Sparsity, discreteness and optimal control

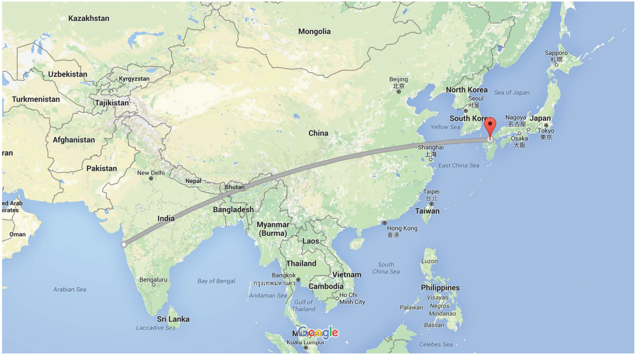
Masaaki Nagahara[1]

[1]The University of Kitakyushu
nagahara@ieee.org

20 May 2016
IIT Bombay

I am ...

- Born in Ehime, Japan
- Graduated from Kyoto University
- Working at The University of Kitakyushu as a full professor
- Interested in control, signal processing, communications
- A blogger of *Welcome to My Sparseland* (very sparse blogging)
  `sparseland.blogspot.com`



Map data ©2016 Google, SK planet, ZENRIN 500 km

The University of Kitakyushu

- I am working with The University of Kitakyushu
- Faculty of Environmental Engineering
- Control theory, signal/image processing, artificial intelligence, autonomous vehicles (including drones), and so on.
- We welcome foreign students for master and PhD degrees.
    - If you are interested, please email me.

About this talk

- Sparsity is useful (as you know)
- Relation between sparsity and discreteness
- Sparsity methods for control
- *Interplay between control and sparse signal processing*

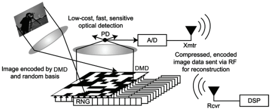
PART I

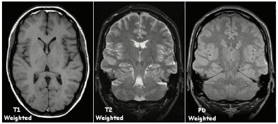
Sparsity and discreteness[1]

[1] M. Nagahara, Discrete signal reconstruction by sum of absolute values, *IEEE Signal Processing Letters*, Vol. 22, no. 10, pp. 1575-1579, Oct. 2015.

Sparsity in Engineering

- Image processing
  - single-pixel camera, compressed sensing MRI
- Statistics
  - big data analysis





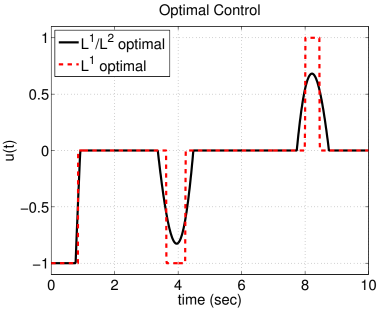
"Your recent Amazon purchases, Tweet score and location history makes you 23.5% welcome here."

Sparsity in Engineering

- Image processing
  - single-pixel camera, compressed sensing MRI
- Statistics
  - big data analysis
- *Discrete signal processing*
  - binary image reconstruction, digital communications

Sparsity in Engineering

- Image processing
  - single-pixel camera, compressed sensing MRI
- Statistics
  - big data analysis
- *Discrete signal processing*
  - binary image reconstruction, digital communications
- *Control*
  - networked control, sparse control, discrete-valued control



What is sparsity?

- A vector x in $\mathbb{R}^n$ is *sparse* if it contains many 0's, or has small ℓ^0 norm

$$\|x\|_0 \triangleq \text{the number of the nonzero elements in } x.$$

- Examples of sparse vectors
  - Frequency domain data of natural signals and images almost all of them are nearly 0 except for low-frequency data.
  - Pulse signals; they are sparse in the time domain.

Sparse signal reconstruction

- Suppose that a sparse signal $x \in \mathbb{R}^n$ is measured by linear measurements
$$y = \Phi x \in \mathbb{R}^m,$$
where $\Phi \in \mathbb{R}^{m \times n}$ is a known matrix.
- Finding the original x is ill-posed if $m < n$.
- To determine *one* vector from y, we adopt *optimization*.

