Control of Networks

Algorithms, Fundamental Limitations, Impossibility Results

Alex Olshevsky

Department of Electrical and Computer Engineering
Boston University

Linear control theory

- The study of the linear differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) + w_2(t) \end{aligned}$$

  is a classic subject of control theory.

Linear control theory

- The study of the linear differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) + w_2(t) \end{aligned}$$

  is a classic subject of control theory.
- Here $x(t) \in \mathbb{R}^n$ is the state, $u(t)$ is the input, $y(t)$ is the observation, and $w_1(t), w_2(t)$ are noises.

Linear control theory

- The study of the linear differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) + w_2(t) \end{aligned}$$

  is a classic subject of control theory.
- Here $x(t) \in \mathbb{R}^n$ is the state, $u(t)$ is the input, $y(t)$ is the observation, and $w_1(t), w_2(t)$ are noises.
- Possible goals: tracking, stabilization, control, ...

Linear control theory

- The study of the linear differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) + w_2(t) \end{aligned}$$

  is a classic subject of control theory.
- Here $x(t) \in \mathbb{R}^n$ is the state, $u(t)$ is the input, $y(t)$ is the observation, and $w_1(t), w_2(t)$ are noises.
- Possible goals: tracking, stabilization, control, ...
- Many aspects are well-understood by now.

Linear control theory

- The study of the linear differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) + w_2(t) \end{aligned}$$

  is a classic subject of control theory.
- Here $x(t) \in \mathbb{R}^n$ is the state, $u(t)$ is the input, $y(t)$ is the observation, and $w_1(t), w_2(t)$ are noises.
- Possible goals: tracking, stabilization, control, ...
- Many aspects are well-understood by now.
- What is still extremely unclear: what if the matrices B and C are not given?

Linear control theory

- The study of the linear differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) + w_2(t) \end{aligned}$$

  is a classic subject of control theory.
- Here $x(t) \in \mathbb{R}^n$ is the state, $u(t)$ is the input, $y(t)$ is the observation, and $w_1(t), w_2(t)$ are noises.
- Possible goals: tracking, stabilization, control, ...
- Many aspects are well-understood by now.
- What is still extremely unclear: what if the matrices B and C are not given?
- This is the subject of this presentation. Designed to be self-contained (no knowledge of control necessary...)

Motivating example: PMU placement

- Goal: move closer to real-time observation of power grids.

Motivating example: PMU placement

- Goal: move closer to real-time observation of power grids.
- Most popular approach is based on installation of Phasor Measurement Units (PMUs) which can sample at high rates (~ 30 samples per second) and have access to accurate GPS for synchronization.

Motivating example: PMU placement

- Goal: move closer to real-time observation of power grids.
- Most popular approach is based on installation of Phasor Measurement Units (PMUs) which can sample at high rates (~ 30 samples per second) and have access to accurate GPS for synchronization.
- Installation cost of a single PMU ranges from $40,000 to $180,000.

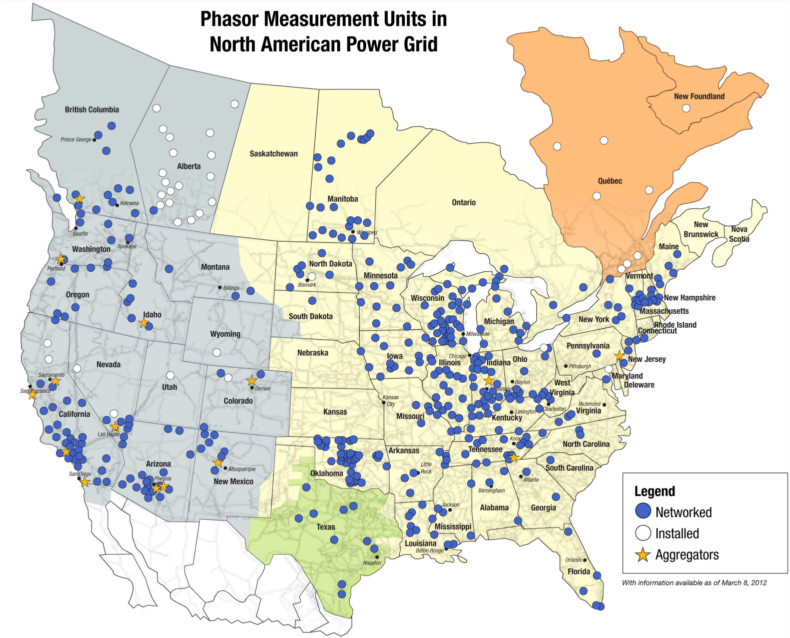
Motivating example: PMU placement

- Goal: move closer to real-time observation of power grids.
- Most popular approach is based on installation of Phasor Measurement Units (PMUs) which can sample at high rates (~ 30 samples per second) and have access to accurate GPS for synchronization.
- Installation cost of a single PMU ranges from \$40,000 to \$180,000.
- Roughly $\sim 1,500$ PMUs have been installed in the United States in the past 15 years, with a total cost on the order of $\sim \$100M$

Motivating example: PMU placement

- Goal: move closer to real-time observation of power grids.
- Most popular approach is based on installation of Phasor Measurement Units (PMUs) which can sample at high rates (~ 30 samples per second) and have access to accurate GPS for synchronization.
- Installation cost of a single PMU ranges from $40,000 to $180,000.
- Roughly $\sim 1,500$ PMUs have been installed in the United States in the past 15 years, with a total cost on the order of $\sim \$100M$
- This is part of the North American Synchronophasor Initiative. Goal is described as 100% coverage of important transmission lines.

