Marginal Density Averaging for Distributed Node Localization from Local Edge Measurements

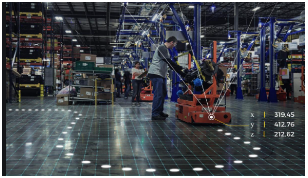
Presented by: Parth Paritosh
In collaboration with: Nikolay Atanasov, Sonia Martínez

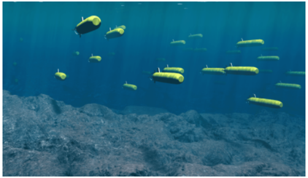
University of California San Diego (UCSD)
Mechanical and Aerospace Engineering

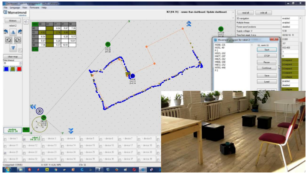
IEEE CDC 2020

Distributed sensing with relative measurements

What are networks based on relative measurements?





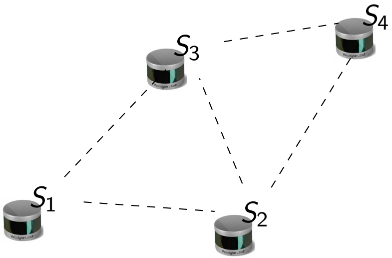


Key features for inference algorithms:

- Rely on localized signals
- Fast computation and low storage at nodes
- Communication efficiency
- Large scale networks with temporal variations

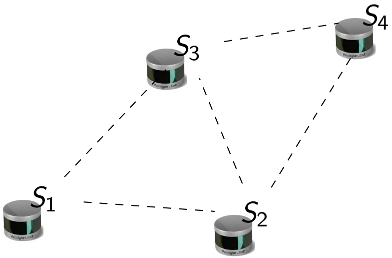
Problem setup

- Sensing agents $\mathcal{N} = (S_1, \cdots, S_n)$ with neighbor set $\mathcal{N}_i$



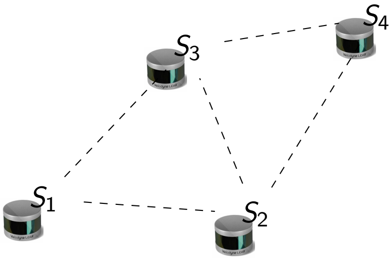
Problem setup

- Sensing agents $\mathcal{N} = (S_1, \cdots, S_n)$ with neighbor set $\mathcal{N}_i$
- Agent state: $\boldsymbol{x}_i \in \mathbb{R}^m$



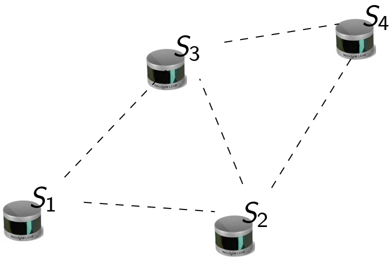
Problem setup

- Sensing agents $\mathcal{N} = (S_1, \cdots, S_n)$ with neighbor set $\mathcal{N}_i$
- Agent state: $\boldsymbol{x}_i \in \mathbb{R}^m$
- Neighbor based measurement models $\mathrm{p}_i(\boldsymbol{z}_i | \{\boldsymbol{x}_j\}_{j \in \mathcal{N}_i}) = \prod_{j \in \mathcal{N}_i} \mathrm{p}_i(\boldsymbol{z}_{ij} | \boldsymbol{x}_i, \boldsymbol{x}_j)$



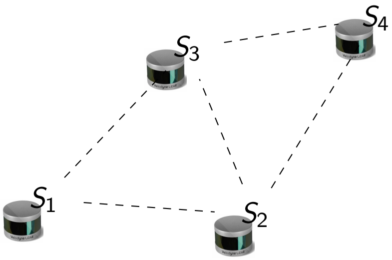
Problem setup

- Sensing agents $\mathcal{N} = (S_1, \cdots, S_n)$ with neighbor set $\mathcal{N}_i$
- Agent state: $\boldsymbol{x}_i \in \mathbb{R}^m$
- Neighbor based measurement models $\mathrm{p}_i(\boldsymbol{z}_i | \{\boldsymbol{x}_j\}_{j\in\mathcal{N}_i}) = \prod_{j\in\mathcal{N}_i} \mathrm{p}_i(\boldsymbol{z}_{ij} | \boldsymbol{x}_i, \boldsymbol{x}_j)$
- Local communication network, (Weighted adjacency matrix: A)



