Semistability of Randomly Switched Systems and Its Applications in Networked Systems

Jianghai Hu

Purdue University

July 8, 2013
Outline

• Semistability definition
• Networked systems examples
  • Consensus problems
  • Network localization
• Semistability characterization
• Convergence characterization
• Conclusion and future directions
A discrete-time switched linear system

\[ x(t + 1) = A_{\sigma(t)}x(t), \quad t = 0, 1, \ldots \]

- A finite number of modes \( \mathcal{M} = \{1, \ldots, m\} \)
- The switching sequence \( \sigma = \{\sigma(t)\}_{t=0,1,...} \) with \( \sigma(t) \in \mathcal{M} \)
- Solution starting from \( z \) under switching sequence \( \sigma \) is \( x(t; z, \sigma) \)

Different perspectives of switching sequence \( \sigma \)
- Control input: one can specify \( \sigma \) fully (e.g. switching stabilization)
- Perturbation: need to assume the worst (e.g. absolute stability)
- Random signal (mean square stability)
**Assumption:** Assume all $A_i$, $i \in \mathcal{M}$, have a common eigenvalue 1 with a common corresponding eigenvector $v \neq 0$

- Common eigenspace $\Omega_e := \bigcap_{i \in \mathcal{M}} \mathcal{N}(I - A_i) \neq \{0\}$
- A continuum of equilibrium points in $\Omega_e$

**Definition (Exponential Semistability)**

The SLS is exponentially semistable under arbitrary switching if starting from any initial state $z$ and under any switching sequence $\sigma$, there exist $x_e(z, \sigma) \in \Omega_e$ and constants $\rho > 0$, $0 \leq r < 1$ such that

$$\|x(t; z, \sigma) - x_e(z, \sigma)\| \leq \rho r^t \|z - x_e(z, \sigma)\|, \forall t$$
Suppose the switching sequence $\sigma(t)$ is random

- i.i.d. sequence with $P(\sigma(t) = i) = p_i$
- A Markov chain

**Definition (Mean Square Exponential Semistability)**

The SLS is mean square exponentially semistable if starting from any initial state $z$, there exists a random $x_e(z) \in \Omega_e$ such that

$$\mathbb{E} \left[ \| x(t; z) - x_e(z) \|^2 \right] \leq \rho r^t \| z - \mathbb{E}[x_e(z)] \|^2, \forall t$$
Consensus Problem

- A sensor network with a communication graph $G = (V, E)$
- Initial sensor data $x(0) = [x_1(0) \cdots x_n(0)]^T$
- A distributed algorithm to compute an average of sensor data: at each round each sensor updates its value to average of neighbors:

$$x(t+1) = \begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\
0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\
0 & 0 & 0 & \frac{1}{2} & \frac{1}{2}
\end{bmatrix} x(t)$$

- Laplacian matrix $L_G$ always has an eigenvalue 1 with eigenvector

$$1 = [1 \cdots 1]^T \in \mathcal{N}(I - L_G)$$
If the graph $G$ is connected, then $\dim \mathcal{N}(I - L_G) = 1$, with all other eigenvalues $\lambda$ satisfying $|\lambda| < 1$. The distributed algorithm will reach consensus $\bar{x} = w^T x(0)$ where $w$ is a left eigenvector of $L_G$ for $\lambda = 1$

$x(t) \rightarrow \bar{x} \cdot \mathbf{1}$ as $t \rightarrow \infty$

- Each sensor data $x_i(t) \rightarrow \bar{x} = \begin{bmatrix} 2/13 & 4/13 & 2/13 & 3/13 & 2/13 \end{bmatrix} x(0)$
Unconnected Graph $G$

If the graph $G$ is unconnected, then $L_G$ has multiple eigenvalues at 1:

$$\dim \mathcal{N}(I - L_G) = \#\{\text{connected components of } G\} > 1$$

The algorithm will reach consensus within each connected component, but not a global one

- $x(t) \rightarrow [a \ a \ a \ b \ b]^T \in \mathcal{N}(I - L_G)$ as $t \rightarrow \infty$
Suppose the network switches between two topologies $G_1$ and $G_2$ according to some switching signal $\sigma(t) \in \{1, 2\}$, $t = 0, 1, \ldots$

$$x(t + 1) = L_{G_{\sigma(t)}} x(t), \quad t = 0, 1, \ldots$$

$$\dim \mathcal{N}(I - L_{G_1}) = 2$$

$$\dim \mathcal{N}(I - L_{G_2}) = 4$$

**Observation:** $\mathcal{N}(I - L_{G_1}) \cap \mathcal{N}(I - L_{G_2}) = \text{span}(1)$ is of dimension one
Suppose the graph switches among \( \{G_i, i \in I\} \)