Sparse optimization

- The following optimization will do for sparse signal reconstruction:

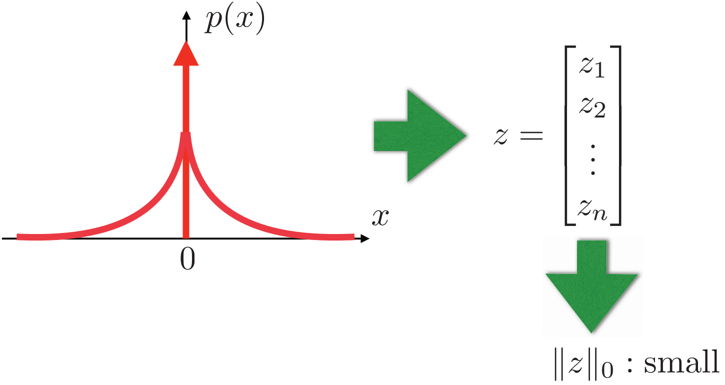
$$\min_{z\in\mathbb{R}^n} \|z\|_0 \text{ subject to } y = \Phi z.$$

- This gives the exact reconstruction (with assumptions on x and Φ).
- However, it is hard to solve if n is very large (e.g. 1 milion).
- In many cases, the following ℓ^1 optimization solves the problem:

$$\min_{z\in\mathbb{R}^n} \|z\|_1 \text{ subject to } y = \Phi z.$$

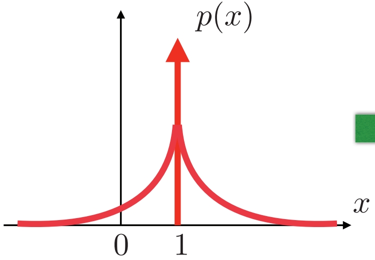
Sparse signals

- Probability distribution of sparse vectors
  - Dirac delta at $x = 0$ (discrete distribution)
  - continuous distribution for $x \neq 0$



Signals that contain many 1's

- Probability distribution of many-1 vectors
  - Dirac delta at $x = 1$ (discrete distribution)
  - continuous distribution for $x \neq 1$

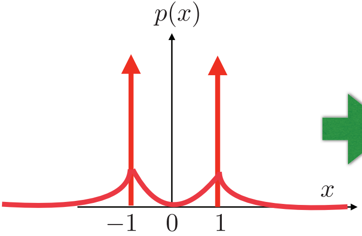


$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

$$\|z - 1\|_0 : \text{small}$$

Signals that contain many binary values ± 1

- Probability distribution
  - Dirac deltas at $x = \pm 1$ (discrete distribution)
  - continuous distribution for $x \neq \pm 1$
- If $\mathbb{P}[x = 1] = \mathbb{P}[x = -1]$, then



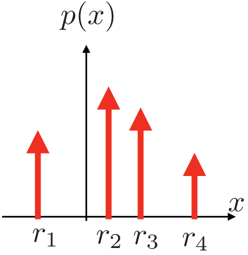
$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

$$\frac{1}{2}\|z - 1\|_0 + \frac{1}{2}\|z + 1\|_0 : \text{small}$$

Discrete signals

- Discrete signal z on a finite alphabet, $\{r_1, r_2, \dots, r_L\}$
- Probability distribution is Dirac deltas at $x = r_1, r_2, \dots, r_L$.

$$\mathbb{P}[x = r_j] = p_j, \quad p_j > 0, \quad p_1 + p_2 + \cdots + p_L = 1.$$



- The weighted sum of ℓ^0 norms

$$p_1\|z - r_1\|_0 + p_2\|z - r_2\|_0 + \cdots + p_L\|z - r_L\|_0$$

is small.