Problem setup

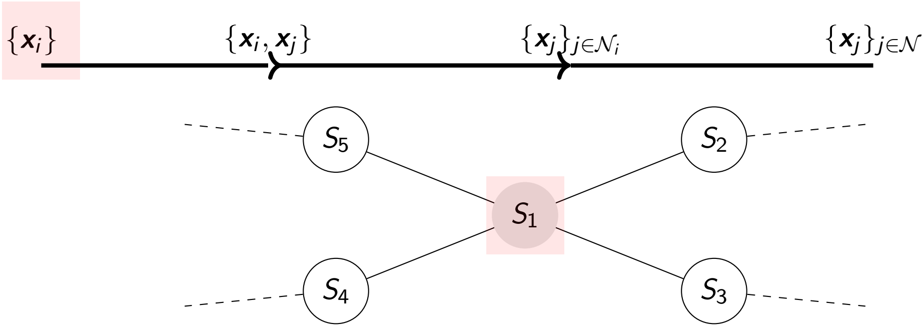
- Sensing agents $\mathcal{N} = (S_1, \cdots, S_n)$ with neighbor set $\mathcal{N}_i$
- Agent state: $\boldsymbol{x}_i \in \mathbb{R}^m$
- Neighbor based measurement models $\mathrm{p}_i(\boldsymbol{z}_i | \{\boldsymbol{x}_j\}_{j \in \mathcal{N}_i}) = \prod_{j \in \mathcal{N}_i} \mathrm{p}_i(\boldsymbol{z}_{ij} | \boldsymbol{x}_i, \boldsymbol{x}_j)$
- Local communication network, (Weighted adjacency matrix: A)



- How to find the true value of agent states $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ with relative measurements received over the given communication network?

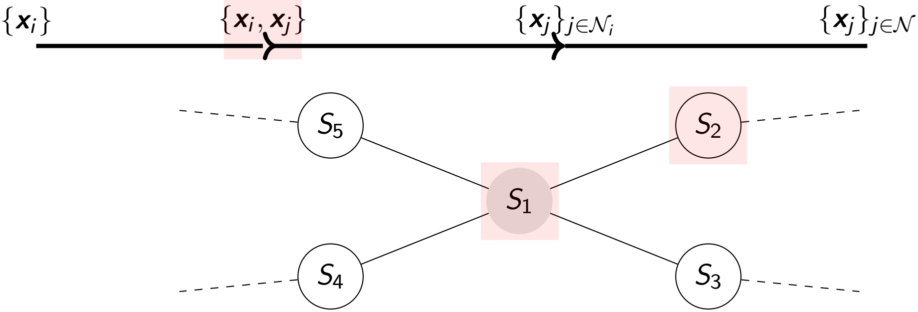
Agent domains for distributed estimation

How do we select the domain $\mathcal{X}_i$ of agent i's estimate?



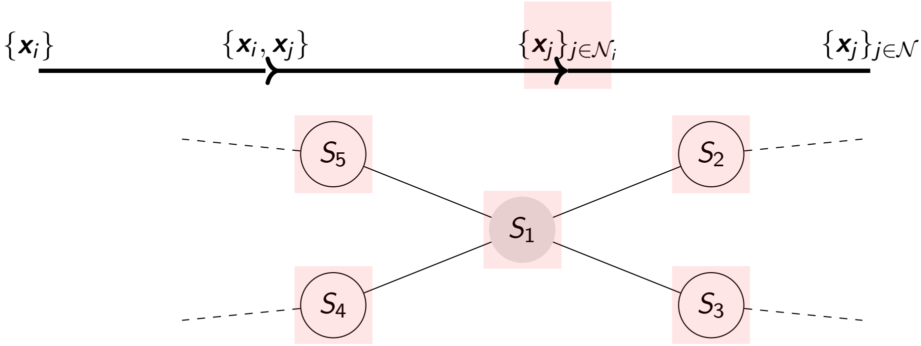
Agent domains for distributed estimation

How do we select the domain $\mathcal{X}_i$ of agent i's estimate?



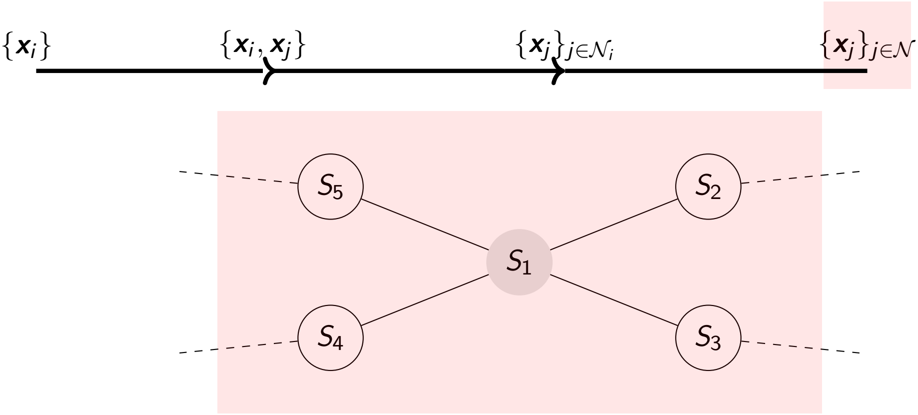
Agent domains for distributed estimation

How do we select the domain $\mathcal{X}_i$ of agent i's estimate?