PMU Placement as of 2015



Problem statement (noiseless case)

- We are given a system of differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, n.$$

Problem statement (noiseless case)

- We are given a system of differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, n.$$

- We have the ability to install actuators and sensors, meaning that we can transform the system into

$$\begin{aligned} \dot{x}_i &= \sum_j a_{ij} x_j + u_i, \quad i \in I \\ \dot{x}_i &= \sum_j a_{ij} x_j, \quad i \notin I \\ y_i &= x_i \quad i \in O \end{aligned}$$

Problem statement (noiseless case)

- We are given a system of differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, n.$$

- We have the ability to install actuators and sensors, meaning that we can transform the system into

$$\begin{aligned}
\dot{x}_i &= \sum_j a_{ij} x_j + u_i, \quad i \in I \\
\dot{x}_i &= \sum_j a_{ij} x_j, \quad i \notin I \\
y_i &= x_i \quad i \in O
\end{aligned}$$

- We want to choose the sets I and O as sparse as possible to achieve:

Problem statement (noiseless case)

- We are given a system of differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, n.$$

- We have the ability to install actuators and sensors, meaning that we can transform the system into

$$\begin{aligned} \dot{x}_i &= \sum_j a_{ij} x_j + u_i, \quad i \in I \\ \dot{x}_i &= \sum_j a_{ij} x_j, \quad i \notin I \\ y_i &= x_i \quad i \in O \end{aligned}$$

- We want to choose the sets I and O as sparse as possible to achieve:
    1. Controllability: can move the state from any $x(0)$ to any $x(T)$

Problem statement (noiseless case)

- We are given a system of differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, n.$$

- We have the ability to install actuators and sensors, meaning that we can transform the system into

$$\begin{aligned} \dot{x}_i &= \sum_j a_{ij} x_j + u_i, \quad i \in I \\ \dot{x}_i &= \sum_j a_{ij} x_j, \quad i \notin I \\ y_i &= x_i \quad i \in O \end{aligned}$$

- We want to choose the sets I and O as sparse as possible to achieve:
  1. Controllability: can move the state from any $x(0)$ to any $x(T)$
  2. Reachability: only care about moving the system in some directions.

Problem statement (noiseless case)

- We are given a system of differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, n.$$

- We have the ability to install actuators and sensors, meaning that we can transform the system into

$$\begin{aligned} \dot{x}_i &= \sum_j a_{ij} x_j + u_i, \quad i \in I \\ \dot{x}_i &= \sum_j a_{ij} x_j, \quad i \notin I \\ y_i &= x_i \quad i \in O \end{aligned}$$

- We want to choose the sets I and O as sparse as possible to achieve:
  1. Controllability: can move the state from any $x(0)$ to any $x(T)$
  2. Reachability: only care about moving the system in some directions.
  3. Energy constrained control: controllability with a bound on control energy (for example, to move from the origin to a random point on the unit sphere).

Energy considerations

- Given
$$\dot{x} = Ax + Bu,$$
let $E(x_i \to x_f, T)$ be the energy it takes to drive the system from x_i to x_f:
$$E(x_i \to x_f, T) = \inf\{\int_0^T ||u(t)||_2^2 \, dt \mid u \text{ drives the system from } x_i \text{ to } x_f\}$$

Energy considerations

- Given
$$\dot{x} = Ax + Bu,$$
let $E(x_i \to x_f, T)$ be the energy it takes to drive the system from x_i to x_f:
$$E(x_i \to x_f, T) = \inf\{\int_0^T ||u(t)||_2^2 \, dt \mid u \text{ drives the system from } x_i \text{ to } x_f\}$$
- In every real world scenario, use of arbitrarily large inputs in unphysical.

Energy considerations

- Given
$$\dot{x} = Ax + Bu,$$
let $E(x_i \to x_f, T)$ be the energy it takes to drive the system from x_i to x_f:
$$E(x_i \to x_f, T) = \inf\{\int_0^T ||u(t)||_2^2 \, dt \mid u \text{ drives the system from } x_i \text{ to } x_f\}$$
- In every real world scenario, use of arbitrarily large inputs in unphysical.
- Very easy to write down reasonable-looking real 10×10 systems where the energy is of the magnitude 10^{30} or more.

Energy considerations

- Given

$$\dot{x} = Ax + Bu,$$

  let $E(x_i \to x_f, T)$ be the energy it takes to drive the system from x_i to x_f:

$$E(x_i \to x_f, T) = \inf\{\int_0^T ||u(t)||_2^2 \, dt \mid u \text{ drives the system from } x_i \text{ to } x_f\}$$

- In every real world scenario, use of arbitrarily large inputs in unphysical.
- Very easy to write down reasonable-looking real 10×10 systems where the energy is of the magnitude 10^{30} or more.
- Want to measure "difficulty of controllability" through just one number. Standard choice:

$$\mathcal{E}(T) = \frac{1}{S_1} \int_{||z||_2=1} E(0 \to z, T) \, dz,$$

  where S_1 is the surface area of the unit sphere.

Time-varying actuator scheduling

- Another variation: allow the set of actuators and sensors to be time-varying.

Time-varying actuator scheduling

- Another variation: allow the set of actuators and sensors to be time-varying.
- Introduced in a paper published in *Automatica* in 1972:

ON THE DETERMINATION OF OPTIMAL COSTLY MEASUREMENT STRATEGIES
FOR LINEAR STOCHASTIC SYSTEMS*

by

Michael Athans
Massachusetts Institute of Technology
Cambridge, Mass. 02139, U.S.A.