**Proposition**

\[
\dim \bigcap_{i \in I} \mathcal{N}(I - L_{G_i}) = 1 \text{ if and only if the union graph } \bigcup_{i \in I} G_i \text{ is connected. In this case if the switching signal } \sigma \text{ is such that each } G_i, \ i \in I, \text{ appears infinitely many times, then consensus will be reached}
\]

- Consensus \( \bar{x} \) depends on both \( x(0) \) and switching signal \( \sigma \)
- What is \( \bar{x} \) and how fast will the algorithm converge to consensus?
Suppose $G_{\sigma(t)}$, $t = 0, 1, \ldots$, is a sequence of random graphs in $\{G_i, i \in I\}$

- i.i.d.; Markov chain; non-Markov switching policy, etc.

**Definition (Recurrent Graphs)**

For each realization of $G_{\sigma}$, its recurrent graphs are those graphs $G_i$ that appear infinitely often in the sequence $G_{\sigma}$

**Proposition**

*If the union graph of all recurrent graphs is connected with probability one, then* $x(t) \to \bar{x} \cdot 1$ *as* $t \to \infty$ *a.s.*

- The consensus $\bar{x}$ is a random variable whose distribution depends on $x(0)$ and the random switching policy $\sigma$
- How fast is consensus reached and how is $\bar{x}$ distributed?
Network Localization Problem

- Localization is essential in networks (of sensors, robots, vehicles)
- Distributed localization using relative measurements
  - **Distance-based**: relative distances between neighbors
  - **Direction-based**: relative angles, i.e. Angle of Arrival (AOA)

(NCSU WILAN Lab) (DARPA)
Distance-Based vs. Direction-Based Localization

**Distance-Based Localization**
- Easier implementation (RSS [Savarese'01], TDOA [Savvides'01])
- Nonlinear equations, ambiguity
- [Eren’04], [Aspnes’06], [Priyantha’03]

**Direction-Based Localization**
- Easier to solve (linear equations)
- Higher equipment cost
- Sensitivity to measurement errors
- [Eren’06], [Rong’06], [Ash’07]
A formation graph $G = (V, p, K)$ is given by

- Set of vertices: $V = \{1, \ldots, n\}$
- Vertices' positions $p = [p_1 \cdots p_n]^T$
- Connectivity matrix $K = [k_{ij}]$ with $k_{ij} \geq 0$. Two vertices are connected if $k_{ij} > 0$

**Problem (AOA Localization)**

Suppose each vertex can measure the orientations of its neighbors. Recover the shape of the formation graph

- Ambiguity: absolute position, scale
At least two anchors needed to remove ambiguity

An anchored formation graph is a formation graph with a subset $A$ of vertices identified as anchors

**Problem (AOA Localization with Anchors)**

Knowing the absolute positions of anchors and orientations of edges, recover the absolute positions of all vertices

AOA localizable

Not AOA localizable
Mechanical analogy: formation graph as a spring network

- After perturbing vertex positions by \( \Delta \mathbf{p} \), the resistance force
  \[
  \mathbf{f} = \mathbf{S} \cdot \Delta \mathbf{p} + o(\|\Delta \mathbf{p}\|)
  \]

- \( \mathbf{S} \in \mathbb{R}^{2n \times 2n} \) is the **stiffness matrix**
  - Elastic energy stored is \( J \sim \frac{1}{2} \Delta \mathbf{p}^T \mathbf{S} \Delta \mathbf{p} \)
  - Null space of \( \mathbf{S} \) at least 3-dimensional: translations, rotations
Stiffness Matrix Example

Stiffness matrix $S$ is structurally similar to Laplacian matrix $L_G$

![Graph showing the structure of the stiffness matrix](image)

- **Observation:** $S$ is completely determined by AOA information
  - The $(i, j)$-th 2-by-2 block is the projection matrix onto $p_i - p_j$
  - Each row blocks add up to zero

- For anchored formation, partition $S = \begin{bmatrix} S_{ff} & S_{fa} \\ S_{af} & S_{aa} \end{bmatrix}$
AOA Localizability

**Theorem ([Zhu’13])**

For an anchored formation, its AOA localization problem has a unique solution if and only if $S_{ff}$ is nonsingular.