Agent domains for distributed estimation

How do we select the domain $\mathcal{X}_i$ of agent i's estimate?



Existing solutions

How do we select the domain $\mathcal{X}_i$ of agent i's estimate?

$\{\boldsymbol{x}_i\}$ — $\{\boldsymbol{x}_i, \boldsymbol{x}_j\}$ — $\{\boldsymbol{x}_j\}_{j \in \mathcal{N}_i}$ — $\{\boldsymbol{x}_j\}_{j \in \mathcal{N}}$

Belief propagation - - - - - - - - - - - - - - - - - - ? - Geometric updates

- Yedidia et al.[1](2003): Learning marginal density at each agent state via Belief propagation in forest type graphs
- Nedich et al.[2](2017): Convergence rates of geometric averaging of inferences for the decentralized communication problem

[1] Jonathan S. Yedidia, William T. Freeman, and Yair Weiss. "Understanding belief propagation and its generalizations." Exploring artificial intelligence in the new millennium 8 (2003): 236-239.

[2] Angelia Nedić, Alex Olshevsky, and César A. Uribe. "Fast convergence rates for distributed non-bayesian learning." IEEE Transactions on Automatic Control 62.11 (2017): 5538-5553.

Research question

How to design an inference algorithm to learn true value of variables in $\mathcal{X}_i$ at agent i using noisy measurements and neighbor estimates available at each time step?

Outline

1 Introduction
- Problem setup
- Research Question

2 Objective and proposed algorithms
- Centralized
- Decentralized communication
- Decentralized communication and storage
- Gaussian marginal consensus algorithm

3 Simulation

4 Summary and future work

Consistent inference with optimization

- $\bar{p}(\mathcal{X})$: Unknown density over $\mathcal{X} = \boldsymbol{x}_{1:n}$
- $\mathrm{q}(\boldsymbol{z}_{1:n,t}) = \prod_{i\in\mathcal{N}} \mathrm{q}_i(\boldsymbol{z}_{i,t}|\mathcal{X}^*)$: Sampled data generating density
- $\mathrm{q}(\boldsymbol{z}_{1:n,t}|\mathcal{X}) = \prod_{i\in\mathcal{N}} \mathrm{q}(\boldsymbol{z}_{i,t}|\mathcal{X}_i)$: Known likelihood functions for agent i

Divergence based objective

- KL-divergence objective function

$$\arg\min_{\bar{p}} \left\{ \mathop{\mathbb{E}}_{\mathcal{X}\sim\bar{p}} [\mathrm{D}_{\mathrm{KL}}(\mathrm{q}(\boldsymbol{z}_{1:n,t}) || \, \mathrm{q}(\boldsymbol{z}_{1:n,t}|\mathcal{X}))] \right\}$$

$$\equiv \arg\min_{\bar{p}} \left\{ \mathop{\mathbb{E}}_{\boldsymbol{z}_{1:n,t}\sim\mathrm{q}(\boldsymbol{z}_{1:n,t})} \mathop{\mathbb{E}}_{\mathcal{X}\sim\bar{p}} [-\log(\mathrm{q}(\boldsymbol{z}_{1:n,t}|\mathcal{X}))] \right\}$$

Objective function

Centralized objective

- Time-averaged objective function

$$\arg\min_{\bar{p}\in\mathcal{F}} \left\{ \frac{1}{T} \sum_{t=1}^{T} F[\bar{p}; \boldsymbol{z}_{1:n,t}] \right\}$$

$$F[\bar{p}; \boldsymbol{z}_{1:n,t}] = \mathop{\mathbb{E}}_{\mathcal{X}\sim\bar{p}}[-\log(\mathrm{q}(\boldsymbol{z}_{1:n,t}|\mathcal{X}))]$$

Summable property of the objective function

Using independence between agent observations $\boldsymbol{z}_{i,t}$,

$$F[\bar{p}; \boldsymbol{z}_{1:n,t}] = \sum_{i=1}^{n} F_i[\bar{p}_i, \boldsymbol{z}_{i,t}] = \sum_{i=1}^{n} \mathop{\mathbb{E}}_{\mathcal{X}\sim\bar{p}_i}[-\log(\mathrm{q}_i(\boldsymbol{z}_{i,t}|\mathcal{X}))]$$

Centralized algorithm

- p_t: Estimated probability density function at time t
- $\frac{\delta F}{\delta p}[p_t, \boldsymbol{z}_{1:n,t}]$: Gradient of centralized objective function
- The sequence $\{\alpha_t\}$ is square-summable but non-summable.