Discrete signal reconstruction

- A binary signal $x \in \{1, -1\}^n$ whose entries are drawn from

$$\mathbb{P}[x = \pm 1] = 1/2.$$

- Incomplete linear measurement

$$y = \Phi x \in \mathbb{C}^m, \quad \text{with } m \ll n$$

- Reconstruct x from y (discrete signal reconstruction)

Sum-of-absolute-values optimization

- Observing that

$$\frac{1}{2}\|x-1\|_0 + \frac{1}{2}\|x+1\|_0$$

  is small, we can say that *the sum of absolute values (SOAV)*

$$\frac{1}{2}\|x-1\|_1 + \frac{1}{2}\|x+1\|_1$$

  is also small.

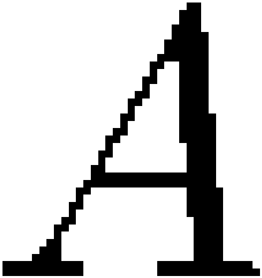
- Solve the SOAV optimization

$$\min_{z\in\mathbb{R}^n} \frac{1}{2}\|z-1\|_1 + \frac{1}{2}\|z+1\|_1 \text{ subject to } y = \Phi z$$

- In many cases, this will also do!
  - See [Nagahara, IEEE SPL, Oct. 2015]

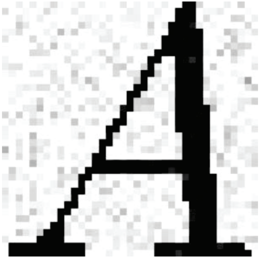
Binary image reconstruction

- Original image



Binary image reconstruction

- Original image disturbed by Gaussian noise



- Measurement: FFT and downsampling by 2
  - incomplete linear measurement

Binary image reconstruction

- Reconstruction by SOAV

Binary image reconstruction

- Reconstruction by Basis Pursuit (ℓ^1 optimization)

Discrete signal reconstruction

- Binary (or low-bit) image reconstruction
- Digital communications
- Discrete-valued control
- etc
- If you are interested, please email me. I can give you my preprints.

PART II

Sparsity Methods for Control

Maximum hands-off control

joint work with
Debasish Chatterjee (IITB)
K. S. Mallikarjuna Rao (IITB)
Daniel E. Quevedo (University of Paderborn)
Dragan Nešić (University of Melbourne)

History of Control

- Classical control for stabilization (1960—)

History of Control

- Classical control for stabilization (1960—)

History of Control

- Optimal control for enhancing performance (1970—1980)

History of Control

- Optimal control for enhancing performance (1970—1980)

History of Control

- Robust control against uncertainties (1990—2000)

History of Control

- Robust control against uncertainties (1990—2000)

History of Control

- Now: *Green control*
  - uses less fuel and electric power
  - reduces CO_2, noise, and vibration

History of Control

- Now: *Green control*
  - uses less fuel and electric power
  - reduces CO2, noise, and vibration
- *Maximum hands-off control gives a smart solution!*
  - maximizes the time duration on which the control value is 0
    - L^0 optimal: non-convex
  - can be obtained via L^1 optimal control (convex)

Outline

An example

Plant: $G(s) = 1/s^2$

$$\frac{dx(t)}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \quad x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

An example

Plant: $G(s) = 1/s^2$

$$\frac{dx(t)}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \quad x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Maximum hands-off control problem

Fix $T > 0$. Find an admissible control $u(t)$, $t \in [0, T]$ that drives the state from $x(0)$ to $x(T) = 0$, that satisfies

$$|u(t)| \leq 1, \quad \forall\, t \in [0, T],$$

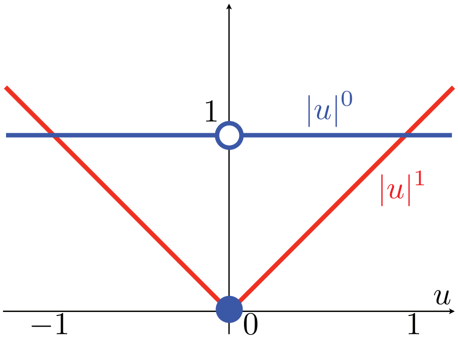
and that minimizes

$$J_0(u) = \mu(\mathrm{supp}(u)) = \int_0^T |u(t)|^0 dt,$$

the length of the support of u (L^0 norm).

