Optimal input design for nonlinear dynamical systems: a graph-theory approach

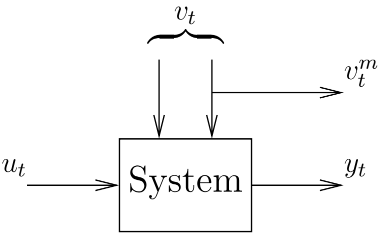
ROYAL INSTITUTE OF TECHNOLOGY

Patricio E. Valenzuela

Department of Automatic Control and ACCESS Linnaeus Centre
KTH Royal Institute of Technology, Stockholm, Sweden

Seminar Uppsala university
January 16, 2015

System identification



- System identification: Modeling based on input-output data.
- Basic entities: data set, model structure, identification method.

ROYAL INSTITUTE OF TECHNOLOGY

System identification

The maximum likelihood method:

$$\hat{\theta}_{n_{\text{seq}}} = \arg\max_{\theta \in \Theta} l_\theta(y_{1:n_{\text{seq}}})$$

where

$$l_\theta(y_{1:n_{\text{seq}}}) := \log p_\theta(y_{1:n_{\text{seq}}})$$

As $n_{\text{seq}} \to \infty$, we have:

- $\hat{\theta}_{n_{\text{seq}}} \to \theta_0$
- $\sqrt{n_{\text{seq}}}\left(\hat{\theta}_{n_{\text{seq}}} - \theta_0\right) \to$ Normal with zero mean and covariance $\{\mathcal{I}_F^e\}^{-1}$

System identification

- $\sqrt{n_{\text{seq}}}\left(\hat{\theta}_{n_{\text{seq}}} - \theta_0\right) \to$ Normal with zero mean and covariance $\{\mathcal{I}_F^e\}^{-1}$

where

$$\mathcal{I}_F^e := \mathbf{E}\left\{ \frac{\partial}{\partial\theta} l_\theta(y_{1:n_{\text{seq}}})\Big|_{\theta=\theta_0} \frac{\partial}{\partial\theta^\top} l_\theta(y_{1:n_{\text{seq}}})\Big|_{\theta=\theta_0} \Big| u_{1:n_{\text{seq}}} \right\}$$

Different $u_{1:n_{\text{seq}}}$s $\Rightarrow$ different $\mathcal{I}_F^e$s.

Covariance matrix of $\sqrt{n_{\text{seq}}}\left(\hat{\theta}_{n_{\text{seq}}} - \theta_0\right)$ affected by $u_{1:n_{\text{seq}}}$!

Input design for dynamic systems

- Input design: Maximize information from an experiment.
- Existing methods: focused on linear systems.
- Recent developments for nonlinear systems (Hjalmarsson 2007, Larsson 2010, Gopaluni 2011, De Cock-Gevers-Schoukens 2013, Forgione-Bombois-Van den Hof-Hjalmarsson 2014).

Input design for dynamic systems

Challenges for nonlinear systems:

- Problem complexity (usually non-convex).
- Model restrictions.
- Input restrictions.

How could we overcome these limitations?

ROYAL INSTITUTE OF TECHNOLOGY

Summary

1. We present a method for input design for dynamic systems.

2. The method is also suitable for nonlinear systems.

Outline

Problem formulation for output-error models

Input design based on graph theory

Extension to nonlinear SSM

Closed-loop application oriented input design

Conclusions and future work

ROYAL INSTITUTE OF TECHNOLOGY

Problem formulation for output-error models

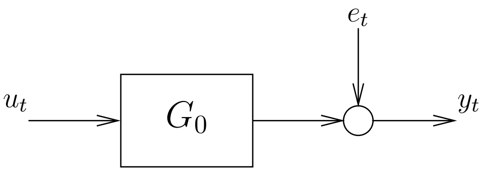
Input design based on graph theory

Extension to nonlinear SSM

Closed-loop application oriented input design

Conclusions and future work

Problem formulation for output-error models



G_0 is a system, where:

- e_t: white noise (variance λ_e)
- u_t: input
- y_t: system output

Goal: Design

$$u_{1:n_{\mathrm{seq}}} := (u_1, \ldots, u_{n_{\mathrm{seq}}})$$

as a realization of a stationary process *maximizing* $\mathcal{I}_F$.

Problem formulation for output-error models

Here,

$$\mathcal{I}_F := \frac{1}{\lambda_e}\mathbf{E}\left\{\sum_{t=1}^{n_{\text{seq}}} \psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top\right\}$$

$$= \frac{1}{\lambda_e}\int \sum_{t=1}^{n_{\text{seq}}} \psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top \, dP(u_{1:n_{\text{seq}}})$$

$$\psi_t^{\theta_0}(u_t) := \nabla_\theta \hat{y}_t(u_t)\big|_{\theta=\theta_0}$$

$$\hat{y}_t(u_t) := G(u_t;\, \theta)$$

Design $u_{1:n_{\text{seq}}} \in \mathbb{R}^{n_{\text{seq}}} \Leftrightarrow$ Design $P(u_{1:n_{\text{seq}}}) \in \mathcal{P}$.

ROYAL INSTITUTE OF TECHNOLOGY

Problem formulation for output-error models

Here,

$$\begin{aligned}\mathcal{I}_F &:= \frac{1}{\lambda_e}\mathbf{E}\left\{\sum_{t=1}^{n_{\text{seq}}} \psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top\right\} \\ &= \frac{1}{\lambda_e}\int \sum_{t=1}^{n_{\text{seq}}} \psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top \, dP(u_{1:n_{\text{seq}}})\end{aligned}$$

Assumption

$u_t \in \mathcal{C}$ *($\mathcal{C}$ finite set)*

$$\mathcal{I}_F = \frac{1}{\lambda_e}\sum_{u_{1:n_{\text{seq}}}\in\mathcal{C}^{n_{\text{seq}}}}\sum_{t=1}^{n_{\text{seq}}} \psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top p(u_{1:n_{\text{seq}}})$$

Problem formulation for output-error models

Characterizing $p(u_{1:n_{\text{seq}}}) \in \mathcal{P}_{\mathcal{C}}$:

- p nonnegative,
- $\sum p(\mathbf{x}) = 1$,
- p is shift invariant.