Stochastic mirror descent

$$p_{t+1} = \arg\min_{p \in \mathcal{F}} \left\{ \left\langle \frac{\delta F}{\delta p}[p_t, \boldsymbol{z}_{1:n,t}], p \right\rangle + \frac{1}{\alpha_t} \mathrm{D}_{\mathrm{KL}}(p || p_t) \right\}$$

$$\implies p_{t+1}(\mathcal{X}) \propto \mathrm{q}(\boldsymbol{z}_{1:n,t} | \mathcal{X})^{\alpha_t} p_t(\mathcal{X})$$

Decentralized objective

Decentralized communication with full state estimates ($\mathcal{X}_i = \mathcal{X}$)

- Consensus: Ensuring that agents achieve the same estimate
- Likelihood update: Including likelihood information at each time step

$$p_i(\mathcal{X}) = \arg\min_{\bar{p}_i} \left[\mathop{\mathbb{E}}_{\boldsymbol{z}_{i,t} \sim \mathrm{q}_i(\boldsymbol{z}_{i,t})} \mathop{\mathbb{E}}_{\mathcal{X} \sim \bar{p}_i} [-\log(\mathrm{q}_i(\boldsymbol{z}_{i,t} | \mathcal{X}))] \right] \qquad \text{(Agent objective)}$$

$$p_i(\mathcal{X}) = p_j(\mathcal{X}), \quad \forall j \in \mathcal{N}_i \qquad \text{(Consensus constraint)}$$

- (Assumption) The communication network is connected and the graph adjacency matrix A satisfies $A\mathbb{1} = \mathbb{1}, A = A^\top$, and diagonal entries $A_{ii} > 0\,, \forall i \in \{1, \ldots, n\}$, where $\mathbb{1}$ is a vector of ones.

Decentralized communication algorithm

$$v_{i,t} = \frac{1}{Z^v_{i,t}} \prod_{j=1}^{n} p_{j,t}^{A_{ij}}, \quad Z^v_{i,t} = \int_{\mathbf{x}\in\mathcal{X}} \left(\prod_{j=1}^{n} p_{j,t}^{A_{ij}} \right) \qquad \text{(Mixing step)}$$

$$p_{i,t+1} = \exp\left(\alpha_t \frac{\delta F}{\delta p}\right) v_{i,t} \Bigg/ \left(\int \exp(\alpha_t \frac{\delta F}{\delta p}) v_{i,t}(\mathbf{x}) d\mathbf{x} \right) \qquad \text{(Likelihood update)}$$

Decentralized marginal objective

Decentralized communication with marginal state estimates ($\mathcal{X}_i = \{\boldsymbol{x}_j\}_{j \in \mathcal{N}_i}$)

- Consensus: Ensuring that agents achieve the same estimate on common domain $\mathcal{X}_{ij} = \mathcal{X}_i \cap \mathcal{X}_j$
- Likelihood update: Including likelihood information at each time step

$$p_i = \arg\min_{\bar{p}_i} \left[\mathop{\mathbb{E}}_{\boldsymbol{z}_{i,t} \sim \mathrm{q}_i(\boldsymbol{z}_{i,t})} \mathop{\mathbb{E}}_{\mathcal{X} \sim \bar{p}_i} [-\log(\mathrm{q}_i(\boldsymbol{z}_{i,t} | \mathcal{X}_i))] \right] \quad \text{(Agent objective)}$$

$$s.t. \quad p_i(\mathcal{X}_{ij}) = p_j(\mathcal{X}_{ij}), \quad \forall j \in \mathcal{N}_i \quad \text{(Marginal consensus constraint)}$$

Marginal consensus step

- $p_{i,t}(\mathcal{X}_i)$: Estimated density function by agent i on variables contained in $\mathcal{X}_i$
- $p_{i,t}(\mathcal{X}_{ij})$: Marginal density of $p_{i,t}(\mathcal{X}_i)$ computed over the common set of variables at agent j

Geometric marginal mixing with stochastic weights

$$p_{i,t}(\mathcal{X}_{ij}) = \int_{\mathcal{X}_i \setminus \mathcal{X}_{ij}} p_{i,t}(\mathcal{X}_i) \qquad \text{(Common marginal)}$$

$$p_{i,t}(\mathcal{X}_i | \mathcal{X}_{ij}) = \frac{p_{i,t}(\mathcal{X}_i)}{p_{i,t}(\mathcal{X}_{ij})} \qquad \text{(Conditional density)}$$

$$\tilde{p}_{ji,t}(\mathcal{X}_i) = p_{i,t}(\mathcal{X}_i | \mathcal{X}_{ij}) p_{j,t}(\mathcal{X}_{ij}) \qquad \text{(Include marginal information)}$$

$$v_{i,t}(\mathcal{X}_i) \propto \prod_{j \in \mathcal{N}_i} \tilde{p}_{ji,t}^{A_{ij}}(\mathcal{X}_i) \qquad \text{(Mixing step)}$$