ABSTRACT

This paper presents the formulation of a class of optimization problems dealing with selecting, at each instant of time, one measurement provided by one out of many sensors. Each measurement has an associated measurement cost. The basic problem is then to select an optimal measurement policy, during a specified observation time interval, so that a weighted combination of "prediction accuracy" and accumulated "observation cost" is optimized. The current analysis is limited to the class of linear stochastic dynamic systems and measurement subsystems. The problem of selecting the optimal measurement strategy can be transformed into a deterministic optimal control problem. An iterative digital computer algorithm is suggested for obtaining numerical results. It is shown that the optimal measurement policy and the associated "matched" Kalman-type filter can be precomputed, i.e., specified before the measurements actually occur.

Time-varying actuator scheduling

- Another variation: allow the set of actuators and sensors to be time-varying.
- Introduced in a paper published in *Automatica* in 1972:

ON THE DETERMINATION OF OPTIMAL COSTLY MEASUREMENT STRATEGIES FOR LINEAR STOCHASTIC SYSTEMS*

by

Michael Athans
Massachusetts Institute of Technology
Cambridge, Mass. 02139, U.S.A.

ABSTRACT

This paper presents the formulation of a class of optimization problems dealing with selecting, at each instant of time, one measurement provided by one out of many sensors. Each measurement has an associated measurement cost. The basic problem is then to select an optimal measurement policy, during a specified observation time interval, so that a weighted combination of "prediction accuracy" and accumulated "observation cost" is optimized. The current analysis is limited to the class of linear stochastic dynamic systems and measurement subsystems. The problem of selecting the optimal measurement strategy can be transformed into a deterministic optimal control problem. An iterative digital computer algorithm is suggested for obtaining numerical results. It is shown that the optimal measurement policy and the associated "matched" Kalman-type filter can be precomputed, i.e., specified before the measurements actually occur.

- Makes sense when the act of measurement itself is costly, or the transmission of measurement is costly.

Time-varying actuator scheduling

- Another variation: allow the set of actuators and sensors to be time-varying.
- Introduced in a paper published in *Automatica* in 1972:

ON THE DETERMINATION OF OPTIMAL COSTLY MEASUREMENT STRATEGIES FOR LINEAR STOCHASTIC SYSTEMS*

by

Michael Athans
Massachusetts Institute of Technology
Cambridge, Mass. 02139, U.S.A.

ABSTRACT

This paper presents the formulation of a class of optimization problems dealing with selecting, at each instant of time, one measurement provided by one out of many sensors. Each measurement has an associated measurement cost. The basic problem is then to select an optimal measurement policy, during a specified observation time interval, so that a weighted combination of "prediction accuracy" and accumulated "observation cost" is optimized. The current analysis is limited to the class of linear stochastic dynamic systems and measurement subsystems. The problem of selecting the optimal measurement strategy can be transformed into a deterministic optimal control problem. An iterative digital computer algorithm is suggested for obtaining numerical results. It is shown that the optimal measurement policy and the associated "matched" Kalman-type filter can be precomputed, i.e., specified before the measurements actually occur.

- Makes sense when the act of measurement itself is costly, or the transmission of measurement is costly.
- I will consider this in discrete time. Will dispense with formal problem statement.

A very partial literature review

- [Simon and Mitter, *Information and Computation*, 1968]. Considers minimizing the number of driver nodes.
- [Athans, *Automatica*, 1972]. An optimal control approach to time-varying actuator scheduling.
- Many works on structural controllability of networks in the 1980-early 1990s by Shields, Pearson, Glover, Willems, Siljak, Commault, Dion...
- Much work (recent and old) on sensor placement for observation of PDEs.
- [Liu, Slotine, Barabasi, *Nature*, 2011] explain to minimize the number of driver nodes generically for controllability.
- [Muller, Schuppert, *Nature*, 2011] argues most practical results involve affecting only a small number of key variables.
- [O., *IEEE Trans. on Control of Network Systems*, 2014] How to achieve controllability while minimizing the number of variables affected?
- Generalizations by O., Pappas, Jadbabaie, Bushnell, Poovendran, Lygeros, Cortes, Belabbas, Pasqualetti, Pequito, and their students to reachability and minimizing control energy.
- [Jadbabaie, O., Siami, *IEEE Trans. on Automatic Control* submission]

First main result

- **Theorem:** [O., *IEEE Trans. on Control of Network Systems*, 2014]
  The minimal controllability problem (i.e., choosing the sparsest set I of variables to affect to achieve controllability) is NP-hard.

First main result

- **Theorem:** [O., *IEEE Trans. on Control of Network Systems*, 2014] The minimal controllability problem (i.e., choosing the sparsest set I of variables to affect to achieve controllability) is NP-hard.
- In fact, the smallest number of variables of $\dot{x} = Ax$ that need to be affected for controllability – let's call this $I^*(A)$ – cannot be approximated to a multiplicative factor better than $O(\log n)$ in polynomial time.

First main result

- **Theorem:** [O., *IEEE Trans. on Control of Network Systems*, 2014] The minimal controllability problem (i.e., choosing the sparsest set I of variables to affect to achieve controllability) is NP-hard.
- In fact, the smallest number of variables of $\dot{x} = Ax$ that need to be affected for controllability – let's call this $I^*(A)$ – cannot be approximated to a multiplicative factor better than $O(\log n)$ in polynomial time.
- Arguably, this explains why no results until recently.

