boldsymbol{\theta}}^{k} - \alpha \sum_{i=1}^{n} \nabla f_i\left(\hat{\boldsymbol{\theta}}^{k}\right), \quad \hat{\boldsymbol{\theta}}^{0}$$

where $\nabla f_i\left(\hat{\boldsymbol{\theta}}^{k}\right) = \left(H_i^k\right)^T R_i^{-1}\left(\mathbf{h}_i\left(\hat{\boldsymbol{\theta}}^{k}\right) - \mathbf{z}_i\right)$

$\alpha > 0$ is the step size, and $H_i^k = \dfrac{\partial \mathbf{h}_i}{\partial \boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}^{k})$

Distributed implementation of the algorithm requires

- Agents reach consensus on initial value
- Calculating the global gradient sum (average)

Proposed solution

- Static average consensus algorithm reach agreement on initial value
- Dynamic average consensus algorithms to reach consensus on the descent direction

## Static average consensus

- Exponential convergence of the algorithm

**Lemma 1** *Let the vector of the values of all the nodes* $\mathbf{y}(t) \in \mathbb{R}^n$ *be the solution of the following differential equation:*

$$\dot{\mathbf{y}}(t) = -\mathcal{L}\mathbf{y}(t), \quad \mathbf{y}(t_0) = \mathbf{y}_0.$$

*Then, all the nodes of the graph globally, asymptotically reach an average consensus value* $\bar{y} = \frac{1}{n}\mathbf{1}_n^T\mathbf{y}_0$ *with an exponential rate of* $\kappa = \lambda_2(\mathcal{L})$*, i.e.,* $\|\boldsymbol{\delta}(t)\| \leq \|\boldsymbol{\delta}(t_0)\| \exp(-\kappa t)$*, where* $\boldsymbol{\delta}(t) = \mathbf{y}(t) - \bar{y}\mathbf{1}_n$.

## Dynamic average consensus

- Continuous-time version of centralized algorithm

$$\dot{\hat{\boldsymbol{\theta}}}(t) = -\alpha \sum_{i=1}^{n} \nabla f_i\left(\hat{\boldsymbol{\theta}}(t)\right), \quad \hat{\boldsymbol{\theta}}(t_0) = \hat{\boldsymbol{\theta}}^0.$$

- Let $\phi_i(t) = \nabla f_i\left(\hat{\boldsymbol{\theta}}_i(t)\right) \in \mathbb{R}^r$, $\hat{\boldsymbol{\theta}}_i(t)$'s are the local solutions
- DAC algorithm allows the agents to reach consensus on the average

$$\bar{\phi}(t) = \frac{1}{n}\sum_{i=1}^{n} \phi_i(t)$$

- DAC-error:

$$\tilde{\mathbf{x}}(t) \triangleq \mathbf{x}(t) - \mathbf{1}_n \otimes \bar{\phi}(t)$$

where $x_i(t) \in \mathbb{R}^r$ denote node $i$'s estimate of $\bar{\phi}(t)$,

$\boldsymbol{\phi}(t) \triangleq \left[\phi_1^T(t) \quad \ldots \quad \phi_n^T(t)\right]^T \in \mathbb{R}^{nr}$,

$\mathbf{x}(t) \triangleq \left[x_1^T(t) \quad \ldots \quad x_n^T(t)\right]^T \in \mathbb{R}^{nr}$

- Proposed DAC algorithm

$$\dot{\mathbf{z}}(t) = -\beta \text{sgn}\left\{\left(\mathcal{B}^T \otimes I_r\right)\mathbf{x}(t)\right\}, \quad \mathbf{z}(t_0) = \mathbf{z}_0,$$

$$\mathbf{x}(t) = \left(\mathcal{B} \otimes I_r\right)\mathbf{z}(t) + \boldsymbol{\phi}(t),$$

$\mathbf{z}(t) \in \mathbb{R}^{r\ell}$ is the internal estimator state

**Theorem 1** *For any connected undirected network, given* $\sup_{t\in[t_0,\infty)} \|\mathcal{B}^T\dot{\boldsymbol{\phi}}(t)\|_\infty \leq \varphi < \infty$*, the robust DAC algorithm guarantees that the average consensus error,* $\tilde{\mathbf{x}}(t)$*, converges to zero in finite time for any initial condition* $\mathbf{z}_0$*, if the estimator input gain* $\beta$ *is selected that* $\beta > \frac{\varphi}{\lambda_2(\mathcal{B}\mathcal{B}^T)}$*, where* $\lambda_2(\cdot)$ *is the smallest non-zero eigenvalue. More specifically, we have* $\tilde{\mathbf{x}}(t) = 0$ *for all* $t \geq \frac{\|\tilde{\mathbf{x}}(t_0)\|_2}{\sqrt{\lambda_2(\mathcal{B}\mathcal{B}^T)}}$.

Approved for Public Release. Distribution Unlimited.