Marginal consensus algorithm

$$v_{i,t}(\mathcal{X}_i) \propto \prod_{j \in \mathcal{N}_i} \left(\frac{p_{i,t}(\mathcal{X}_i)}{p_{i,t}(\mathcal{X}_{ij})} p_{j,t}(\mathcal{X}_{ij}) \right)^{A_{ij}} \qquad \text{(Mixing step)}$$

$$p_{i,t+1}(\mathcal{X}_i) = \arg\min_{p \in \mathcal{F}_m} \left\{ \alpha_t \left\langle \frac{\delta F}{\delta p}(p_{i,t}, \boldsymbol{z}_{i,t}), p \right\rangle + \mathrm{D}_{\mathrm{KL}}(p || v_{i,t}) \right\}$$

$$= \mathrm{q}_i(\boldsymbol{z}_{i,t}|\mathcal{X}_i)^{\alpha_t} v_{i,t} \Big/ \left(\int \mathrm{q}_i(\boldsymbol{z}_{i,t}|\mathcal{X})^{\alpha_t} v_{i,t}(\boldsymbol{x}) d\boldsymbol{x} \right)$$

Gaussian marginal consensus algorithm

Gaussian pdf $\phi\left(\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{bmatrix} \middle| \begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix}, \begin{bmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{bmatrix}^{-1}\right)$

Gaussian marginal consensus algorithm

Common marginal

$$p_{i,t}(\mathcal{X}_{ij}) = \int_{\mathcal{X}_i \setminus \mathcal{X}_{ij}} p_{i,t}(\mathcal{X}_i)$$

Conditional density

$$p_{i,t}(\mathcal{X}_i|\mathcal{X}_{ij}) = \frac{p_{i,t}(\mathcal{X}_i)}{p_{i,t}(\mathcal{X}_{ij})}$$

Conditional marginal product:

$$\tilde{p}_{ji,t}(\mathcal{X}_i) = p_{i,t}(\mathcal{X}_i|\mathcal{X}_{ij}) p_{j,t}(\mathcal{X}_{ij})$$

Marginal density w.r.t. $\boldsymbol{x}_1$:
$\phi\left(\boldsymbol{x}_1 \middle| \boldsymbol{\mu}_1, \left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right)^{-1}\right)$
Conditional distribution

$$(X_1|X_2 = \boldsymbol{x}_2) \sim \mathcal{N}\left(\mu_1 - \Omega_{11}^{-1}\Omega_{12}(\boldsymbol{x}_2 - \mu_2), \Omega_{11}^{-1}\right)$$

Marginal distribution of X_2: $\mathcal{N}(\bar{\mu}_2, \bar{\Omega}_{22}^{-1})$
Joint distribution of (X_1, X_2)

$$\left(\begin{bmatrix} \mu_1 + \Omega_{11}^{-1}\Omega_{12}(\mu_2 - \bar{\mu}_2) \\ \bar{\mu}_2 \end{bmatrix}, \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12}^\top & \bar{\Omega}_{22} + \Omega_{12}^\top\Omega_{11}^{-1}\Omega_{12} \end{bmatrix}^{-1}\right)$$

Gaussian marginal consensus

- Mixing step

$$v_{i,t}(\mathcal{X}_i) \propto \prod_{j \in \mathcal{N}_i} \tilde{p}_{ji,t}^{A_{ij}}(\mathcal{X}_i)$$

- Likelihood update

$$p_{i,t+1}(\mathcal{X}_i) \propto \mathrm{q}_i(\boldsymbol{z}_{i,t}|\mathcal{X}_i)^{\alpha_t} v_{i,t}(\mathcal{X}_i)$$

- Gaussians $\phi(\boldsymbol{x}|\mu_i, \Omega_i^{-1})$, $\Omega_w = \sum_{i=1}^n A_i \Omega_i$,

$$\prod_{i=1}^n \phi(\boldsymbol{x}|\mu_i, \Omega_i^{-1})^{A_i} = \phi\left(\boldsymbol{x}\middle|\Omega_w^{-1}\sum_{i=1}^n A_i \Omega_i \mu_i, \Omega_w^{-1}\right)$$

