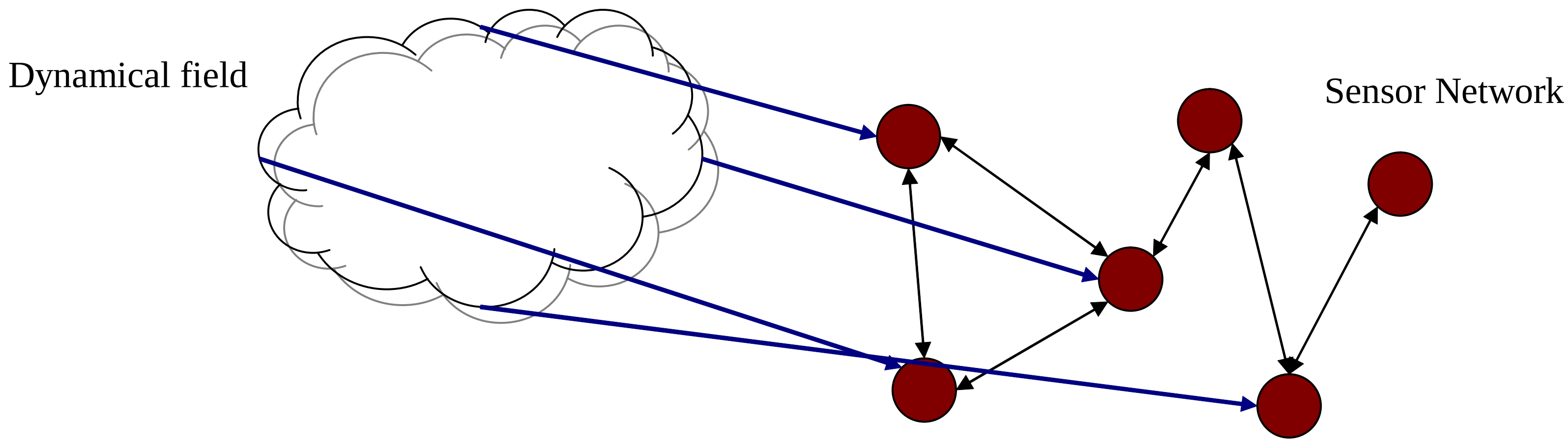
Convergence Results in Distributed Kalman Filtering

Shuguang (Robert) Cui

Texas A&M University

- Joint Work With: S. Kar (CMU), H. V. Poor (Princeton) and J. M. F. Moura (CMU)

Linear Dynamical Systems: Distributed Filtering



Field/Signal Process: $\{\mathbf{x}_k\} \in \mathbb{R}^M$

$$\mathbf{x}_{(k+1)\Delta} = \mathcal{F}\mathbf{x}_{k\Delta} + \mathbf{w}_k$$

Local observation at sensor n: $\{\mathbf{y}^n_{k\Delta}\} \in \mathbb{R}^{m_n}$

$$\mathbf{y}^n_{k\Delta} = \mathcal{C}_n\mathbf{x}_{k\Delta} + \mathbf{v}^n_k$$

- Applications: Smart Grid, Social Network, etc...

Existing Work

- Consensus or Distributed Optimization Based Approaches ([Olfati-Saber,2007], [Anandkumar *et al*, 2008] [Giannakis *et al*, 2008], [Ribiero *et al*, 2008], [Khan *et al*, 2008], [Scaglione *et al*, 2008], [Nedich *et al*, 2008], [Moura *et al*, 2008])
  - Based on linear estimate fusion among neighboring sensors
  - Error stability is not guaranteed (with only global but not local detectability) in general due to time-varying correlations among neighboring estimates
  - Asymptotic stability may only be obtained under specific margins of instability of system dynamics ([Khan-Kar-Jadbabaie-Moura, 2009])

Existing Work (*Contd.*)

- An estimate swapping approach ([Kar-Moura, 2010])
  - Neighborhood estimates are swapped over time guaranteeing asymptotic stability only under global detectability
  - Uses the machinery of Random Dynamical Systems to establish distributional convergence of related switched Riccati iterates
  - Does not exploit the (possibly) variable rate communication or any form of observation fusion to improve performance
  - Does not quantify how to approach the centralized estimation performance

Distributed Filtering Architecture

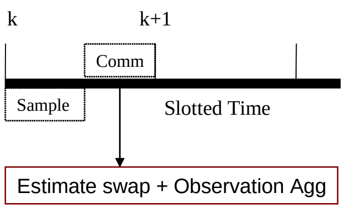
- The agents (sensors) need to collaborate for *successful* estimation of $\{\mathbf{x}_k\}$.