ROYAL INSTITUTE OF TECHNOLOGY

Problem formulation for output-error models

Problem

Design $u_{1:n_{\text{seq}}}^{\text{opt}} \in \mathcal{C}^{n_{\text{seq}}}$ as a realization from $p^{\text{opt}}(u_{1:n_{\text{seq}}})$, where

$$p^{\text{opt}}(u_{1:n_{\text{seq}}}) := \arg\max_{p \in \mathcal{P}_{\mathcal{C}}} h(\mathcal{I}_F(p))$$

where $h : \mathbb{R}^{n_\theta \times n_\theta} \to \mathbb{R}$ is a matrix concave function, and

$$\mathcal{I}_F(p) = \frac{1}{\lambda_e} \sum_{u_{1:n_{\text{seq}}} \in \mathcal{C}^{n_{\text{seq}}}} \sum_{t=1}^{n_{\text{seq}}} \psi_t^{\theta_0}(u_t) \psi_t^{\theta_0}(u_t)^\top p(u_{1:n_{\text{seq}}})$$

ROYAL INSTITUTE OF TECHNOLOGY

Problem formulation for output-error models

Input design based on graph theory

Extension to nonlinear SSM

Closed-loop application oriented input design

Conclusions and future work

ROYAL INSTITUTE OF TECHNOLOGY

Input design problem

Problem

Design $u_{1:n_{\text{seq}}}^{\text{opt}} \in \mathcal{C}^{n_{\text{seq}}}$ as a realization from $p^{\text{opt}}(u_{1:n_{\text{seq}}})$, where

$$p^{\text{opt}}(u_{1:n_{\text{seq}}}) := \arg\max_{p \in \mathcal{P}_{\mathcal{C}}} h(\mathcal{I}_F(p))$$

where $h : \mathbb{R}^{n_\theta \times n_\theta} \to \mathbb{R}$ is a matrix concave function, and

$$\mathcal{I}_F(p) = \frac{1}{\lambda_e} \sum_{u_{1:n_{\text{seq}}} \in \mathcal{C}^{n_{\text{seq}}}} \sum_{t=1}^{n_{\text{seq}}} \psi_t^{\theta_0}(u_t) \psi_t^{\theta_0}(u_t)^\top p(u_{1:n_{\text{seq}}})$$

Issues:

1. $\mathcal{I}_F(p)$ requires a sum of n_{seq}-dimensional terms (n_{seq} large).
2. How could we represent an element in $\mathcal{P}_{\mathcal{C}}$?

ROYAL INSTITUTE OF TECHNOLOGY

Input design problem

Solving the issues:

1. $\mathcal{I}_F(p)$ requires a sum of n_{seq}-dimensional terms (n_{seq} large).

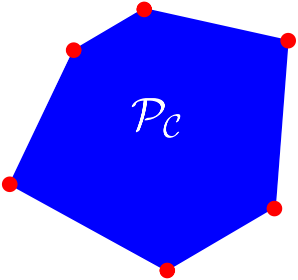
Assumption

$u_{1:n_{\text{seq}}}$ is a realization of a stationary process with memory n_m ($n_m < n_{\text{seq}}$).

$\Rightarrow \mathcal{I}_F(p)$ requires a sum of n_m-dimensional terms.

Minimum n_m: related with the memory of the system.

Input design problem



Solving the issues:

2. How could we represent an element in $\mathcal{P}_{\mathcal{C}}$?

$\mathcal{P}_{\mathcal{C}}$ is a polyhedron.

$\mathcal{V}_{\mathcal{P}_{\mathcal{C}}}$: Set of extreme points of $\mathcal{P}_{\mathcal{C}}$.

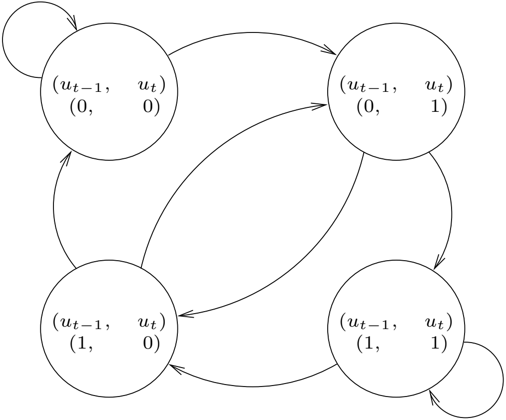
$\Rightarrow \mathcal{P}_{\mathcal{C}}$ can be described as a convex combination of $\mathcal{V}_{\mathcal{P}_{\mathcal{C}}}$.

The elements in $\mathcal{V}_{\mathcal{P}_{\mathcal{C}}}$ can be found by using Graph theory!

ROYAL INSTITUTE OF TECHNOLOGY

Graph theory in input design

Example: de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$.

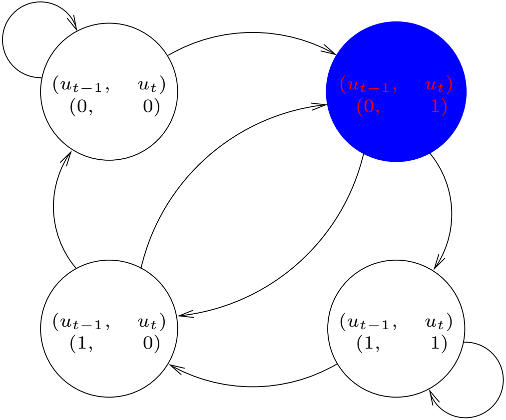


Elements in $\mathcal{V}_{\mathcal{P}_\mathcal{C}}$ $\Leftrightarrow$ Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$

Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$ $\Leftrightarrow$ Elementary cycles in $\mathcal{G}_{\mathcal{C}^{(n_m-1)}}$

Graph theory in input design

Example: de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$.

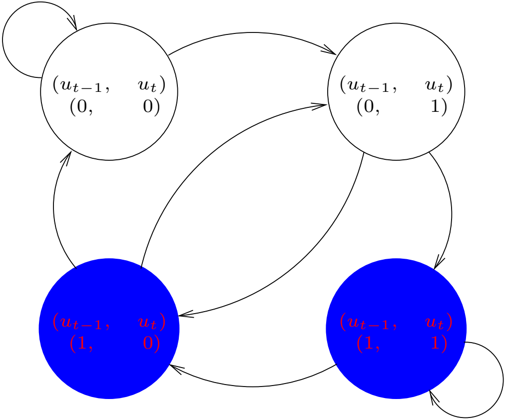


Elements in $\mathcal{V}_{\mathcal{P}_\mathcal{C}}$ $\Leftrightarrow$ Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$

Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$ $\Leftrightarrow$ Elementary cycles in $\mathcal{G}_{\mathcal{C}^{(n_m-1)}}$

Graph theory in input design

Example: de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$.



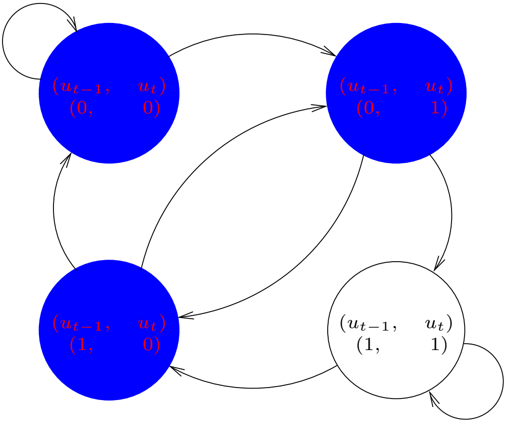
Elements in $\mathcal{V}_{\mathcal{P}_\mathcal{C}}$ $\Leftrightarrow$ Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$

Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$ $\Leftrightarrow$ Elementary cycles in $\mathcal{G}_{\mathcal{C}^{(n_m-1)}}$

ROYAL INSTITUTE OF TECHNOLOGY

Graph theory in input design

Example: de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$.