- If $S_{ff}$ is nonsingular, the anchored formation is called **fixable**.
AOA Localization Algorithm

Rotations are in null space of stiffness matrix $S$

$$S \begin{bmatrix} p_1^{\perp} \\ \vdots \\ p_n^{\perp} \end{bmatrix} = \begin{bmatrix} S_{ff} & S_{fa} \\ S_{af} & S_{aa} \end{bmatrix} \begin{bmatrix} p_f^{\perp} \\ p_a^{\perp} \end{bmatrix} = 0$$

where $p_i^{\perp}$ is $p_i$ rotated $90^\circ$ counterclockwise

If the formation is fixable, then

$$p_f^{\perp} = -S_{ff}^{-1}S_{fa}p_a^{\perp}$$

- Due to structural resemblance to $L_G$, a consensus-type distributed algorithm can be designed to solve the equation iteratively

G. Zhu and J. Hu, Distributed network localization using angle-of-arrival information, ACC'2013
Anchor-less AOA Localization

**Simpler Problem:** With AOA information only and no anchors, recover the formation shape (but not its size and absolute position).

Decompose the stiffness matrix $S$ as $S = D + F$

- Diagonal part $D = \text{diag}(S_{ii})$
- Off diagonal part $F$

From $Sp^\perp = 0$, we have $p^\perp = -D^{-1}F p^\perp$

- $A_G$ has eigenvalue 1, with other eigenvalues satisfying $|\lambda| < 1$
- $\dim \mathcal{N}(I - A_G) = \mathcal{N}(S) \geq 3$, representing ambiguity:
  - Translations and rotations of $p^\perp$
  - Translations and scalings of $p$

**Iterative AOA localization algorithm:**

$$p^\perp(t + 1) = A_G \cdot p^\perp(t), \quad t = 0, 1, \ldots$$
Rigid Graph

**Definition (Rigid Graph)**

A formation graph is called rigid if \( \dim \mathcal{N}(S) = 3 \)

**Proposition**

For a rigid formation graph, starting from any initial \( p^\perp(0) \), the iteration

\[
p^\perp(t + 1) = A_G \cdot p^\perp(t), \quad t = 0, 1, \ldots
\]

will converge to some \( p(\infty) \) that differs from the true vertex positions \( p \) by a translation and a scaling.
Simulation

initial guess

5 iterations

10 iterations

50 iterations

100 iterations

ground truth
For nonrigid graphs, $\dim \mathcal{N}(I - A_G) \geq 4$

$$p^\perp(t + 1) = A_G \cdot p^\perp(t)$$

may converge to a wrong formation shape

- In general, $\dim \mathcal{N}(I - A_G) - 3$ is the number of linearly independent directions of infinitesimal perturbations of vertex positions that preserve edge lengths
Switching Formation Graphs

Suppose the formation graph switches among \( \{G_i, i \in \mathcal{I}\} \)

**Proposition**

\[
\dim \cap_{i \in \mathcal{I}} \mathcal{N}(I - A_G) = 3 \text{ if and only if the union graph } \bigcup_{i \in \mathcal{I}} G_i \text{ is rigid. In this case if the switching signal } \sigma \text{ is such that each } G_i \text{ appears infinitely often, then the localization algorithm }
\]

\[
p^\perp(t + 1) = A_{G_{\sigma(t)}} p^\perp(t)
\]

will converge to the correct formation shape
Suppose the formation graph $G_\sigma$ switches randomly among $\{G_i, i \in \mathcal{I}\}$.

**Proposition ([Zhu’13])**

*If with probability one the union of the recurrent formation graphs for the random switching signal $\sigma$ is rigid, then the localization algorithm

$$p^\perp(t + 1) = A_{G_{\sigma(t)}} p^\perp(t)$$

will converge to the correct formation shape a.s.*

- What is the mean square convergence rate?
- What is the equilibrium distribution?
Characterizing Semistability

**Assumption:** Assume $\Omega_e := \cap_{i \in \mathcal{M}} \mathcal{N}(I - A_i) \neq \{0\}$ for the SLS

$$x(t + 1) = A_{\sigma(t)}x(t), \quad t = 0, 1, \ldots$$

**Change of coordinates:** Let $\hat{x} := Ox$ where $O = \begin{bmatrix} O_e^\perp & O_e \end{bmatrix} \in \mathbb{R}^{n \times n}$