An example

$$J_0(u) = \mu(\mathrm{supp}(u)) = \int_0^T |u(t)|^0 dt,$$



$$J_1(u) = \int_0^T |u(t)| dt,$$

An example

Plant: $G(s) = 1/s^2$

$$\frac{dx(t)}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \quad x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

L^1-optimal control

Fix $T > 0$. Find an admissible control $u(t)$, $t \in [0, T]$ that drives the state from $x(0)$ to $x(T) = 0$, that satisfies

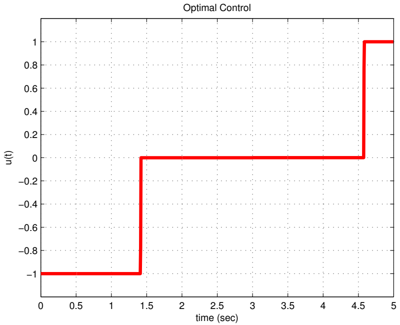
$$|u(t)| \leq 1, \quad \forall\, t \in [0, T],$$

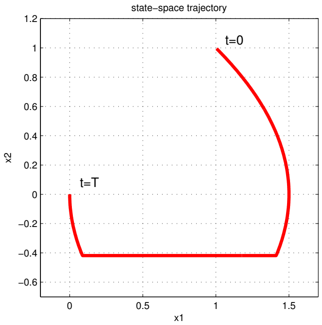
and that minimizes the L^1 norm

$$J_1(u) = \int_0^T |u(t)| dt.$$

An example

L^1-optimal control $u^*(t)$ and trajectory $x^*(t)$ [Athans and Falb, 1966]

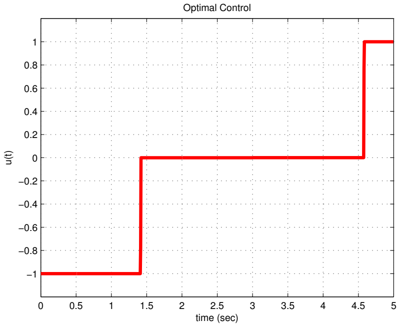


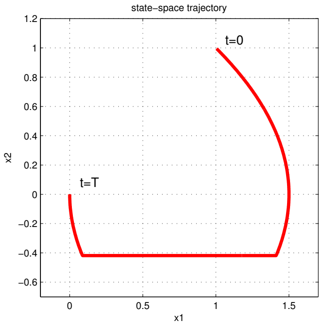


- $u^*(t) \equiv 0$ over $[3 - \sqrt{10}/2, 3 + \sqrt{10}/2] \approx [1.4, 4.6]$

An example

L^1-optimal control $u^*(t)$ and trajectory $x^*(t)$ [Athans and Falb, 1966]

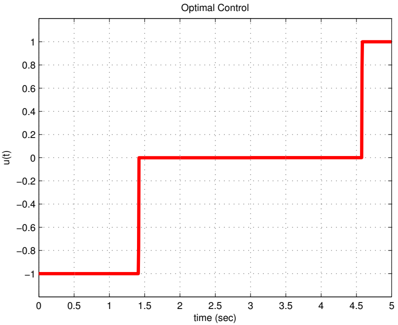


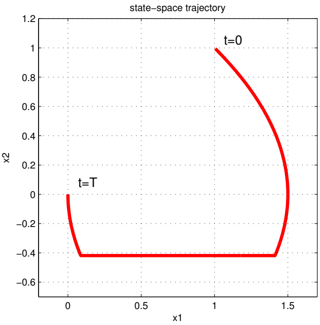


- $u^*(t) \equiv 0$ over $[3-\sqrt{10}/2, 3+\sqrt{10}/2] \approx [1.4, 4.6]$
- $u^*(t)$ is *sparse* ($\|u^*\|_0 = |\mathrm{supp}(u^*)| \approx 1.84 < 5 = T$)