- Likelihood $\mathrm{q}_i(\boldsymbol{z}_{i,t}|\mathcal{X}_i) = \phi\left(\boldsymbol{z}_{i,t}|H_i\mathcal{X}_i, V_i\right)$,

$$\phi(\boldsymbol{z}_{i,t}|H_i\mathcal{X}_i, V_i^{-1})\phi\left(\mathcal{X}_i; \mu, \Omega_i^{-1}\right) =$$
$$\mathcal{N}\left(\left(H_i^\top V_i H_i + \Omega_i\right)^{-1}(H_i^\top V_i \boldsymbol{z}_{i,t} + \Omega_i \mu_i)\right.$$
$$\left., \left(H_i^\top V_i H_i + \Omega_i\right)^{-1}\right)$$

Simulation details

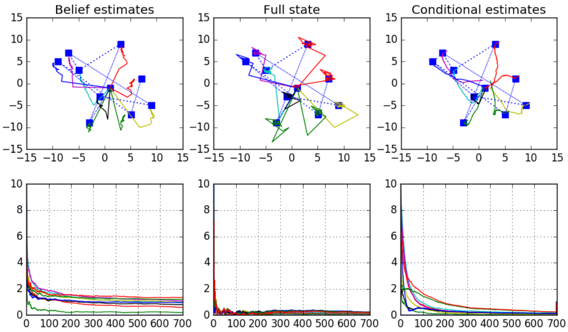
- A 10-node network with unknown locations $\boldsymbol{x} = [\boldsymbol{x}_i]_{i \in \mathcal{N}}, \boldsymbol{x}_i \in \mathbb{R}^2$.
- Observations: $\boldsymbol{z}_{ij} = (\boldsymbol{x}_j - \boldsymbol{x}_i) + \epsilon, \epsilon \sim \mathcal{N}(0, V_i), V_i = \mathbb{I}_2$
- Agent observation model $\mathrm{q}_i(\boldsymbol{z}_i|\boldsymbol{x}) = \phi(\boldsymbol{z}_i - H_i\boldsymbol{x}, V_i)$, $H_i \in \{-1, 0, 1\}^{d_z|\mathcal{X}_i| \times 2|\mathcal{X}_i|}$
- $p_{i,t}(\mathcal{X}_i) = \phi(\mathcal{X}_i|\mu_{i,t}, \Omega_{i,t}^{-1})$: Estimated normal density representing variables in $\mathcal{X}_i$ with mean $\mu_{i,t}$ and covariance $\Omega_{i,t}^{-1}$

Existing algorithms

- Belief propagation
- Full state updates
- Proposed algorithm

Self-state estimates

Belief propagation(BP), full state(FS) and marginal state estimates(MS) for a 10-agent ring network



(Row 1) Convergence to true positions
(Row 2) Estimation error across time

Effect of edge density

Estimation in multiple 10-node graphs with number of edges in $\{9, \cdots, 45\}$.

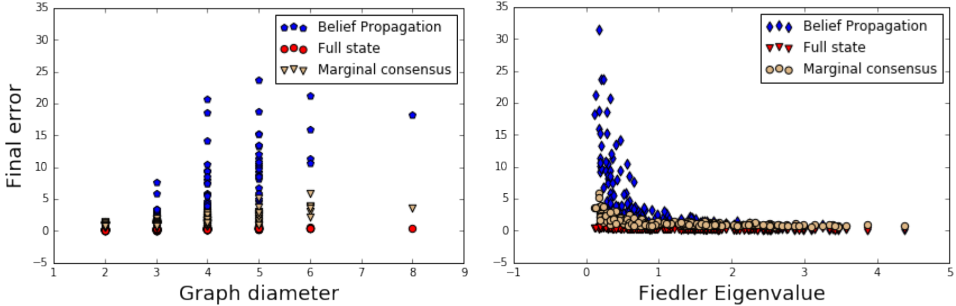


Figure: Error in self position estimates via BP, full and partial state estimation algorithms. (a) With increasing graph diameter after 500 steps. (b) With increasing connectivity captured by Fiedler Eigenvalue.

Comparing the amount of communicated information

Transmitting a d-dimensional Gaussian density requires transmitting $d + d^2$ floating point numbers.

Table: Comparing the iterations and communicated numbers for convergence to a fixed error $\epsilon = 0.1$ in a 25-node graph

	Iterations			Information units		
	BP	FS	CS	BP	FS	CS
Line	NA	18	1356	NA	2203k	1301k
100 edges	9	2	28	29.8k	846.6k	291k
287 edges	7	2	15	51k	1856k	2709k

Summary

- Introduced inference algorithm on **partial set of variables** with distributed communication
- Convergence for any connected graph
- Developed Gaussian version of the marginal consensus algorithm
- Studied simulations emphasizing trade-offs with Belief propagation and Full-state algorithms

Future work

- Convergence analysis of marginal consensus algorithm
- Extending implementation to non-Gaussian densities and particle methods