- $O_e^\perp$ has orthonormal columns spanning $\Omega_e^\perp$
- $O_e$ has orthonormal columns spanning $\Omega_e$

In the new coordinates, the SLS becomes $\hat{x}(t + 1) = \hat{A}_{\sigma}\hat{x}(t)$, with

$$\hat{A}_i = \begin{bmatrix} \hat{A}_{i,11} & 0 \\ \hat{A}_{i,21} & I \end{bmatrix}, \quad i \in \mathcal{M}$$

**Theorem ([Shen’CDC01])**

The SLS is exp. semistable under arbitrary switching if and only if the SLS $\{\hat{A}_i\}_{i \in \mathcal{M}}$ is exp. semistable under arbitrary switching. Moreover, the convergence rate is given by the joint spectrum radius of $\{\hat{A}_i\}_{i \in \mathcal{M}}$.
Joint Spectral Radius

Definition (Joint Spectral Radius)

The joint spectral radius of a set of square matrices \( \mathcal{A} = \{ A_i \}_{i \in \mathcal{M}} \) is

\[
\rho_\mathcal{A} := \lim_{t \to \infty} \sup_{\sigma} \left\{ \| A_{\sigma(t-1)} \cdots A_{\sigma(0)} \|^{1/t} \right\}
\]

- \( \rho_\mathcal{A} \) is the maximal trajectory growth rate of SLS with subsystems \( \mathcal{A} \)
- SLS is exponentially stable if and only if \( \rho_\mathcal{A} < 1 \)
- Computing \( \rho_\mathcal{A} \) is NP-hard [Tsitsiklis&Blondel’97]
- Many approximation algorithms [Theys’05, Junger’09, Hu’11]

Example:

\[
A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \rho_\mathcal{A} = \frac{1 + \sqrt{5}}{2} = 1.618
\]
Characterizing M.S. Semistability

Proposition ([Shen’13])

The random SLS $x(t+1) = A_{\sigma(t)}x(t)$ is m.s. semistable if and only if the following SLS is output m.s. exponentially stable

$$x(t+1) = A_{\sigma(t)}x(t), \quad y(t) = Cx(t)$$

where $C = O_{e\perp}$ is the projection matrix onto subspace $\Omega_e^{\perp}$

- If $\sigma(t)$ is i.i.d. with $\mathbb{P}(\sigma(t) = i) = p_i$, above is equivalent to

  $$\mathbb{E} \left[ \sum_{t=0}^{\infty} C^T A_{\sigma(t)} C \right] < \infty$$

- Difficult to characterize for general case
Convergence Point in $\Omega_e$

SLS in new coordinates: $\hat{x}(t + 1) = \hat{A}_\sigma \hat{x}(t)$, with

$$\hat{A}_i = \begin{bmatrix} \hat{A}_{i,11} & 0 \\ \hat{A}_{i,21} & 1 \end{bmatrix} \quad \text{and} \quad \hat{x}(t) = \begin{bmatrix} \hat{x}_e^\perp(t) \\ \hat{x}_e(t) \end{bmatrix}$$

Suppose SLS is (m.s.) exponentially semistable. Then $\hat{x}(t) \to \begin{bmatrix} 0 \\ \hat{x}_e^\perp(\infty) \end{bmatrix}$

$$\hat{x}_e^\perp(\infty) = \hat{x}_e^\perp(0) + \sum_{t=0}^{\infty} A_{\sigma(t)} \hat{x}_e(t)$$

- The above converges as $\hat{x}_e(t) \to 0$
- $\hat{x}_e^\perp(\infty)$ depends on $\hat{x}_e(0)$ and $\sigma$
- For random switching policy $\sigma$, $\hat{x}_e^\perp(\infty)$ is a random variable
- Distribution of $\hat{x}_e^\perp(\infty)$ difficult to characterize (may be fractional)
Conclusions

- Propose the concepts of semistability of SLS
- Two networked systems examples where semistability is relevant
  - Consensus problem
  - AOA localization problem
- Conditions for characterizing semistability and convergence point

Future Directions:

- More efficient approaches for characterizing semistability
- General random switching policies
- More application examples

Main references:

- Semistability of switched linear systems with applications to distributed sensor networks: A generating function approach, J. Shen, J. Hu and Q. Hui, CDC'11
- A distributed continuous-time algorithm for network localization using Angle-Of-Arrival information, G. Zhu, Y. Kim and J. Hu, Automatica’13