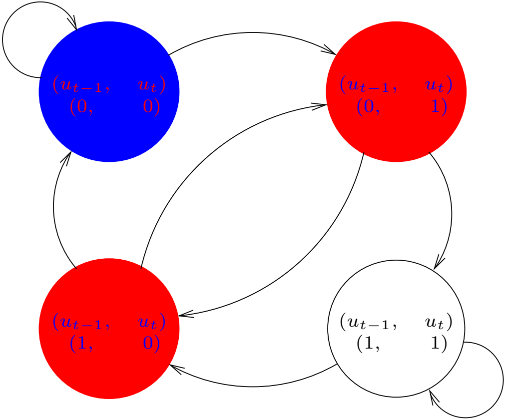


Elements in $\mathcal{V}_{\mathcal{P}_\mathcal{C}}$ $\Leftrightarrow$ Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$

Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$ $\Leftrightarrow$ Elementary cycles in $\mathcal{G}_{\mathcal{C}^{(n_m-1)}}$

Graph theory in input design

Example: de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$.



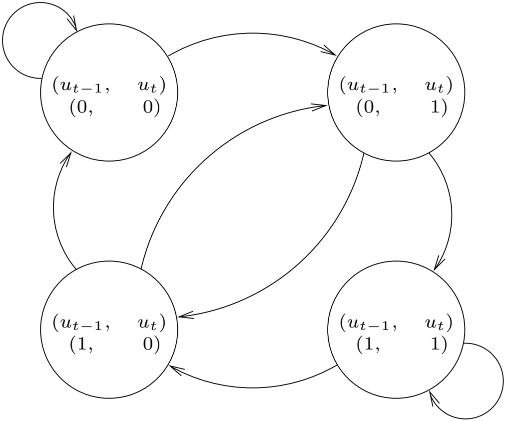
Elements in $\mathcal{V}_{\mathcal{P}_\mathcal{C}}$ $\Leftrightarrow$ Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$

Prime cycles in $\mathcal{G}_{\mathcal{C}^{n_m}}$ $\Leftrightarrow$ Elementary cycles in $\mathcal{G}_{\mathcal{C}^{(n_m-1)}}$

ROYAL INSTITUTE OF TECHNOLOGY

Graph theory in input design

Example: de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$.



There are algorithms to find elementary cycles (Johnson 1975, Tarjan 1972).

Graph theory in input design

Once $v_i \in \mathcal{V}_{\mathcal{P}_\mathcal{C}}$ is known

$\Rightarrow$ The distribution for each v_i is known.

$\Rightarrow$ An input signal $\{u_t^i\}_{t=0}^{t=N}$ can be drawn from v_i.

Therefore,

$$\mathcal{I}_F^{(i)} := \frac{1}{\lambda_e} \sum_{u_{1:n_m} \in \mathcal{C}^{n_m}} \sum_{t=1}^{n_m} \psi_t^{\theta_0}(u_t) \psi_t^{\theta_0}(u_t)^\top v_i(u_{1:n_m})$$
$$\approx \frac{1}{\lambda_e N} \sum_{t=1}^{N} \psi_t^{\theta_0}(u_t) \psi_t^{\theta_0}(u_t)^\top$$

for all $v_i \in \mathcal{V}_{\mathcal{P}_\mathcal{C}}$.

Graph theory in input design

Therefore,

$$\mathcal{I}_F^{(i)} := \frac{1}{\lambda_e} \sum_{u_{1:n_m} \in \mathcal{C}^{n_m}} \sum_{t=1}^{n_m} \psi_t^{\theta_0}(u_t) \psi_t^{\theta_0}(u_t)^\top v_i(u_{1:n_m})$$
$$\approx \frac{1}{\lambda_e\, N} \sum_{t=1}^{N} \psi_t^{\theta_0}(u_t) \psi_t^{\theta_0}(u_t)^\top$$

for all $v_i \in \mathcal{V}_{\mathcal{P}_\mathcal{C}}$.

The sum is approximated by Monte-Carlo!

Input design based on graph theory

To design an experiment in $\mathcal{C}^{n_m}$:

1. Compute all the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$.
2. Generate the input signals $\{u_t^i\}_{t=0}^{t=N}$ from the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$, for each $i \in \{1, \dots, n_{\mathcal{V}}\}$.
3. For each $i \in \{1, \dots, n_{\mathcal{V}}\}$, approximate $\mathcal{I}_F^{(i)}$ by using

$$\mathcal{I}_F^{(i)} \approx \frac{1}{\lambda_e N} \sum_{t=1}^{N} \psi_t^{\theta_0}(u_t) \psi_t^{\theta_0}(u_t)^\top$$

Input design based on graph theory

To design an experiment in $\mathcal{C}^{n_m}$:

4. Define $\gamma := \{\alpha_1, \dots, \alpha_{n_\mathcal{V}}\} \in \mathbb{R}^{n_\mathcal{V}}$.
   Solve
   $$\gamma^{\text{opt}} := \arg \max_{\gamma \in \mathbb{R}^{n_\mathcal{V}}} h(\mathcal{I}_F^{\text{app}}(\gamma))$$
   where
   $$\mathcal{I}_F^{\text{app}}(\gamma) := \sum_{i=1}^{n_\mathcal{V}} \alpha_i \, \mathcal{I}_F^{(i)}$$
   $$\sum_{i=1}^{n_\mathcal{V}} \alpha_i = 1$$
   $$\alpha_i \geq 0 \,, \text{ for all } i \in \{1, \dots, n_\mathcal{V}\}$$

Input design based on graph theory

ROYAL INSTITUTE OF TECHNOLOGY

To design an experiment in $\mathcal{C}^{n_m}$:

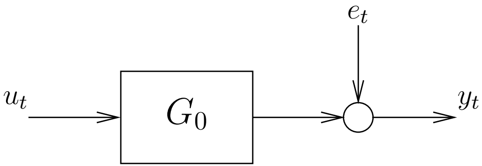
5. The optimal pmf p^{opt} is given by

$$p^{\text{opt}} = \sum_{i=1}^{n_{\mathcal{V}}} \alpha_i^{\text{opt}} \, v_i$$

6. Sample $u_{1:n_{\text{seq}}}$ from p^{opt} using Markov chains.

$\mathcal{I}_F^{\text{app}}(\gamma)$ linear in the decision variables $\Rightarrow$ The problem is convex!

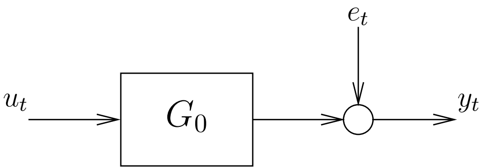
Example I



$$G(u_t;\theta) = \begin{cases} x_{t+1} = \dfrac{1}{\theta_1 + x_t^2} + u_t \\ y_t = \theta_2\, x_t^2 + e_t \\ x_1 = 0 \end{cases}$$

with $\theta = \begin{bmatrix}\theta_1 & \theta_2\end{bmatrix}^\top = \theta_0 = \begin{bmatrix}0.8 & 2\end{bmatrix}^\top$.
e_t: white noise, Gaussian, zero mean, variance $\lambda_e = 1$.