First main result

- **Theorem:** [O., *IEEE Trans. on Control of Network Systems*, 2014] The minimal controllability problem (i.e., choosing the sparsest set I of variables to affect to achieve controllability) is NP-hard.
- In fact, the smallest number of variables of $\dot{x} = Ax$ that need to be affected for controllability – let's call this $I^*(A)$ – cannot be approximated to a multiplicative factor better than $O(\log n)$ in polynomial time.
- Arguably, this explains why no results until recently.
- As observed in follow-up papers, as a consequence reachability and energy-efficient control are also NP-hard problems.

Second main result

- Is there an algorithm that matches this $O(\ln n)$ inapproximability barrier?

Second main result

- Is there an algorithm that matches this $O(\ln n)$ inapproximability barrier?
- Most experiments on real-world systems suggest that they are controllable from a constant number of variables.

Second main result

- Is there an algorithm that matches this $O(\ln n)$ inapproximability barrier?
- Most experiments on real-world systems suggest that they are controllable from a constant number of variables.
- **Theorem**: [O., *IEEE Transactions on Control of Network Systems*, 2014] There exists a polynomial time algorithm which, given the matrix A, outputs a set I so that $\dot{x} = Ax + B(I)u$ is controllable, and the number of entries in I is at most $I^*(A)(1 + \ln n)$.

Second main result

- Is there an algorithm that matches this $O(\ln n)$ inapproximability barrier?
- Most experiments on real-world systems suggest that they are controllable from a constant number of variables.
- **Theorem**: [O., *IEEE Transactions on Control of Network Systems*, 2014] There exists a polynomial time algorithm which, given the matrix A, outputs a set I so that $\dot{x} = Ax + B(I)u$ is controllable, and the number of entries in I is at most $I^*(A)(1 + \ln n)$.
- ...optimal up to constant factors.

Second main result

- Is there an algorithm that matches this $O(\ln n)$ inapproximability barrier?
- Most experiments on real-world systems suggest that they are controllable from a constant number of variables.
- **Theorem**: [O., *IEEE Transactions on Control of Network Systems*, 2014] There exists a polynomial time algorithm which, given the matrix A, outputs a set I so that $\dot{x} = Ax + B(I)u$ is controllable, and the number of entries in I is at most $I^*(A)(1 + \ln n)$.
- ...optimal up to constant factors.
- ...in many cases, this is good enough! For example, if the linear system is controllable from $O(1)$ entries, this finds $O(\ln n)$ entries.

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define

$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define

$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$

- The function ϕ is called supermodular if $X \subset Y$ implies $\Delta(X, a) \geq \Delta(Y, a)$.

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define

$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$

- The function ϕ is called supermodular if $X \subset Y$ implies $\Delta(X, a) \geq \Delta(Y, a)$.
- In other words, supermodularity is about diminishing returns.

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define

$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$

- The function ϕ is called supermodular if $X \subset Y$ implies $\Delta(X, a) \geq \Delta(Y, a)$.
- In other words, supermodularity is about diminishing returns.
- Can be thought of as a discrete version of concavity.

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define
$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$
- The function ϕ is called supermodular if $X \subset Y$ implies $\Delta(X, a) \geq \Delta(Y, a)$.
- In other words, supermodularity is about diminishing returns.
- Can be thought of as a discrete version of concavity.
- Key idea: the dimension of the reachable space (the set of $x(T)$ reachable from 0) is a supermodular function of I.

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define
$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$
- The function ϕ is called supermodular if $X \subset Y$ implies $\Delta(X, a) \geq \Delta(Y, a)$.
- In other words, supermodularity is about diminishing returns.
- Can be thought of as a discrete version of concavity.
- Key idea: the dimension of the reachable space (the set of $x(T)$ reachable from 0) is a supermodular function of I.
- Each variable actuated increases the reachable space, but variables have less effect when added later.

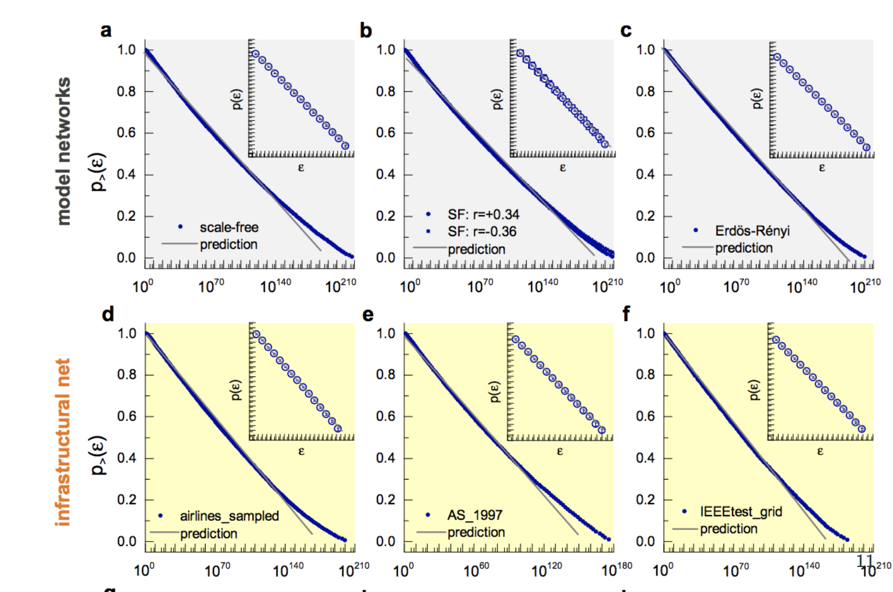
Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define
$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$
- The function ϕ is called supermodular if $X \subset Y$ implies $\Delta(X, a) \geq \Delta(Y, a)$.
- In other words, supermodularity is about diminishing returns.
- Can be thought of as a discrete version of concavity.
- Key idea: the dimension of the reachable space (the set of $x(T)$ reachable from 0) is a supermodular function of I.
- Each variable actuated increases the reachable space, but variables have less effect when added later.
- Algorithm is simple: keep adding variables to greedily maximize the dimension of the reachable space.