An example

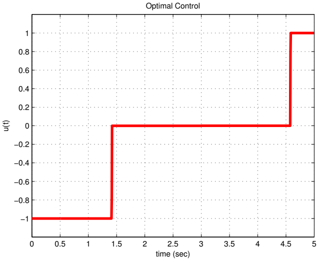
L^1-optimal control $u^*(t)$ and trajectory $x^*(t)$ [Athans and Falb, 1966]





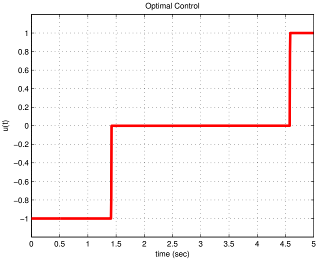
- $u^*(t) \equiv 0$ over $[3-\sqrt{10}/2, 3+\sqrt{10}/2] \approx [1.4, 4.6]$
- $u^*(t)$ is *sparse* $(\|u^*\|_0 = |\mathrm{supp}(u^*)| \approx 1.84 < 5 = T)$
- In fact, it is *the sparsest* (i.e., *maximum hands-off control*).

Why hands-off control is *green*?



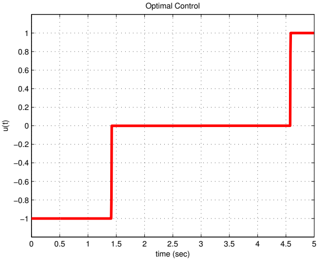
- Reduced fuel and electric power consumption

Why hands-off control is *green*?



- Reduced fuel and electric power consumption
- Reduced CO2, noise, and vibration

Why hands-off control is *green*?



- Reduced fuel and electric power consumption
- Reduced CO2, noise, and vibration
- Data compression
  - Sparse signals can be effectively compressed; see e.g. [Nagahara, Quevedo, Østergaard, IEEE Trans. AC 2014]

Maximum hands-off control and L^1 optimality

Plant

$$\frac{dx(t)}{dt} = f(x(t)) + g(x(t))u(t), \quad t \geq 0, \quad x(0) = x_0$$

$$x(t) \in \mathbb{R}^n, \quad u(t) \in \mathbb{R}$$

Maximum hands-off control and L^1 optimality

Plant

$$\frac{dx(t)}{dt} = f(x(t)) + g(x(t))u(t), \quad t \geq 0, \quad x(0) = x_0$$

$$x(t) \in \mathbb{R}^n, \quad u(t) \in \mathbb{R}$$

Theorem

Assume that the L^1-optimal control problem is *normal* [a] (or *non singular*) and has at least one solution. Then

$$\{L^0 \text{ optimal controls}\} = \{L^1 \text{ optimal controls}\}$$

---

[a] When the optimal control is *uniquely determined almost everywhere* from the minimum principle, the control problem is called *normal*.

Maximum hands-off control and L^1 optimality

Plant

$$\frac{dx(t)}{dt} = f(x(t)) + g(x(t))u(t), \quad t \geq 0, \quad x(0) = x_0$$

$$x(t) \in \mathbb{R}^n, \quad u(t) \in \mathbb{R}$$

Theorem

Assume that the L^1-optimal control problem is *normal* [a] (or *non singular*) and has at least one solution. Then

$$\{L^0 \text{ optimal controls}\} = \{L^1 \text{ optimal controls}\}$$

[a] When the optimal control is *uniquely determined almost everywhere* from the minimum principle, the control problem is called *normal*.

A maximum hands-off control problem (non convex optimization) can be solved via a related L^1 optimal control problem (convex)!

Sufficient condition for normality

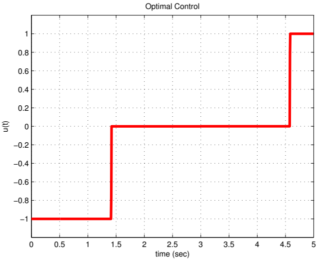
Lemma [Athans & Falb, 1966]

Assume the plant is given by

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t), \quad t \geq 0.$$

If the plant is *controllable* and *A is non singular*, then for any initial state $x(0) \in \mathbb{R}^n$, the L^1-optimal control problem is normal.