Example I



$$G(u_t;\theta) = \begin{cases} x_{t+1} = \dfrac{1}{\theta_1 + x_t^2} + u_t \\ y_t = \theta_2\, x_t^2 + e_t \\ x_1 = 0 \end{cases}$$

with $\theta = \begin{bmatrix}\theta_1 & \theta_2\end{bmatrix}^\top = \theta_0 = \begin{bmatrix}0.8 & 2\end{bmatrix}^\top$.
We consider $h(\cdot) = \log\det(\cdot)$, and $n_{\text{seq}} = 10^4$.

Example I

$$G(u_t;\theta) = \begin{cases} x_{t+1} = \dfrac{1}{\theta_1 + x_t^2} + u_t \\ y_t = \theta_2\, x_t^2 + e_t \\ x_1 = 0 \end{cases}$$

Results:

$h(\mathcal{I}_F)$	Case 1	Case 2	Case 3	Binary
$\log\{\det(\mathcal{I}_F)\}$	3.82	4.50	4.48	3.47

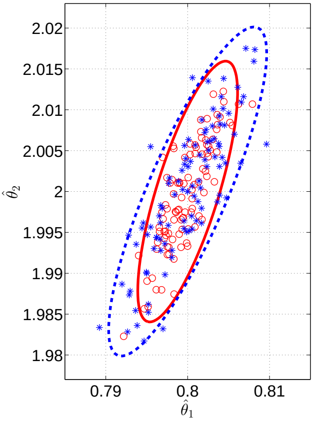
Case 1: $n_m = 2$, $\mathcal{C} = \{-1,\, 0,\, 1\}$

Case 2: $n_m = 1$, $\mathcal{C} = \{-1,\, -1/3,\, 1/3,\, 1\}$

Case 3: $n_m = 1$, $\mathcal{C} = \{-1,\, -0.5,\, 0,\, 0.5,\, 1\}$

Example I

Results ($95\,\%$ confidence ellipsoids):



Red: Case 2; **Blue:** Binary input.

Problem formulation for output-error models

Input design based on graph theory

Extension to nonlinear SSM

Closed-loop application oriented input design

Conclusions and future work

Extension to nonlinear SSM

Nonlinear state space model:

$$x_0 \sim \mu(x_0)$$
$$x_t | x_{t-1} \sim f_\theta(x_t | x_{t-1}, u_{t-1})$$
$$y_t | x_t \sim g_\theta(y_t | x_t, u_t)$$

where $\theta \in \Theta$.

- f_θ, g_θ, μ: pdfs
- x_t: states
- u_t: input
- y_t: system output

Goal: Design

$$u_{1:n_{\text{seq}}} := (u_1, \ldots, u_{n_{\text{seq}}})$$

as a realization of a stationary process *maximizing* $\mathcal{I}_F$.

Extension to nonlinear SSM

Here,

$$\mathcal{I}_F = \mathbf{E}\left\{\mathcal{S}(\theta_0)\mathcal{S}^\top(\theta_0)\right\}$$
$$\mathcal{S}(\theta_0) = \nabla_\theta \log p_\theta(y_{1:n_{\text{seq}}}|u_{1:n_{\text{seq}}})\big|_{\theta=\theta_0}$$

Design $u_{1:n_{\text{seq}}} \in \mathbb{R}^{n_{\text{seq}}} \Leftrightarrow$ Design $P(u_{1:n_{\text{seq}}}) \in \mathcal{P}$.

Extension to nonlinear SSM

Fisher's identity:

$$\nabla_\theta \log p_\theta(y_{1:T}|u_{1:T}) = \mathbf{E}\left\{\nabla_\theta \log p_\theta(x_{1:T}, y_{1:T}|u_{1:T})\middle|y_{1:T}, u_{1:T}\right\}$$

$$\nabla_\theta \log p_\theta(y_{1:T}|u_{1:T}) = \sum_{t=1}^{T} \int_{\mathcal{X}^2} \xi_\theta(x_{t-1:t},\, u_t) p_\theta(x_{t-1:t}|y_{1:T}) \mathrm{d}x_{t-1:t}$$

with

$$\xi_\theta(x_{t-1:t},\, u_t) = \nabla_\theta \Big[\log f_\theta(x_t|x_{t-1}, u_{t-1}) + \log g_\theta(y_t|x_t, u_t) \Big]$$

Assumption

$u_t \in \mathcal{C}$ *($\mathcal{C}$ finite set)*

Extension to nonlinear SSM

Recall $\mathcal{P}_{\mathcal{C}}$:

- p nonnegative,
- $\sum p(\mathbf{x}) = 1$,
- p is shift invariant.

ROYAL INSTITUTE OF TECHNOLOGY

Extension to nonlinear SSM

Problem

Design $u^{\text{opt}}_{1:n_{\text{seq}}} \in \mathcal{C}^{n_{\text{seq}}}$ as a realization from $p^{\text{opt}}(u_{1:n_{\text{seq}}})$, where

$$p^{\text{opt}}(u_{1:n_{\text{seq}}}) := \arg\max_{p \in \mathcal{P}_{\mathcal{C}}} h(\mathcal{I}_F(p))$$

where $h : \mathbb{R}^{n_\theta \times n_\theta} \to \mathbb{R}$ is a matrix concave function, and

$$\mathcal{I}_F(p) = \mathbf{E}\left\{\mathcal{S}(\theta_0)\mathcal{S}^\top(\theta_0)\right\}$$

Input design problem for nonlinear SSM

Problem

Design $u_{1:n_{\text{seq}}}^{\text{opt}} \in \mathcal{C}^{n_{\text{seq}}}$ as a realization from $p^{\text{opt}}(u_{1:n_{\text{seq}}})$, where

$$p^{\text{opt}}(u_{1:n_{\text{seq}}}) := \arg\max_{p \in \mathcal{P}_{\mathcal{C}}} h(\mathcal{I}_F(p))$$

where $h : \mathbb{R}^{n_\theta \times n_\theta} \to \mathbb{R}$ is a matrix concave function, and

$$\mathcal{I}_F(p) = \mathbf{E}\left\{\mathcal{S}(\theta_0)\mathcal{S}^\top(\theta_0)\right\}$$

Issues:

1. How could we represent an element in $\mathcal{P}_{\mathcal{C}}$? $\Rightarrow$ Use the graph theory approach!
2. How could we compute $\mathcal{I}_F(p)$?

Input design based on graph theory (revisited)

To design an experiment in $\mathcal{C}^{n_m}$:

1. Compute all the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$.
2. Generate the input signals $\{u_t^i\}_{t=0}^{t=N}$ from the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$, for each $i \in \{1, \ldots, n_{\mathcal{V}}\}$.
3. **For each $i \in \{1, \ldots, n_{\mathcal{V}}\}$, approximate $\mathcal{I}_F^{(i)}$ by using**

$$\mathcal{I}_F^{(i)} := \mathbf{E}_{v_i(u_{1:n_m})}\left\{\mathcal{S}(\theta_0)\mathcal{S}^\top(\theta_0)\right\}$$

$\approx$ **(new expression required!)**

Estimating $\mathcal{I}_F$

Approximate $\mathcal{I}_F$ as

$$\hat{\mathcal{I}}_F := \frac{1}{M} \sum_{m=1}^{M} \mathcal{S}_m(\theta_0)\mathcal{S}_m^\top(\theta_0)$$

- Difficulty: $\mathcal{S}_m(\theta_0)$ is not available.
- Solution: Estimate $\mathcal{S}_m(\theta_0)$ using particle methods!