Thank you

Gaussian full state updates

$\mathcal{N}(\mu_{i,t}, \Omega_{i,t}^{-1})$: Normal density representing agent i's estimate over the space $\mathcal{X}$

$$\Omega_{i,t} = \sum_{j \in \mathcal{N}_i} \Omega_{j,t-1};\ \mu_{i,t} = \Omega_{i,t}^{-1}\Big(\sum_{j \in \mathcal{N}_i} \Omega_{j,t-1}\mu_{j,t-1}\Big).$$

Belief propagation

- $m_{t,ij}(\boldsymbol{x}_i)$: Message from agent i to agent j
- $p_{i,t}(\boldsymbol{x}_i)$: Agent i's estimate over the variable $\boldsymbol{x}_i$

$$m_{t,ij}(\boldsymbol{x}_i) = \sum_{x_i} \mathrm{q}_i(\boldsymbol{z}_{ij}|\boldsymbol{x}_i, \boldsymbol{x}_j) p_{i,t}(\boldsymbol{x}_i) \prod_{k \in \mathcal{N}_j \setminus i} m_{t-1,kj}(\boldsymbol{x}_i),$$

$$p_{i,t}(\boldsymbol{x}_i) = \frac{p_{i,t-1}(\boldsymbol{x}_i) \prod_{k \in \mathcal{N}_i} m_{ki}(\boldsymbol{x}_i)}{\sum_{j=1}^{n} p_{j,t-1}(\boldsymbol{x}_j) \prod_{k \in \mathcal{N}_j} m_{kj}(\boldsymbol{x}_j)}.$$

Gaussian belief propagation

A Gaussian BP algorithm for agents with observation model $\boldsymbol{z}_i = H\begin{bmatrix}\boldsymbol{x}_i & \boldsymbol{x}_j\end{bmatrix}^\top + \epsilon, \epsilon \sim \mathcal{N}(0_{d\times 1}, \Omega_i^z)$, with $H = \begin{bmatrix}-1, 1\end{bmatrix} \otimes \mathbb{I}_d$, where $\otimes$ is a kronecker product. The update rule for each agent is given as

$$\Omega_{jj,t} = \sum_{i\in\mathcal{N}_j} \Omega_{ij,t-1};\ \mu_{jj,t} = \Omega_{jj,t}^{-1}\left(\sum_{i\in\mathcal{N}_j} \Omega_{ij,t-1}\mu_{ij,t-1}\right),$$

which depends on the messages sent to j from $i \in \mathcal{N}_j$:

$$\Omega_{ij,t} = \begin{bmatrix}\Omega_{ii,t} - \Omega_{ji,t-1} & 0\\ 0 & 0\end{bmatrix} + H_i^\top \Omega_i^z H_i,$$

$$\mu_{ij,t} = \Omega_{ij,t}^{-1}\left(\begin{bmatrix}\sum_{k\in\{\mathcal{N}_i\setminus j\}} \Omega_{ki,t-1}\mu_{ki,t-1}\\ 0\end{bmatrix} + H_i^\top \Omega_i^z \boldsymbol{z}_{ij,t}\right).$$

Marginal averaging algorithm

We present the Gaussian estimate equivalent to the four algorithm steps in the following lemmas. Here, we denote a Gaussian random variable $\mathcal{N}(\mu, \Omega^{-1})$ with mean μ and information matrix as Ω, and its associated density function as $\phi(\cdot|\mu, \Omega^{-1})$.

Neighbor messages

The marginal density of the Gaussian pdf $\phi\left(\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}^{-1}\right)$ with respect to $\boldsymbol{x}_1$ is given as,

$$\phi(\boldsymbol{x}_1 \big| \boldsymbol{\mu}_1, \left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right)^{-1}).$$

Gaussian marginal algorithm

Pre-edge merging

Let (X_1, X_2) be random vectors represented by a joint Gaussian distribution with mean $\begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}$ and information matrix $\Omega = \begin{bmatrix}\Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22}\end{bmatrix}$. The pdf associated with conditional distribution is,$(X_1|X_2 = \boldsymbol{x}_2) \sim \mathcal{N}\left(\mu_1 - \Omega_{11}^{-1}\Omega_{12}(\boldsymbol{x}_2 - \mu_2), \Omega_{11}^{-1}\right)$

Edge merging

Let X_1, X_2 be random vectors with a joint Gaussian distribution. Assume that X_1 conditioned on $X_2 = x_2$ is distributed as $\mathcal{N}(\mu_1 - \Omega_{11}^{-1}\Omega_{12}(x_2 - \mu_2), \Omega_{11}^{-1})$ and that the marginal distribution of X_2 is $\mathcal{N}(\bar{\mu}_2, \bar{\Omega}_{22}^{-1})$. Then, X_1 and X_2 joint distribution is

$$\mathcal{N}\left(\begin{bmatrix}\mu_1 + \Omega_{11}^{-1}\Omega_{12}(\mu_2 - \bar{\mu}_2) \\ \bar{\mu}_2\end{bmatrix}, \begin{bmatrix}\Omega_{11} & \Omega_{12} \\ \Omega_{12}^\top & \bar{\Omega}_{22} + \Omega_{12}^\top\Omega_{11}^{-1}\Omega_{12}\end{bmatrix}^{-1}\right).$$