- Basic Sensing and Communication Architecture:
  - Inter-agent communication rate is measured as the link activation rate
  - Inter-agent communication is random

- Communication is constrained by the network topology; and only $\overline{\gamma} > 0$ rounds of inter-agent message passing may be allowed per epoch $(k\Delta, (k+1)\Delta]$ on an average.

Basic M-GIKF Tasks

The overall communication rate $\overline{\gamma}$ is split into:



- **Estimate Swapping**:
  - Agents perform pairwise swapping of previous epoch's estimates (if the corresponding link is active)
  - *Estimate Swapping guarantees asymptotic stability of local estimation errors under weak global detectability assumptions*
- **Observation Aggregation**:
  - Agents use the remaining communication to disseminate their instantaneous observations to the neighbors
  - *Observation Aggregation improves estimation performance*

Basic M-GIKF Tasks (*Contd.*)

Define:

- $M^e(k)$: Number of sensor communications at the k-th epoch for estimate swapping
- $M^o(k)$: Number of sensor communications at the k-th epoch for observation aggregation

Communication Constraint: Since $\overline{\gamma} > 0$ is given, we require:

$$\limsup_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} (M^e(k) + M^o(k)) \leq \overline{\gamma} \text{ a.s.}$$

Estimate Swapping: GS Protocol

Estimate Swapping at the k-th epoch is performed as follows:

- *Generate an adjacency matrix* $A^e(k) \in \mathcal{M}$ *from distribution* $\mathcal{D}$ *independently of the past*; within each $A^e(k)$ one active link at most
- *If* (n, l) *is a link in* $A^e_{nl}(k)$, *i.e.,* $A^e_{nl}(k) = 1$, *the agents* n *and* l *exchange (swap) their previous estimates*

Remark 1.

- At most one network link is active at a time.
- The case $A^e(k) = I_N$ corresponds to no swapping at all

Estimate Swapping: GS Protocol (*Contd.*)

Connectivity - **Assumption (C.1)**: The maximal graph $(V, \mathcal{E})$ is connected, where V denotes the set of N agents and $\mathcal{E}$ the set of allowable links.

Proposition 1. *Under (C.1) the following hold:*

- *The sequence $\{A^e(k)\}$ of estimate swapping adjacency matrices is independent and identically distributed as $\mathcal{D}$*
- *The mean adjacency matrix $\overline{A}^e$ is irreducible and aperiodic*
- *The average number of inter-sensor communications due to estimate swapping satisfies almost surely (a.s.)*

$$\overline{M}^e = \lim_{k \to \infty} (1/k) \sum_{i=0}^{k-1} M^e(k) < \overline{\gamma}/2$$

Observation Aggregation

Observation Aggregation is performed as follows after Estimate Swapping:

- Each sensor n, starting with its current observation $\mathbf{y}_k^n$, keeps on exchanging its observation with its neighbor(s) till the end of the epoch.
- Inter-agent communications occur at successive ticks of a Poisson process of rate $\overline{\gamma}/2\Delta$ and at each tick only one of the network links is activated with uniform probability.
- The number of communications $M^o(k)$ (for observation aggregation) during $(k\Delta, (k+1)\Delta]$ then follows a Poisson distribution with mean $\overline{\gamma}/2$ and the corresponding sequence of adjacency matrices $\{A_k^o(i)\}$, $i = 1, \cdots, M^o(k)$, is i.i.d. distributed uniformly on the set $\mathcal{M} \setminus \{I_N\}$ of non-trivial permutation matrices.