ROYAL INSTITUTE OF TECHNOLOGY

Particle methods to estimate $\mathcal{S}_m(\theta_0)$

- Goal: Approximate $\{p_\theta(x_{1:t}|y_{1:t})\}_{t\geq 1}$.
- $\{x_{1:t}^{(i)}, w_t^{(i)}\}_{i=1}^N$: Particle system.
- Approach: Auxiliary particle filter + Fixed-lag smoother.

Particle methods to estimate $\mathcal{S}_m(\theta_0)$

Estimate $\mathcal{S}_m(\theta_0)$ as

$$\hat{S}_m(\theta_0) := \sum_{t=1}^{T}\sum_{i=1}^{N} w_{\kappa_t}^{(i)} \xi_{\theta_0}(x_{t-1}^{a_{\kappa_t,t-1}^{(i)}},\, x_t^{a_{\kappa_t,t}^{(i)}},\, u_t)$$

where

$$\xi_\theta(x_{t-1:t},\, u_t) = \nabla_\theta\Big[\log f_\theta(x_t|x_{t-1}, u_{t-1}) + \log g_\theta(y_t|x_t, u_t)\Big]$$

Input design based on graph theory (revisited)

KTH VETENSKAP OCH KONST
ROYAL INSTITUTE OF TECHNOLOGY

To design an experiment in $\mathcal{C}^{n_m}$:

1. Compute all the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$.
2. Generate the input signals $\{u_t^i\}_{t=0}^{t=N}$ from the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$, for each $i \in \{1, \ldots, n_\mathcal{V}\}$.
3. **For each $i \in \{1, \ldots, n_\mathcal{V}\}$, approximate $\mathcal{I}_F^{(i)}$ by using**

$$\mathcal{I}_F^{(i)} := \mathbf{E}_{v_i(u_{1:n_m})} \left\{ \mathcal{S}(\theta_0) \mathcal{S}^\top(\theta_0) \right\}$$
$$\approx \frac{1}{M} \sum_{m=1}^{M} \hat{\mathcal{S}}_m(\theta_0) \hat{\mathcal{S}}_m^\top(\theta_0)$$

ROYAL INSTITUTE OF TECHNOLOGY

Input design based on graph theory (revisited)

To design an experiment in $\mathcal{C}^{n_m}$:

4. Define $\gamma := \{\alpha_1, \ldots, \alpha_{n_\mathcal{V}}\} \in \mathbb{R}^{n_\mathcal{V}}$.
   For $k \in \{1, \ldots, K\}$**, solve**

$$\gamma^{\mathrm{opt},k} := \arg \max_{\gamma \in \mathbb{R}^{n_\mathcal{V}}} h(\mathcal{I}_F^{\mathrm{app},k}(\gamma_k))$$

   where

$$\mathcal{I}_F^{\mathrm{app},k}(\gamma_k) := \sum_{i=1}^{n_\mathcal{V}} \alpha_{i,k}\, \mathcal{I}_F^{(i),k}$$

$$\sum_{i=1}^{n_\mathcal{V}} \alpha_{i,k} = 1$$

$$\alpha_{i,k} \geq 0\,, \text{ for all } i \in \{1, \ldots, n_\mathcal{V}\}$$

   Compute $\overline{\gamma}$ **as the sample mean of** $\{\gamma^{\mathrm{opt},k}\}_{k=1}^{K}$**.**

ROYAL INSTITUTE OF TECHNOLOGY

Input design based on graph theory (revisited)

To design an experiment in $\mathcal{C}^{n_m}$:

5. The optimal pmf p^{opt} is given by

$$p^{\text{opt}} = \sum_{i=1}^{n_{\mathcal{V}}} \overline{\alpha}_i^{\text{opt}} \, v_i$$

6. Sample $u_{1:n_{\text{seq}}}$ from p^{opt} using Markov chains.

$\mathcal{I}_F^{\text{app}}(\gamma)$ linear in the decision variables $\Rightarrow$ The problem is convex!

Example II

Nonlinear state space model:

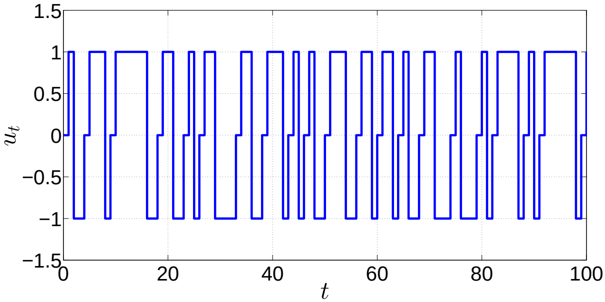
$$
\begin{aligned}
x_{t+1} &= \theta_1 x_t + \frac{x_t}{\theta_2 + x_t^2} + u_t + v_t, && v_t \sim \mathcal{N}(0, 0.1^2) \\
y_t &= \frac{1}{2} x_t + \frac{2}{5} x_t^2 + e_t, && e_t \sim \mathcal{N}(0, 0.1^2)
\end{aligned}
$$

where $\theta = \begin{bmatrix} \theta_1 & \theta_2 \end{bmatrix}^\top$, $\theta_0 = \begin{bmatrix} 0.7 & 0.6 \end{bmatrix}^\top$.

Input design: $n_{\text{seq}} = 5 \cdot 10^3$, $n_m = 2$, $\mathcal{C} = \{-1, 0, 1\}$, and $h(\cdot) = \log\det(\cdot)$.

Example II

Input sequence:



Example II

Results:

Input / $h(\widehat{\mathcal{I}}_F)$	$\log\det(\widehat{\mathcal{I}}_F)$
Optimal	$\mathbf{25.34}$
Binary	24.75
Uniform	24.38

Problem formulation for output-error models

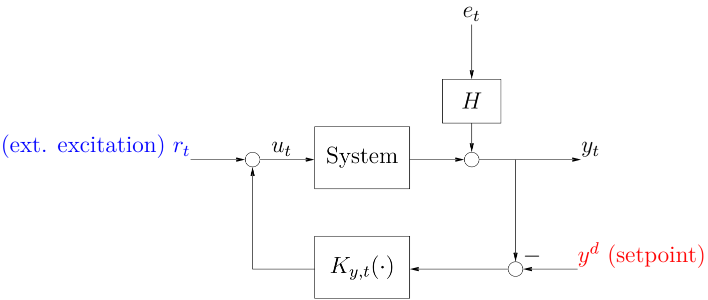
Input design based on graph theory

Extension to nonlinear SSM

Closed-loop application oriented input design

Conclusions and future work

Closed-loop application oriented input design



System ($\theta_0 \in \Theta$):

$$
\begin{aligned}
x_{t+1} &= A_{\theta_0} x_t + B_{\theta_0} u_t \\
y_t &= C_{\theta_0} x_t + \nu_t \\
\nu_t &= H(q;\, \theta_0) e_t
\end{aligned}
$$

$\{e_t\}$: white noise, known distribution.