L^1/L^2-optimal control for continuous control



- Maximum-hands off control is discontinuous
  - the "bang-off-bang" property
- Smoothing by adding L^2 norm:

$$J_{12} = \|u\|_1 + \frac{1}{2} r \|u\|_2^2$$

- L^1/L^2-optimal control is *continuous in t*.

L^1/L^2-optimal control for continuous control

L^1/L^2-optimal control

Plant: $\frac{dx}{dt} = f(x) + g(x)u$

Assumption: f, g, $\frac{df}{dx}$, $\frac{dg}{dx}$ are continuous in x.

Constraints: $x(0) = x_0$; $x(T) = 0$; $|u(t)| \leq 1 \ \forall t \in [0, T]$

Cost function: $J_{12} = \|u\|_1 + \frac{1}{2} r \|u\|_2^2$

L^1/L^2-optimal control for continuous control

L^1/L^2-optimal control

Plant: $\frac{dx}{dt} = f(x) + g(x)u$

Assumption: f, g, $\frac{df}{dx}$, $\frac{dg}{dx}$ are continuous in x.

Constraints: $x(0) = x_0$; $x(T) = 0$; $|u(t)| \leq 1 \ \forall t \in [0, T]$

Cost function: $J_{12} = \|u\|_1 + \frac{1}{2} r \|u\|_2^2$

Proposition

The L^1/L^2-optimal control $u_{12}^*(t)$ is continuous in t over $[0, T]$.

L^1/L^2-optimal control for continuous control

L^1/L^2-optimal control

Plant: $\frac{dx}{dt} = f(x) + g(x)u$

Assumption: f, g, $\frac{df}{dx}$, $\frac{dg}{dx}$ are continuous in x.

Constraints: $x(0) = x_0$; $x(T) = 0$; $|u(t)| \leq 1 \ \forall t \in [0, T]$

Cost function: $J_{12} = \|u\|_1 + \frac{1}{2} r \|u\|_2^2$

Proposition

The L^1/L^2-optimal control $u_{12}^*(t)$ is continuous in t over $[0, T]$.

Proposition

Assume the L^1-optimal control problem is normal and its solution exists. Then

$$u_{12}^*(t) \to u_1^*(t) = u_0^*(t), \quad \text{a.a. } t \in [0, T],$$

as $r \to 0$.