Proof idea: supermodularity

- Let ϕ be a function from subsets of $\{1, \ldots, p\}$ to $\mathbb{R}$.
- Suppose ϕ is increasing: if $X \subset Y$, then $\phi(X) \leq \phi(Y)$. For $a \notin X$, define
$$\Delta(X, a) = \phi(X \cup a) - \phi(X)$$
- The function ϕ is called supermodular if $X \subset Y$ implies $\Delta(X, a) \geq \Delta(Y, a)$.
- In other words, supermodularity is about diminishing returns.
- Can be thought of as a discrete version of concavity.
- Key idea: the dimension of the reachable space (the set of $x(T)$ reachable from 0) is a supermodular function of I.
- Each variable actuated increases the reachable space, but variables have less effect when added later.
- Algorithm is simple: keep adding variables to greedily maximize the dimension of the reachable space.
- The $O(\log n)$ approximability is a general result about greedy supermodular optimization.

Unfortunately, energy grows exponentially with these methods



What's going on?

- Optimizing for controllability turns out to be not quite the right thing to do.

What's going on?

- Optimizing for controllability turns out to be not quite the right thing to do.
- Need to explicitly consider energy in the problem formulation.

What's going on?

- Optimizing for controllability turns out to be not quite the right thing to do.
- Need to explicitly consider energy in the problem formulation.
- Maybe we can generalize the supermodularity-based approach discussed earlier?

What's going on?

- Optimizing for controllability turns out to be not quite the right thing to do.
- Need to explicitly consider energy in the problem formulation.
- Maybe we can generalize the supermodularity-based approach discussed earlier?
- Two recent papers
  T. Summers, F. Cortesi, J. Lygeros, "On submodularity and controllability in complex dynamical networks,"
  IEEE Transactions on Control of Network Systems, 2016
  V. Tzoumas, M. A. Rahimian, G. J. Pappas, A. Jadbabaie, "Minimal actuator placement with bounds on control effort,"
  IEEE Transactions on Control of Network Systems, 2016
  claimed that $-\mathcal{E}(T)$ is a supermodular function of the actuated variables.

What's going on?

- Optimizing for controllability turns out to be not quite the right thing to do.
- Need to explicitly consider energy in the problem formulation.
- Maybe we can generalize the supermodularity-based approach discussed earlier?
- Two recent papers
  T. Summers, F. Cortesi, J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Transactions on Control of Network Systems*, 2016
  V. Tzoumas, M. A. Rahimian, G. J. Pappas, A. Jadbabaie, "Minimal actuator placement with bounds on control effort," *IEEE Transactions on Control of Network Systems*, 2016
  claimed that $-\mathcal{E}(T)$ is a supermodular function of the actuated variables.
- If true, the same guarantees would effortlessly carry over. Unfortunately, I constructed a counterexample in [O., *IEEE Transactions on Control of Network Systems*, 2018].

Non-supermodularity of control energy

- Somewhat of a counterintuitive phenomenon. Counterexample:

$$A = \begin{pmatrix} -182 & 0 & -565 & 0 & -11 & -736 \\ 0 & -1075 & 831 & -276 & -1752 & -612 \\ -565 & 831 & -2435 & 214 & 1321 & -1853 \\ 0 & -276 & 214 & -73 & -453 & -158 \\ -11 & -1752 & 1321 & -453 & -2864 & -1045 \\ -736 & -612 & -1853 & -158 & -1045 & -3371 \end{pmatrix}$$

$$\begin{aligned} \mathcal{E}_{I=\{1,2,3\}}(\infty) - \mathcal{E}_{I=\{1,2,3,4\}}(\infty) &\approx 2.5 \cdot 10^4 \\ \mathcal{E}_{I=\{1,2,3,5\}}(\infty) - \mathcal{E}_{I=\{1,2,3,4,5\}}(\infty) &\approx 2.52 \cdot 10^4 \end{aligned}$$

Non-supermodularity of control energy

- Somewhat of a counterintuitive phenomenon. Counterexample:

$$A = \begin{pmatrix} -182 & 0 & -565 & 0 & -11 & -736 \\ 0 & -1075 & 831 & -276 & -1752 & -612 \\ -565 & 831 & -2435 & 214 & 1321 & -1853 \\ 0 & -276 & 214 & -73 & -453 & -158 \\ -11 & -1752 & 1321 & -453 & -2864 & -1045 \\ -736 & -612 & -1853 & -158 & -1045 & -3371 \end{pmatrix}$$

$$\begin{aligned} \mathcal{E}_{I=\{1,2,3\}}(\infty) - \mathcal{E}_{I=\{1,2,3,4\}}(\infty) &\approx 2.5 \cdot 10^4 \\ \mathcal{E}_{I=\{1,2,3,5\}}(\infty) - \mathcal{E}_{I=\{1,2,3,4,5\}}(\infty) &\approx 2.52 \cdot 10^4 \end{aligned}$$

- Found by working backwards from the error in paper, which is the assertion that $A \succeq B$ implies $A^2 \succeq B^2$.