Feedback: $u_t = r_t + K_{y,t}(y_t - y^d)$

Closed-loop application oriented input design

Model:

$$\begin{aligned} x_{t+1} &= A(\theta)x_t + B(\theta)u_t \\ y_t &= C(\theta)x_t + \nu_t \\ \nu_t &= H(q;\,\theta)e_t \end{aligned}$$

$\theta \in \Theta$.

Goal: Perform an experiment to obtain $\hat{\theta}_{n_{\mathrm{seq}}}$.
$\Rightarrow$ design $r_{1:n_{\mathrm{seq}}}$!

Requirements:

1. y_t, u_t should not be perturbed excessively.
2. $\hat{\theta}_{n_{\mathrm{seq}}}$ must satisfy quality constraints.

ROYAL INSTITUTE OF TECHNOLOGY

Closed-loop application oriented input design

Minimize control objective:

$$J = \mathbf{E}\left\{\sum_{t=1}^{n_{\text{seq}}} \left\|y_t - y^d\right\|_Q^2 + \|u_t - u_{t-1}\|_R^2\right\}$$

Requirements:

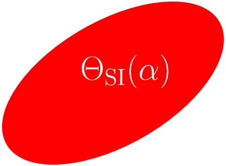
1. y_t, u_t should not be perturbed excessively.

Probabilistic bounds:

$$\mathbf{P}\{|y_t - y^d| \leq y_{\max}\} > 1 - \epsilon_y$$
$$\mathbf{P}\{|u_t| \leq u_{\max}\} > 1 - \epsilon_x$$

for $t = 1, \ldots, n_{\text{seq}}$

Closed-loop application oriented input design



Requirements:

2. $\hat{\theta}_{n_{\mathrm{seq}}}$ must satisfy quality constraints.

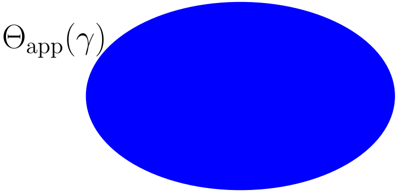
Maximum likelihood: As $n_{\mathrm{seq}} \to \infty$, we have:

$$\sqrt{n_{\mathrm{seq}}}\left(\hat{\theta}_{n_{\mathrm{seq}}} - \theta_0\right) \in \mathit{As}\mathcal{N}(0, \{\mathcal{I}_F^e\}^{-1})$$

$\Rightarrow$ Identification set:

$$\Theta_{\mathrm{SI}}(\alpha) = \left\{\theta : (\theta - \theta_0)^\top \mathcal{I}_F^e \, (\theta - \theta_0) \leq \chi^2_\alpha(n_\theta)\right\}$$

Closed-loop application oriented input design



Requirements:

2. $\hat{\theta}_{n_{\text{seq}}}$ must satisfy quality constraints.

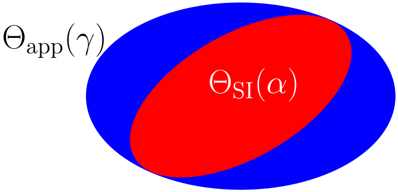
Quality constraint: Application set

$$\Theta(\gamma) = \left\{ \theta : V_{\text{app}}(\theta) \leq \frac{1}{\gamma} \right\}$$

Relaxation: Application ellipsoid

$$\Theta_{\text{app}}(\gamma) := \left\{ \theta : (\theta - \theta_0)^\top \, \nabla_\theta^2 V_{\text{app}}(\theta)\Big|_{\theta=\theta_0} (\theta - \theta_0) \leq \frac{2}{\gamma} \right\}$$

Closed-loop application oriented input design



Requirements:

2. $\hat{\theta}_{n_{\text{seq}}}$ must satisfy quality constraints.

Quality constraint:

$$\Theta_{\text{SI}}(\alpha) \subseteq \Theta_{\text{app}}(\gamma)$$

achieved by

$$\frac{1}{\chi^2_\alpha(n_\theta)} \mathcal{I}_F^e \succeq \frac{\gamma}{2} \left. \nabla^2_\theta V_{\text{app}}(\theta) \right|_{\theta=\theta_0}$$

ROYAL INSTITUTE OF TECHNOLOGY

Closed-loop application oriented input design

Optimization problem:

$$\begin{aligned}
\min_{\{r_t\}_{t=1}^{n_{\text{seq}}}} \quad & J = \mathbf{E}\left\{\sum_{t=1}^{n_{\text{seq}}} \|y_t - y_d\|_Q^2 + \|\Delta u_t\|_R^2\right\} \\
\text{s.t.} \quad & \text{System constraints} \\
& \mathbf{P}\{|y_t - y^d| \leq y_{\max}\} > 1 - \epsilon_y \\
& \mathbf{P}\{|u_t| \leq u_{\max}\} > 1 - \epsilon_x \\
& \mathcal{I}_F \succeq \frac{\gamma \chi_\alpha^2(n_\theta)}{2} \nabla_\theta^2 V_{\text{app}}(\theta)
\end{aligned}$$

- Difficulty: $\mathbf{P}$ (and $\mathcal{I}_F^e$) hard to optimize.
- Solution: Use the graph-theory approach!

Closed-loop application oriented input design

Graph theory approach:

$r_{1:n_{\text{seq}}}$ realization from $p(r_{1:n_m})$ with alphabet $\mathcal{C}$.

To design an experiment in $\mathcal{C}^{n_m}$:

1. Compute all the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$.
2. Generate the signals $\{r_t^i\}_{t=0}^{t=N}$ from the prime cycles of $\mathcal{G}_{\mathcal{C}^{n_m}}$, for each $i \in \{1, \ldots, n_{\mathcal{V}}\}$.

ROYAL INSTITUTE OF TECHNOLOGY

Closed-loop application oriented input design

Graph theory approach:

$r_{1:n_{\text{seq}}}$ realization from $p(r_{1:n_m})$ with alphabet $\mathcal{C}$.

Given $e_{1:N_{\text{sim}}}$, $r^{(j)}_{1:N_{\text{sim}}}$, approximate

$$J^{(j)} \approx \frac{1}{N_{\text{sim}}} \sum_{t=1}^{N_{\text{sim}}} \left\| y_t^{(j)} - y^d \right\|_Q^2 + \left\| u_t^{(j)} - u_{t-1}^{(j)} \right\|_R^2$$

$$\mathbf{P}_{e_t, r_t^{(j)}} \{ |u_t^{(j)}| \leq u_{\max} \} \approx \text{Monte Carlo}$$

$$\mathbf{P}_{e_t, r_t^{(j)}} \{ |y_t^{(j)} - y^d| \leq y_{\max} \} \approx \text{Monte Carlo}$$

$\mathcal{I}_F^{(j)}$ computed as in previous parts.