Observation Aggregation: GS Protocol (*Contd.*)

Proposition 2. *Under (C.1) the following hold:*

- *The average number of inter-agent communications due to observation aggregation per epoch is* $\overline{\gamma}/2$.
- *The sequence* $\{\mathcal{I}_k^n\}$ *denotes the set of observations (w.r.t. node indices) available at sensor* n *at the end of the epochs*
- *For every epoch* k *and every sensor* n,

$$\mathbb{P}(\mathcal{I}_k^n = [1, \cdots, N]) > 0$$

M-GIKF: Reasonable Communication Scheme

Definition 1. *A communication scheme (for estimate swapping and observation aggregation) is said to be* reasonable *if*

(E.1) *The estimate swapping adjacency matrices* $\{A^e(k)\}$ *are i.i.d. The mean matrix* $\overline{A}$ *is doubly stochastic, irreducible, and aperiodic.*

(E.2) *The sequences* $\{\mathcal{I}_k^n\}$ *are i.i.d. for each* n*, independent of the estimate swapping and satisfy* $\mathbb{P}(\mathcal{I}_k^n = [1, \cdots, N]) > 0$*.*

(E.3) *The average number of inter-sensor communications (including both the estimate swapping and observation aggregation steps) per epoch is less than or equal to* $\overline{\gamma}$*, where* $\overline{\gamma} > 0$ *is a predefined upper bound on the communication rate.*

Remark 2. *The previous GS protocol is a reasonable scheme under* ***(C.1)****.*

M-GIKF: Estimate Update

Define:

- $(\widehat{\mathbf{x}}^n_{k|k-1}, \widehat{P}^n_k)$: Estimate (state) at sensor n of $\mathbf{x}_k$ based on *information* till time $k-1$
- $n_{\overrightarrow{k}}$: Neighbor of sensor n at time k w.r.t. $A^e(k)$.

With *estimate swapping*: swap $\left(\widehat{\mathbf{x}}^n_{k|k-1}, \widehat{P}^n_k\right)$ and $\left(\widehat{\mathbf{x}}^{n_{\overrightarrow{k}}}_{k|k-1}, \widehat{P}^{n_{\overrightarrow{k}}}_k\right)$;

With *observation aggregation*: sensor n collects: $\mathbf{y}_k^{\mathcal{I}_n^k}$

Update Rule:

$$\widehat{\mathbf{x}}^n_{k+1|k} = \mathbb{E}\left[\mathbf{x}_{k+1} \;\middle|\; \widehat{\mathbf{x}}^{n_{\overrightarrow{k}}}_{k|k-1}, \widehat{P}^{n_{\overrightarrow{k}}}_k, \mathbf{y}_k^{\mathcal{I}_n^k}\right]$$

$$\widehat{P}^n_{k+1} = \mathbb{E}\left[\left(\mathbf{x}_{k+1} - \widehat{\mathbf{x}}^n_{k+1|k}\right)\left(\mathbf{x}_{k+1} - \widehat{\mathbf{x}}^n_{k+1|k}\right)^T \;\middle|\; \widehat{\mathbf{x}}^{n_{\overrightarrow{k}}}_{k|k-1}, \widehat{P}^{n_{\overrightarrow{k}}}_k, \mathbf{y}_k^{\mathcal{I}_n^k}\right]$$

M-GIKF: Estimate Update (*Contd.*)

- The filtering steps can be implemented through time-varying Kalman filter recursions
- The sequence $\{\widehat{P}_k^n\}$ of conditional predicted error covariance matrices at sensor n satisfies the random Riccati recursion:

$$\widehat{P}_{k+1}^n = \mathcal{F}\widehat{P}_k^{n_{\vec{k}}}\mathcal{F}^T + \mathcal{Q} - \mathcal{F}\widehat{P}_k^{n_{\vec{k}}}\mathcal{C}_n^T\left(\mathcal{C}_n\widehat{P}_k^{n_{\vec{k}}}\mathcal{C}_n^T + \mathcal{R}_n\right)^{-1}\mathcal{C}_n\widehat{P}_k^{n_{\vec{k}}}\mathcal{F}^T$$