## Distributed Algorithm

$$\dot{\Theta}(t) = -\gamma\left(\mathcal{L} \otimes I_r\right)\Theta(t) - \hat{\alpha}\mathbf{x}(t), \quad \Theta(t_0)$$

$$\dot{\mathbf{z}}(t) = -\beta \text{sgn}\left\{\left(\mathcal{B}^T \otimes I_r\right)\mathbf{x}(t)\right\}, \quad \mathbf{z}(t_0) = \mathbf{z}_0,$$

$$\mathbf{x}(t) = \left(\mathcal{B} \otimes I_r\right)\mathbf{z}(t) + \boldsymbol{\phi}(t),$$

where $\gamma$ is the static average consensus gain, $\hat{\alpha}$ is the step size, $\Theta(t) \triangleq \left[\hat{\boldsymbol{\theta}}_1^T(t), \ldots, \hat{\boldsymbol{\theta}}_n^T(t)\right]^T \in \mathbb{R}^{nr}$

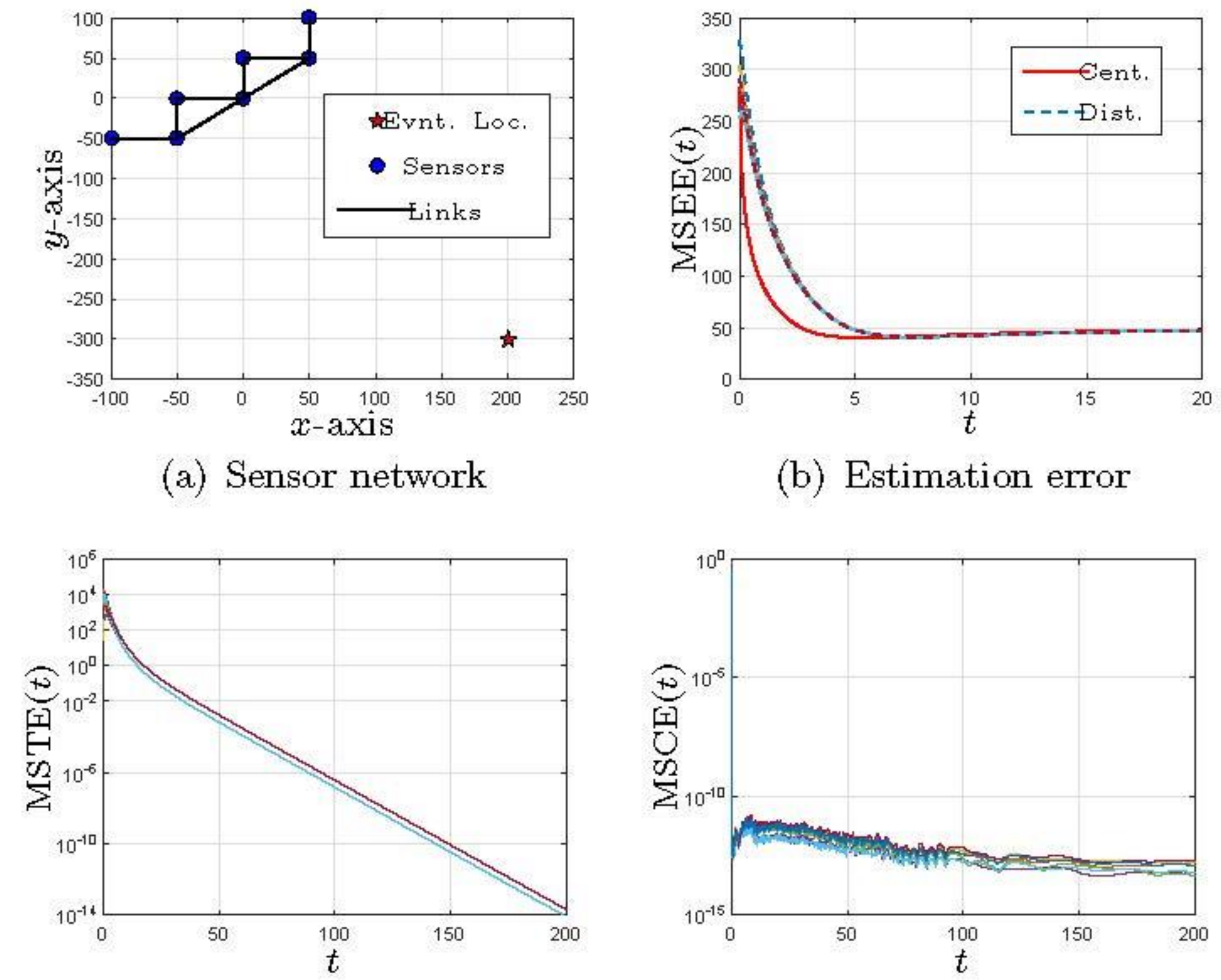
**Theorem 2** *For any connected undirected network, the distributed algorithm converges to the centralized solution trajectories exponentially fast for all* $t \geq \frac{\|\tilde{\mathbf{x}}(t_0)\|_2}{\sqrt{\lambda_2(\mathcal{B}\mathcal{B}^T)}}$.

## Example

- Distributed event localization using acoustic sensor network (direction of arrival measurements)

$$z_i = \arctan\left(\frac{T_y - S_i^y}{T_x - S_i^x}\right) + w_i$$

Here $\boldsymbol{\theta} = [T_x, T_y]^T = [200, -300]^T$, $n = 7$, and $R_i = 10^{-2}, \forall i \in \{1,\ldots,7\}$




(a) Sensor network (b) Estimation error (c) Tracking error (d) Consensus error

Figure 1: Simulation scenario and mean-square errors obtained from $10^3$ Monte Carlo runs.

## Conclusions

- Not an other distributed optimization algorithm
- Can be easily applied to higher order schemes
- The proposed distributed algorithm recovers the centralized localization accuracy exponentially fast
- Future work: Consider privacy preserving protocols & event-triggered communication schemes