Non-supermodularity of control energy

- Somewhat of a counterintuitive phenomenon. Counterexample:

$$A = \begin{pmatrix} -182 & 0 & -565 & 0 & -11 & -736 \\ 0 & -1075 & 831 & -276 & -1752 & -612 \\ -565 & 831 & -2435 & 214 & 1321 & -1853 \\ 0 & -276 & 214 & -73 & -453 & -158 \\ -11 & -1752 & 1321 & -453 & -2864 & -1045 \\ -736 & -612 & -1853 & -158 & -1045 & -3371 \end{pmatrix}$$

$$\begin{aligned} \mathcal{E}_{I=\{1,2,3\}}(\infty) - \mathcal{E}_{I=\{1,2,3,4\}}(\infty) &\approx 2.5 \cdot 10^4 \\ \mathcal{E}_{I=\{1,2,3,5\}}(\infty) - \mathcal{E}_{I=\{1,2,3,4,5\}}(\infty) &\approx 2.52 \cdot 10^4 \end{aligned}$$

- Found by working backwards from the error in paper, which is the assertion that $A \succeq B$ implies $A^2 \succeq B^2$.
- Approximation with energy constraints is currently an open problem.

Minimal reachability

- An alternative approach is to look at reachability. The graphs we showed above have lots of directions which are easily reachable.

Minimal reachability

- An alternative approach is to look at reachability. The graphs we showed above have lots of directions which are easily reachable.
- Can the supermodularity based approach be generalized to this setting?

Minimal reachability

- An alternative approach is to look at reachability. The graphs we showed above have lots of directions which are easily reachable.
- Can the supermodularity based approach be generalized to this setting?
- Suppose we just have one direction y_1 which we want to be reachable. A natural set function is

$$f(I) = ||P_{\text{reachable space}}(y_1)||_2^2$$

Minimal reachability

- An alternative approach is to look at reachability. The graphs we showed above have lots of directions which are easily reachable.
- Can the supermodularity based approach be generalized to this setting?
- Suppose we just have one direction y_1 which we want to be reachable. A natural set function is

$$f(I) = ||P_{\text{reachable space}}(y_1)||_2^2$$

- Two recent papers

  V. Tzoumas, A. Jadbabaie, G. J. Pappas, "Minimal actuator placement with bound on control effort," *IEEE Conference on Decision and Control*, 2015.

  Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, R. Poovendran, "Towards scalable voltage control in the smart grid," *Proc. of the 7th International CPS Conference*, 2016.

  claimed this is a supermodular function.

Minimal reachability

- An alternative approach is to look at reachability. The graphs we showed above have lots of directions which are easily reachable.
- Can the supermodularity based approach be generalized to this setting?
- Suppose we just have one direction y_1 which we want to be reachable. A natural set function is

$$f(I) = ||P_{\text{reachable space}}(y_1)||_2^2$$

- Two recent papers
  V. Tzoumas, A. Jadbabaie, G. J. Pappas, "Minimal actuator placement with bound on control effort," *IEEE Conference on Decision and Control*, 2015.
  Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, R. Poovendran, "Towards scalable voltage control in the smart grid," *Proc. of the 7th International CPS Conference*, 2016.
  claimed this is a supermodular function.
- Unfortunately, we show in a recent preprint [Jadbabaie, O., Pappas, Tzoumas, *IEEE Trans. on Automatic Control*, 2019] that this is false.

Minimal reachability

- An alternative approach is to look at reachability. The graphs we showed above have lots of directions which are easily reachable.
- Can the supermodularity based approach be generalized to this setting?
- Suppose we just have one direction y_1 which we want to be reachable. A natural set function is

$$f(I) = ||P_{\text{reachable space}}(y_1)||_2^2$$

- Two recent papers
  V. Tzoumas, A. Jadbabaie, G. J. Pappas, "Minimal actuator placement with bound on control effort," *IEEE Conference on Decision and Control*, 2015.
  Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, R. Poovendran, "Towards scalable voltage control in the smart grid," *Proc. of the 7th International CPS Conference*, 2016.
  claimed this is a supermodular function.
- Unfortunately, we show in a recent preprint [Jadbabaie, O., Pappas, Tzoumas, *IEEE Trans. on Automatic Control*, 2019] that this is false.
- In fact, we show a bit more....

Third main result

- Definition: $\mathrm{BPTIME}(t(n))$ is the class of problems for which a randomized algorithm can compute the correct answer with probability at least $2/3$ in time $t(n)$.

Third main result

- Definition: $\mathrm{BPTIME}(t(n))$ is the class of problems for which a randomized algorithm can compute the correct answer with probability at least $2/3$ in time $t(n)$.
- Observe: we can always approximate the minimum reachability problem to a multiplicative factor of $n = 2^{\log n}$ (just actuate every variable).

Third main result

- Definition: $\mathrm{BPTIME}(t(n))$ is the class of problems for which a randomized algorithm can compute the correct answer with probability at least $2/3$ in time $t(n)$.
- Observe: we can always approximate the minimum reachability problem to a multiplicative factor of $n = 2^{\log n}$ (just actuate every variable).
- **Theorem:** [Jadbabaie, O., Pappas, Tzoumas, IEEE TAC, 2019] For any $\delta \in (0, 1)$, unless problems in NP can be solved in $\mathrm{BPTIME}(n^{\log \log n})$ there is no polynomial time algorithm which approximates minimal reachability to a multiplicative factor of $2^{(\log n)^{\delta}}$.

Third main result

- Definition: $\mathrm{BPTIME}(t(n))$ is the class of problems for which a randomized algorithm can compute the correct answer with probability at least $2/3$ in time $t(n)$.
- Observe: we can always approximate the minimum reachability problem to a multiplicative factor of $n = 2^{\log n}$ (just actuate every variable).
- **Theorem:** [Jadbabaie, O., Pappas, Tzoumas, IEEE TAC, 2019] For any $\delta \in (0,1)$, unless problems in NP can be solved in $\mathrm{BPTIME}(n^{\log \log n})$ there is no polynomial time algorithm which approximates minimal reachability to a multiplicative factor of $2^{(\log n)^{\delta}}$.
- Punchline: minimal reachability is almost exponentially harder than minimal controllability, which was approximable to a factor of $O(\log n)$.