- The sequence $\{\widehat{P}_k^n\}$ is random due to the random neighborhood selection functions $n_{\vec{k}}$ and $\mathcal{I}_n^k$
- The goal is to study asymptotic properties of $\{\widehat{P}_k^n\}$ at every sensor n
  - In what sense $\{\widehat{P}_k^n\}$ is stable
  - In what sense they reach *agreement*

Switched Riccati Iterates

Let $\mathfrak{P}$ denote a generic subset of $[1, \cdots, N]$.
Define:

- $\mathcal{C}_{\mathfrak{P}}$: The stack of $\mathcal{C}_j$'s for all $j \in \mathfrak{P}$
- $f_{\mathfrak{P}}(\cdot)$: The Riccati operator given by

$$f_{\mathfrak{P}}(X) = \mathcal{F}X\mathcal{F}^T + \mathcal{Q} - \mathcal{F}X\mathcal{C}_{\mathfrak{P}}^T \left(\mathcal{C}_{\mathfrak{P}}X\mathcal{C}_{\mathfrak{P}}^T + \mathcal{R}_n\right)^{-1} \mathcal{C}_{\mathfrak{P}}X\mathcal{F}^T$$

The conditional prediction error covariances are then updated as:

$$\widehat{P}_n(k+1) = f_{\mathcal{I}_k^n}\left(\widehat{P}_{n_k^{\rightarrow}}(k)\right), \quad \forall n$$

Remark 3. *The sequence* $\{\widehat{P}_n(k)\}$ *evolves as a random Riccati iterate with non-stationary switching.*

Assumptions on the Signal Model

We impose the following global assumptions on the signal/observation model:

- **Stabilizability-Assumption (S.1)**: The pair $(\mathcal{F}, \mathcal{Q}^{1/2})$ is stabilizable. The non-degeneracy (positive definiteness) of $\mathcal{Q}$ ensures this.
- **Global Detectability-Assumption (D.1)**: The pair $(\mathcal{C}, \mathcal{F})$ is detectable, where $\mathcal{C} = [\mathcal{C}_1^T \cdots \mathcal{C}_N^T]^T$.

Remark 4. *We do not assume local detectability at each agent. Assumption **(D.1)** is required even by the centralized estimator to achieve a stable estimation error. We show that under rather weak conditions on the inter-sensor communication, the global detectability assumption is also sufficient for our distributed scheme (the M-GIKF) to achieve stable estimation errors at each sensor.*

M-GIKF: Main Results on Convergence

Theorem 1. *Consider the M-GIKF under assumptions* ***(S.1)***, ***(D.1)***, ***(C.1)***-***(C.3)***. *Then,*

(1) For each n, the sequence of conditional error covariances $\left\{\widehat{P}_n(k)\right\}$ is stochastically bounded,

$$\lim_{J\to\infty} \sup_{k\in\mathbb{T}_+} \mathbb{P}(\|\widehat{P}_n(k)\| \geq J) = 0. \tag{1}$$

(2) Let q be a uniformly distributed random variable on $[1, \cdots, N]$ independent of the sequences $\{A^e(k)\}$ and $\{\mathcal{I}_n^k\}$. Then, the sequence $\left\{\widehat{P}_q(k)\right\}$ converges weakly to an invariant distribution $\mu^{\overline{\gamma}}$ on $\mathbb{S}_+^N$, i.e.,

$$\widehat{P}_q(k) \Longrightarrow \mu^{\overline{\gamma}}. \tag{2}$$

(3) The performance approaches the centralized one exponentially over $\bar{\gamma}$.

Proof Methodology

- An appropriate stationary modification of the switched Riccati iterate governing the covariance evolution leads to a hypothetical Random Dynamical System (RDS) in the sense of Arnold.

- The RDS thus constructed is shown to be order preserving and strongly sublinear.

- Together with the global detectability and network assumptions, the weak convergence of the stationary RDS is established.

- The ergodicity of the actual covariance sequence $\{\widehat{P}_n(k)\}$ is then applied to obtain the same convergence results with the above RDS.

Conclusions

- Stability of distributed filtering errors is established under connectivity of the network and global detectability.
- The analysis requires a new approach to the random Riccati equation (with non-stationary switching).
- Distributional convergence of the random Riccati equation is established;
- Approaching centralized performance exponentially fast over the inter-sensor communication rate.