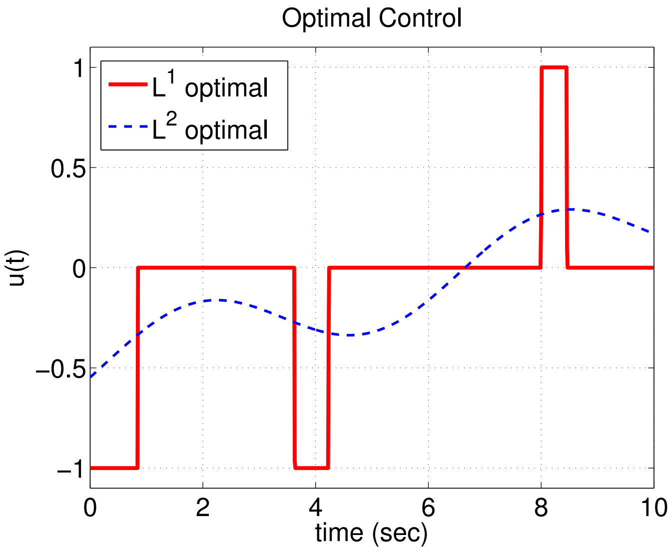
Example: control problem

- Plant: $P(s) = \frac{1}{s^2(s^2+1)}$

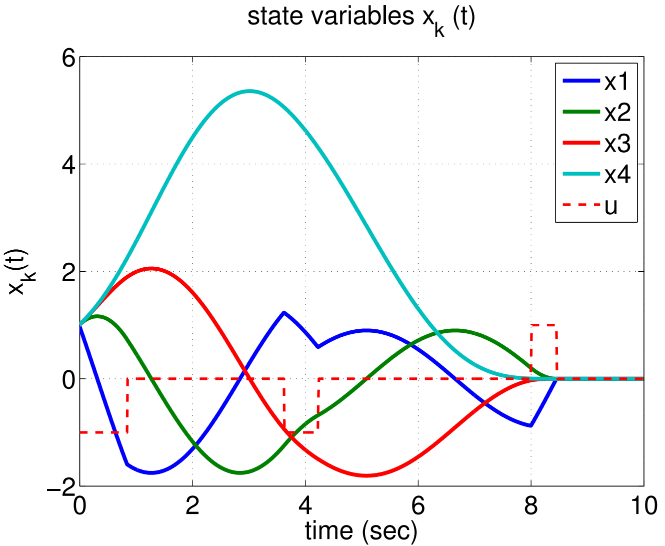
$$\frac{d\boldsymbol{x}(t)}{dt} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} u(t).$$

- Final time: $T = 10$.
- State Constraints: $x(0) = [1, 1, 1, 1]^\top$ and $x(10) = 0$
- Control constraint: $|u(t)| \leq 1, \quad \forall t \in [0, 10]$

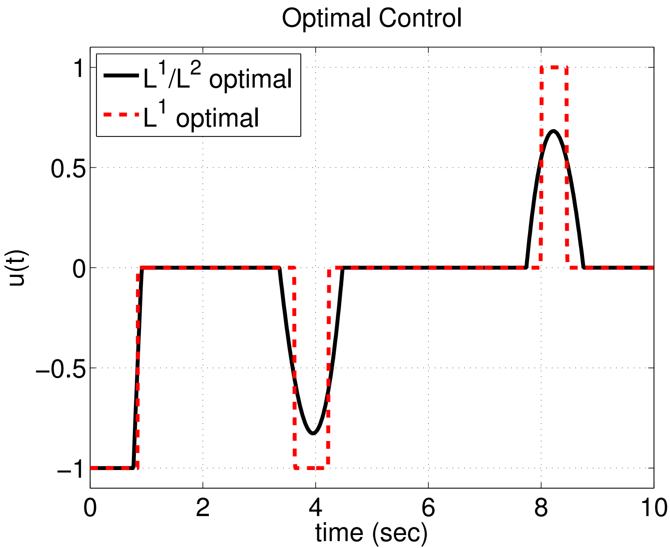
Examples: optimal controls



Examples: states with maximum hands-off control



Examples: L^1/L^2-optimal control



Conclusion

- Maximum hands-off control is *green control*.
  - uses less fuel and electric power
  - reduces CO2, noise, and vibration
  - gives effective data compression for networked control systems
- L^0 optimality $=$ L^1 optimality
  - under the assumption of normality.
- Continuous control by L^1/L^2-optimal control
- Characterization of maximum hands-off control (i.e. L^0 optimal control) is given in the following paper:
  - D. Chatterjee, M. Nagahara, D. E. Quevedo, and K. S. M. Rao, "Characterization of maximum hands-off control," *Systems and Control Letters*, 2016, to be published.

Today's talk was on...

- Relation between sparsity and discreteness

Today's talk was on...

- Relation between sparsity and discreteness
- Sparsity methods for control (maximum hands-off control) for *green* technology

Today's talk was on...

- Relation between sparsity and discreteness
- Sparsity methods for control (maximum hands-off control) for *green* technology
- *Interplay between control and sparse signal processing*
  - Collaborative work by researchers and engineers on control, signal processing, communications, etc is highly important.

Today's talk was on...

- Relation between sparsity and discreteness
- Sparsity methods for control (maximum hands-off control) for *green* technology
- *Interplay between control and sparse signal processing*
  - Collaborative work by researchers and engineers on control, signal processing, communications, etc is highly important.
- *Let's get started!*