Summary and next directions

- Minimal controllability is NP-hard, but has a reasonable approximation guarantee.

Summary and next directions

- Minimal controllability is NP-hard, but has a reasonable approximation guarantee.
- Unfortunately, minimal controllability does not turn out to be a good objective.

Summary and next directions

- Minimal controllability is NP-hard, but has a reasonable approximation guarantee.
- Unfortunately, minimal controllability does not turn out to be a good objective.
- Incorporation of either control energy or desired directions leads to a loss of supermodularity with no crisp approximation results.

Summary and next directions

- Minimal controllability is NP-hard, but has a reasonable approximation guarantee.
- Unfortunately, minimal controllability does not turn out to be a good objective.
- Incorporation of either control energy or desired directions leads to a loss of supermodularity with no crisp approximation results.
- What if we consider the Athans problem, i.e., allow the actuators to change with time?

Summary and next directions

- Minimal controllability is NP-hard, but has a reasonable approximation guarantee.
- Unfortunately, minimal controllability does not turn out to be a good objective.
- Incorporation of either control energy or desired directions leads to a loss of supermodularity with no crisp approximation results.
- What if we consider the Athans problem, i.e., allow the actuators to change with time?
- Main idea: this problem might be a little easier.

Summary and next directions

- Minimal controllability is NP-hard, but has a reasonable approximation guarantee.
- Unfortunately, minimal controllability does not turn out to be a good objective.
- Incorporation of either control energy or desired directions leads to a loss of supermodularity with no crisp approximation results.
- What if we consider the Athans problem, i.e., allow the actuators to change with time?
- Main idea: this problem might be a little easier.
- In fact, lets limit ourselves to choosing an average of d actuators per step.

Random sampling

- Natural idea: choose the actuators randomly!

Random sampling

- Natural idea: choose the actuators randomly!
- Unfortunately, this doesn't work, in either static or dynamic case.

Random sampling

- Natural idea: choose the actuators randomly!
- Unfortunately, this doesn't work, in either static or dynamic case.
- Need some way to quantify which nodes are important. This approach almost seems like beginning the question.

Random sampling

- Natural idea: choose the actuators randomly!
- Unfortunately, this doesn't work, in either static or dynamic case.
- Need some way to quantify which nodes are important. This approach almost seems like beginning the question.
- Let's discuss a seemingly unrelated question: given a weighted graph, can you come up with a sparse subgraph which approximates it?

Random sampling

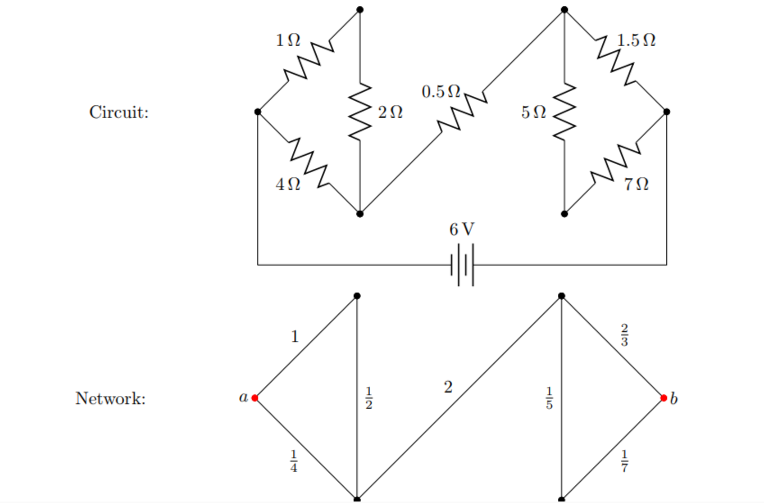
- Natural idea: choose the actuators randomly!
- Unfortunately, this doesn't work, in either static or dynamic case.
- Need some way to quantify which nodes are important. This approach almost seems like beginning the question.
- Let's discuss a seemingly unrelated question: given a weighted graph, can you come up with a sparse subgraph which approximates it?
- A graph can be encoded as a Laplacian matrix (defined by putting w_{ij}, the weight between the i'th and j'th node into the (i, j)'th entry of the matrix.

Random sampling

- Natural idea: choose the actuators randomly!
- Unfortunately, this doesn't work, in either static or dynamic case.
- Need some way to quantify which nodes are important. This approach almost seems like beginning the question.
- Let's discuss a seemingly unrelated question: given a weighted graph, can you come up with a sparse subgraph which approximates it?
- A graph can be encoded as a Laplacian matrix (defined by putting w_{ij}, the weight between the i'th and j'th node into the (i, j)'th entry of the matrix.
- Given a graph with a Laplacian L, we are asking for a sparse subgraph (say with $O(n)$ edges) with a Laplacian L_s such that $L \approx L_s$.

Resistance in a graph

A key observation is that you can define the resistance between any two nodes:



Fourth main result

- Key idea: to sparsify a graph, you can sample edges *proportional to effective resistance across them*.

Fourth main result

- Key idea: to sparsify a graph, you can sample edges *proportional to effective resistance across them*.
- In fact, [Spielman, Srivastava, SICOMP, 2011] showed that you can reduce a graph to $O(n/\epsilon^2)$ edges (where n is the number of nodes) without affecting eigenvectors/eigenvalues of the graph Laplacian by more than $O(\epsilon)$.