Gaussian marginal algorithm

Lemma (Geometric averaging)

Let $\Omega_w = \sum_{i=1}^{n} A_i \Omega_i$. The weighted geometric product of Gaussian density functions $\phi(\boldsymbol{x}|\mu_i, \Omega_i^{-1}), \forall i \in \{1, \dots, n\}$ with corresponding weights A_i is given as,

$$\prod_{i=1}^{n} \phi(\boldsymbol{x}|\mu_i, \Omega_i^{-1})^{A_i} = \phi\left(\boldsymbol{x}\middle|\Omega_w^{-1} \sum_{i=1}^{n} A_i \Omega_i \mu_i, \Omega_w^{-1}\right).$$

Gaussian marginal algorithm

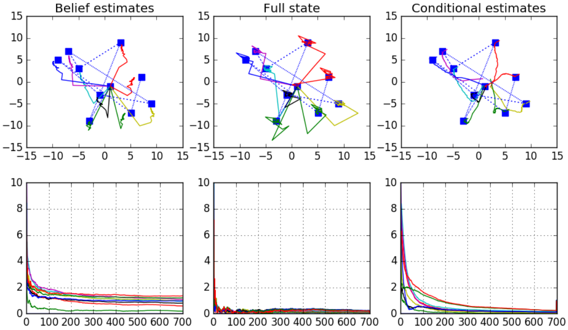
Lemma (Likelihood update)

Let the likelihood density be described as $\mathrm{q}_i(\mathbf{z}_{i,t}|\mathcal{X}_i) = \phi\left(\mathbf{z}_{i,t}|H_i\mathcal{X}_i, V_i\right)$*. Then the posterior Gaussian density obtained as likelihood prior product* $\phi\left(\mathbf{z}_{i,t}|H_i\mathcal{X}_i, V_i^{-1}\right)\phi\left(\mathcal{X}_i;\mu,\Omega_i^{-1}\right)$ *is*

$$\mathcal{N}\left(\left(H_i^\top V_i H_i + \Omega_i\right)^{-1}\left(H_i^\top V_i \mathbf{z}_{i,t} + \Omega_i\mu_i\right),\ \left(H_i^\top V_i H_i + \Omega_i\right)^{-1}\right)$$

Self-state estimates

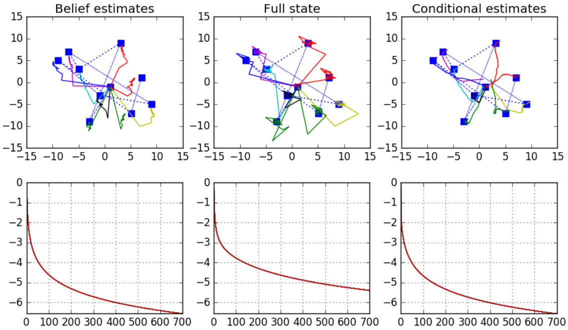
Belief propagation, full state and partial state estimates for a 10-agent ring network



(Row 1) Convergence to true positions
(Row 2) Estimation error across time

Self-state estimates

Belief propagation, full state and partial state estimates for a 10-agent ring network



(Row 1) Convergence to true positions
(Row 2) Logarithm of maximum eigenvalue of the self-covariance estimates

Decentralized update

Consensus: Geometric mixing with stochastic weights

$$v_{i,t} = \frac{1}{Z^v_{i,t}} \prod_{j=1}^{n} p_{j,t}^{A_{ij}}, \quad Z^v_{i,t} = \int_{\boldsymbol{x}\in\mathcal{X}} \left(\prod_{j=1}^{n} p_{j,t}^{A_{ij}} \right) \qquad \text{(Mixing step)}$$

Likelihood update: SMD algorithm

$$\begin{aligned} p_{i,t+1} &= \arg\min_{p\in\mathcal{F}_m} \left\{ \alpha_t \left\langle \frac{\delta F}{\delta p}(p_{i,t}, \boldsymbol{z}_{i,t}), p \right\rangle + \mathrm{D_{KL}}(p \| v_{i,t}) \right\} \\ &= \exp\left(\alpha_t \frac{\delta F}{\delta p} \right) v_{i,t} \Big/ \left(\int \exp(\alpha_t \frac{\delta F}{\delta p}) v_{i,t}(\boldsymbol{x}) d\boldsymbol{x} \right) \end{aligned}$$