Closed-loop application oriented input design

Optimization problem (graph-theory):

$$
\begin{aligned}
\min_{\{\beta_1, \dots, \beta_{n_v}\}} \quad & \sum_{j=1}^{n_v} \beta_j J^{(j)} \\
\text{s.t.} \quad & \text{System constraints} \\
& \text{Constraints on } \{\beta_j\}_{j=1}^{n_v} \\
& \sum_{j=1}^{n_v} \beta_j \mathbf{P}_{e_t, r_t^{(j)}} \{|u_t^{(j)}| \leq u_{\max}\} > 1 - \epsilon_x \\
& \sum_{j=1}^{n_v} \beta_j \mathbf{P}_{e_t, r_t^{(j)}} \{|y_t^{(j)} - y^d| \leq y_{\max}\} > 1 - \epsilon_y \\
& \sum_{j=1}^{n_v} \beta_j \mathcal{I}_F^{(j)} \succeq \frac{\gamma \chi_\alpha^2(n)}{2 n_{\text{seq}}} \nabla_\theta^2 V_{\text{app}}(\theta)
\end{aligned}
$$

ROYAL INSTITUTE OF TECHNOLOGY

Closed-loop application oriented input design

Optimal pmf:

$$p^{\text{opt}} := \sum_{j=1}^{n_v} \beta_j^{\text{opt}} p_j$$

Example III

Consider the open-loop, SISO state space system

$$x_{t+1} = \theta_2^0 \, x_t + u_t$$
$$y_t = \theta_1^0 \, x_t + e_t$$

$\begin{bmatrix} \theta_1^0 & \theta_2^0 \end{bmatrix}^\top = \begin{bmatrix} 0.6 & 0.9 \end{bmatrix}^\top.$

Input:

$$u_t = r_t - k_y \, y_t$$

$k_y = 0.5$ known.

Goal: Estimate $\begin{bmatrix} \theta_1 & \theta_2 \end{bmatrix}^\top$ using indirect identification.

ROYAL INSTITUTE OF TECHNOLOGY

Example III

Design $\{r_t\}_{t=1}^{500}$, $n_m = 2$, $r_t \in \mathcal{C} = \{-0.5, -0.25, 0, 0.25, 0.5\}$.

Performance degradation:

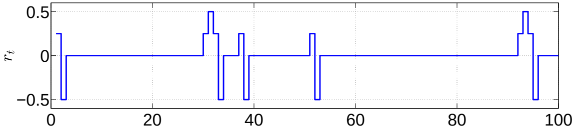
$$V_{\text{app}}(\theta) = \frac{1}{500} \sum_{t=1}^{500} \|y_t(\theta_0) - y_t(\theta)\|_2^2$$

- $y^d = 0$
- $\epsilon_y = \epsilon_x = 0.07$
- $y_{\max} = 2$, $u_{\max} = 1$

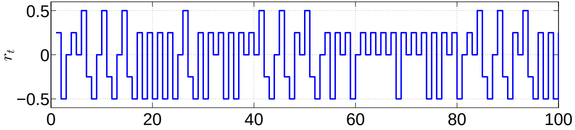
Example III

Reference:

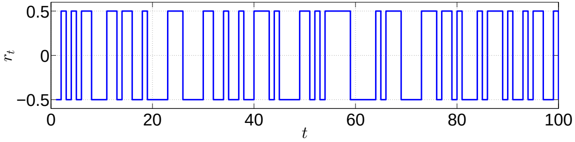
Opt. 1:



Opt. 2:



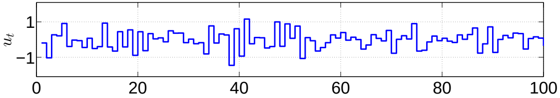
Binary:



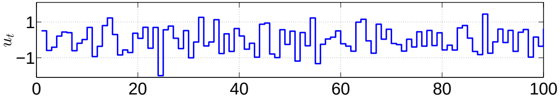
Example III

Input:

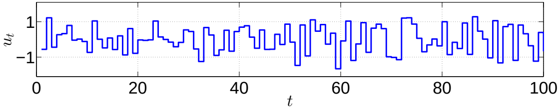
Opt. 1: 93.8%



Opt. 2: 86.6%



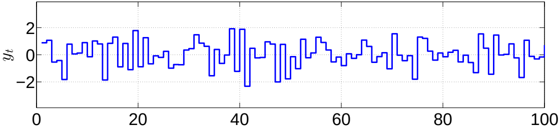
Binary: 90.8%



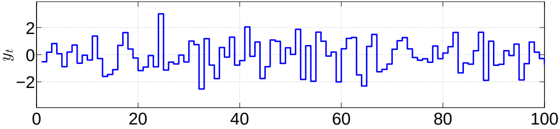
Example III

Output:

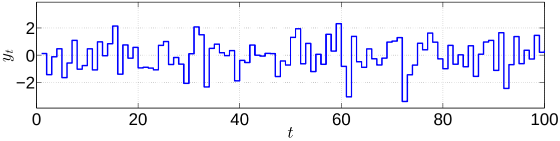
Opt. 1: 96%



Opt. 2: 93.4%

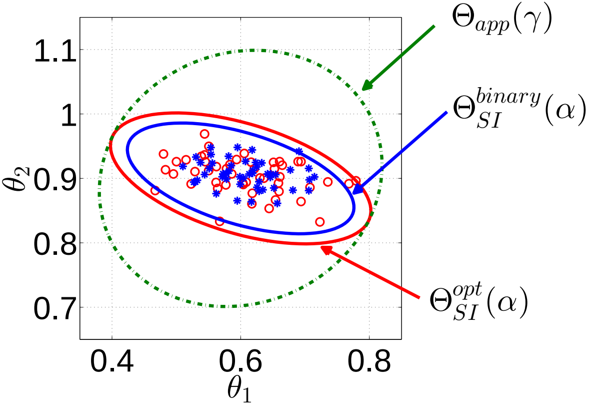


Binary: 79.6%



Example III

Application and identification ellipsoids: (98% confidence level)



ROYAL INSTITUTE OF TECHNOLOGY

ROYAL INSTITUTE OF TECHNOLOGY

Conclusions

- A new method for input design was introduced.
- The method can be used for nonlinear systems.
- Convex problem even for nonlinear systems.

Future work

- Reducible Markov chains.
- Computational complexity.
- Robust input design.
- Application oriented input design for MPC.

ROYAL INSTITUTE OF TECHNOLOGY

Thanks for your attention.

Optimal input design for nonlinear dynamical systems: a graph-theory approach

ROYAL INSTITUTE OF TECHNOLOGY

Patricio E. Valenzuela

Department of Automatic Control and ACCESS Linnaeus Centre
KTH Royal Institute of Technology, Stockholm, Sweden

Seminar Uppsala university
January 16, 2015

Outline

Problem formulation for output-error models

Input design based on graph theory

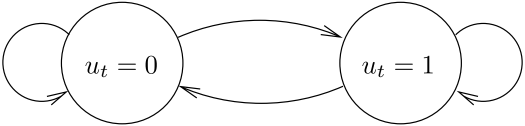
Extension to nonlinear SSM

Closed-loop application oriented input design

Conclusions and future work

Appendix I: Graph theory in input design

Example: Elementary cycles for a de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 1$.