Fourth main result

- Key idea: to sparsify a graph, you can sample edges *proportional to effective resistance across them*.
- In fact, [Spielman, Srivastava, SICOMP, 2011] showed that you can reduce a graph to $O(n/\epsilon^2)$ edges (where n is the number of nodes) without affecting eigenvectors/eigenvalues of the graph Laplacian by more than $O(\epsilon)$.
- How can we use generalize this to our setting?

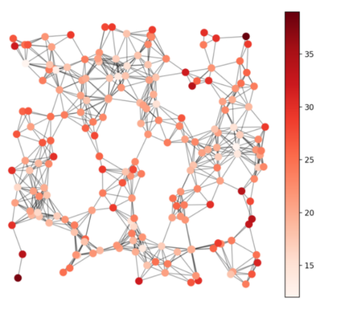
Fourth main result

- Key idea: to sparsify a graph, you can sample edges *proportional to effective resistance across them*.
- In fact, [Spielman, Srivastava, SICOMP, 2011] showed that you can reduce a graph to $O(n/\epsilon^2)$ edges (where n is the number of nodes) without affecting eigenvectors/eigenvalues of the graph Laplacian by more than $O(\epsilon)$.
- How can we use generalize this to our setting?
- Key idea: sample actuator i with probability proportional to $P_i^T W(T)^{-1} P_i$ where P_i is the i'th column of $W(T) = \int_0^T e^{tA} B B^T e^{tA^T}\, dt$. This is independent across time.

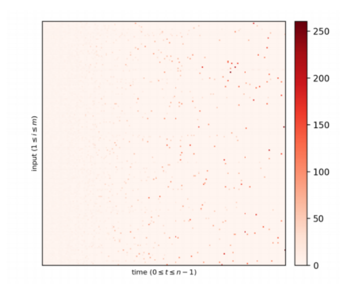
Fourth main result

- Key idea: to sparsify a graph, you can sample edges *proportional to effective resistance across them*.
- In fact, [Spielman, Srivastava, SICOMP, 2011] showed that you can reduce a graph to $O(n/\epsilon^2)$ edges (where n is the number of nodes) without affecting eigenvectors/eigenvalues of the graph Laplacian by more than $O(\epsilon)$.
- How can we use generalize this to our setting?
- Key idea: sample actuator i with probability proportional to $P_i^T W(T)^{-1} P_i$ where P_i is the i'th column of $W(T) = \int_0^T e^{tA} B B^T e^{tA^T} \, dt$. This is independent across time.
- **Theorem:** [Jadbabaie, O., Siami, *IEEE Trans. on Automatic Control* submission, 2019] If $T \geq n$ and d is at least a constant multiple of $\log n/\epsilon^2$ then with high probability the control energy of this scheme is at most $1 + \epsilon$ times the best possible (here, best possible means actuating every variable).

Performance on a 250 agent network

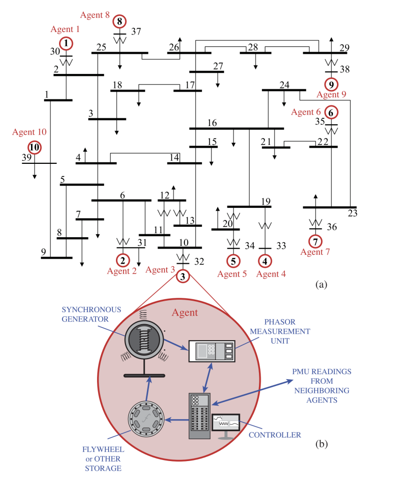


The system matrix A is the Laplacian of this undirected graph, with colors corresponding to weights on edges. The matrix is semistable.



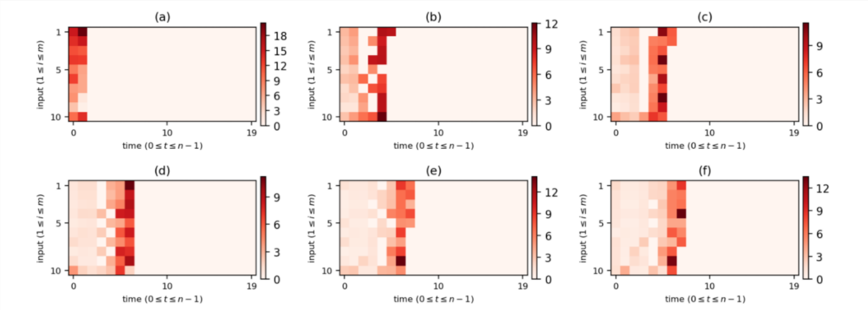
A sparse actuator schedule. A dot corresponds to a used actuator at a given time, and the color corresponds to size of the corresponding input.

The IEEE 39-bus system



This is a way to represent a complete graph with equations $m\ddot{\theta}_i + d_i\dot{\theta}_i = -\sum_j k_{ij}(\theta_i - \theta_j) + u_i$ coupling the nodes. This matrix is semistable.

Performance on the IEEE 39-bus system



Colors represent intensity of the actuator use. In contrast to the previous example, this schedule seems to be "front-loaded."

Conclusion

- Key takeaways:
  1. Exist optimal algorithm for minimal controllability.
  2. Supermodularity is a key property. It's lack makes things difficult.
  3. Minimal reachability is, surprisingly, close to being unsolvable.
  4. Effective control with time-varying actuators has been almost solved.
- Main challenge: find a class of systems for which incorporating energy constraints and desired reachable directions can be done.
- Satisfactory answers could have a transformative impact not only in electricity distribution systems and many other areas.