One elementary cycle: $z = (0, 1, 0)$.

$\Rightarrow$ Prime cycles for a de Bruijn Graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$: $v_1 = ((0, 1), (1, 0), (0, 1))$.

Appendix II: Graph theory in input design

Example: Generation of input signal from a prime cycle.

Consider a de Bruijn graph, $\mathcal{C} := \{0, 1\}$, $n_m := 2$.

- $v_1 = ((0, 1), (1, 0), (0, 1))$
- $\{u_t^1\}_{t=0}^{t=N}$: Take last element of each node.

Finally,

$$\{u_t^i\}_{t=0}^{t=N} = \{1, 0, 1, 0, \ldots, ((-1)^N + 1)/2\}$$

ROYAL INSTITUTE OF TECHNOLOGY

Appendix III: Building A

- For $i \in \mathcal{C}^{n_m}$, define

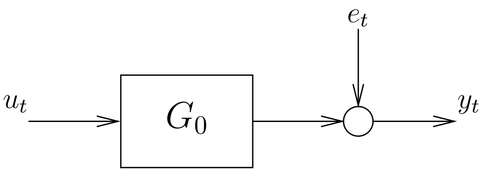
$$\mathcal{A}_i := \{j \in \mathcal{C}^{n_m} : (j, i) \in \mathcal{E}\}.$$

(the set of ancestors of i).

- For each $i \in \mathcal{C}^{n_m}$, let

$$A_{ij} = \begin{cases} \dfrac{\mathbf{P}\{i\}}{\sum\limits_{k \in \mathcal{A}_i} \mathbf{P}\{k\}}, & \text{if } j \in \mathcal{A}_i \text{ and } \sum\limits_{k \in \mathcal{A}_i} \mathbf{P}\{k\} \neq 0 \\ \dfrac{1}{\# \mathcal{A}_i}, & \text{if } j \in \mathcal{A}_i \text{ and } \sum\limits_{k \in \mathcal{A}_i} \mathbf{P}\{k\} = 0 \\ 0, & \text{otherwise.} \end{cases}$$

Apendix IV: Example nonlinear case



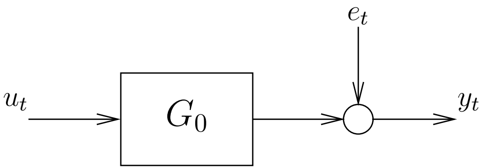
$$G_0(u_t) = G_1(q,\theta)\, u_t + G_2(q,\theta)\, u_t^2$$

where

$$G_1(q,\theta) = \theta_1 + \theta_2\, q^{-1}$$
$$G_2(q,\theta) = \theta_3 + \theta_4\, q^{-1}$$

e_t: Gaussian white noise, zero mean, variance $\lambda_e = 1$.

Apendix IV: Example nonlinear case



$$G_0(u_t) = G_1(q,\theta)\, u_t + G_2(q,\theta)\, u_t^2$$

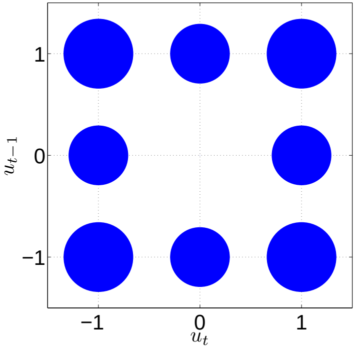
where

$$G_1(q,\theta) = \theta_1 + \theta_2\, q^{-1}$$
$$G_2(q,\theta) = \theta_3 + \theta_4\, q^{-1}$$

We consider $h(\cdot) = \det(\cdot)$, $c_{\text{seq}} = 3$, $n_m = 2$, $\mathcal{C} := \{-1, 0, 1\}$, and $N = 5 \cdot 10^3$.

Apendix IV: Example nonlinear case

Stationary probabilities:



- $\det(\mathcal{I}_F^{\text{app}}) = 0.1796.$
- Results consistent with previous contributions (Larsson et al. 2010).

Appendix V: Particle methods to estimate $\mathcal{S}_m(\theta_0)$

Auxiliary particle filter:

$$\hat{p}_\theta(x_{1:t}|y_{1:t}) := \sum_{i=1}^{N} \frac{w_t^{(i)}}{\sum_{k=1}^{N} w_t^{(k)}} \delta(x_{1:t} - x_{1:t}^{(i)})$$

$\{x_{1:t}^{(i)}, w_t^{(i)}\}_{i=1}^N$: Particle system.

Two step procedure to compute $\{x_{1:t}^{(i)}, w_t^{(i)}\}_{i=1}^N$:

1. Sampling/propagation.

2. Weighting.

Appendix V: Particle methods to estimate $\mathcal{S}_m(\theta_0)$

ROYAL INSTITUTE OF TECHNOLOGY

Two step procedure to compute $\{x_{1:t}^{(i)},\, w_t^{(i)}\}_{i=1}^N$:

1. Sampling/propagation:

$$\{a_t^{(i)},\, x_t^{(i)}\} \sim \frac{w_{t-1}^{a_t}}{\sum_{k=1}^N w_{t-1}^{(k)}} R_{\theta,t}(x_t | x_{t-1}^{a_t}, u_{t-1})$$

Appendix V: Particle methods to estimate $\mathcal{S}_m(\theta_0)$

Two step procedure to compute $\{x_{1:t}^{(i)},\, w_t^{(i)}\}_{i=1}^N$:

2. Weighting:

$$w_t^{(i)} := \frac{g_\theta(y_t|x_t^{(i)}, u_t)\, f_\theta(x_t|x_{t-1}^{(i)}, u_{t-1})}{R_{\theta,t}(x_t|x_{t-1}^{(i)}, u_{t-1})}$$

ROYAL INSTITUTE OF TECHNOLOGY

Appendix V: Particle methods to estimate $\mathcal{S}_m(\theta_0)$

- Difficulty: Auxiliary particle filter suffers particle degeneracy.

- Solution: Use fixed-lag smoother.

Main idea FL-smoother:

$$p_\theta(x_t|y_{1:T}, u_{1:T}) \approx p_\theta(x_t|y_{1:\kappa_t}, u_{1:\kappa_t})$$

for $\kappa_t = \min(t + \Delta, T)$, for some fixed-lag $\Delta > 0$.

Apendix VI: Equivalence time and frequency domain

	Frequency	Time
Design variable	$\Phi_u(\omega)$	$p(u_{1:n_m})$
Restrictions	$\Phi_u(\omega) \geq 0$ $\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_u(\omega) d\omega \leq 1$	$p \geq 0$ $\sum_{u_{1:n_m}} p(u_{1:n_m}) = 1$ p stationary
Information matrix	$\mathcal{I}_F(\Phi_u)$	$\mathcal{I}_F(p)$