Westfälische Wilhelms-Universität Münster

Applied Mathematics Münster

Yet Another Talk About Empirical Gramians

Christian Himpe (christian.himpe@uni-muenster.de)
Mario Ohlberger (mario.ohlberger@uni-muenster.de)

WWU Münster
Institute for Computational and Applied Mathematics

Reduced Basis Summer School 2014

Outline

- (Mathematical) Model
- Gramian-Based Model Reduction
- Empirical Gramians
- State Reduction
- Parameter Reduction
- Combined Reduction

Linear Control System

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$
$$x(0) = x_0$$

Input / Control:	$u \in \mathbb{R}^M$
State:	$x \in \mathbb{R}^N$
Output:	$y \in \mathbb{R}^O$
System Matrix:	$A \in \mathbb{R}^{N \times N}$
Input Matrix:	$B \in \mathbb{R}^{N \times M}$
Output Matrix:	$C \in \mathbb{R}^{O \times N}$

Adjoint System

$$\dot{x}^*(t) = A^T x^*(t) + C^T u^*(t)$$
$$y^*(t) = B^T x^*(t)$$
$$x^*(0) = x_0^*$$

Input / Control:	$u^* \in \mathbb{R}^M$
State:	$x^* \in \mathbb{R}^N$
Output:	$y^* \in \mathbb{R}^O$
System Matrix:	$A^T \in \mathbb{R}^{N \times N}$
Input Matrix:	$B^T \in \mathbb{R}^{M \times N}$
Output Matrix:	$C^T \in \mathbb{R}^{N \times O}$

Model Reduction

Large-Scale Control System:

$$\dot{x} = Ax + Bu$$
$$y = Cx$$
$$x(0) = x_0$$

Have:

- $\dim(x) \gg 1$
- $\dim(u) \ll \dim(x)$
- $\dim(y) \ll \dim(x)$

Reduced Order System:

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u$$
$$\tilde{y} = \tilde{C}\tilde{x}$$
$$\tilde{x}(0) = \tilde{x}_0$$

Want:

- $\dim(\tilde{x}) \ll \dim(x)$
- $\|y - \tilde{y}\|_? \ll 1$

Controllability & Observability

Controllability:
Quantifies how well a state is driven by the input.

Controllability Operator:

$$\mathcal{C}(u) := \int_0^{\infty} e^{At} B u(t) \mathrm{d}t$$

Observability:
Quantifies how well a change in state is visible in the output.

Observability Operator:

$$\mathcal{O}(x_0) := C e^{At} x_0$$

Controllability Gramian & Observability Gramian

Controllability Gramian:

$$
\begin{aligned}
W_C &:= \mathcal{C}\mathcal{C}^T \\
&= \int_0^\infty e^{At} B B^T e^{A^T t} \mathrm{d}t \\
&\Rightarrow A W_C + W_C A^T = -B B^T
\end{aligned}
$$

Observability Gramian:

$$
\begin{aligned}
W_O &:= \mathcal{O}^T \mathcal{O} \\
&= \int_0^\infty e^{A^T t} C^T C e^{At} \mathrm{d}t \\
&\Rightarrow W_O A^T + A W_O = -C^T C
\end{aligned}
$$

- Controllability and observability are dual operators.
- The adjoints system controllability is the systems observability.

Balanced Truncation [Moore'81]

Balance Controllability and Observability:

$$W_O \overset{Cholesky}{=} L_O L_O^T$$

$$W_C \overset{Cholesky}{=} L_C L_C^T$$

$$L_C L_O^T \overset{SVD}{=} UDV^T$$

Partition and Truncate:

$$\sigma_1 > \ldots > \sigma_N$$

$$\rightarrow U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix},\ V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}^T$$

Reduced Order Model:

$$\tilde{A} = V_1 A U_1$$

$$\tilde{B} = V_1 B$$

$$\tilde{C} = C U_1$$

$$\tilde{x}_0 = V_1 x_0$$

Cross Gramian [Fernando & Nicholson'83]

Cross Gramian:

$$
\begin{aligned}
W_X &:= \mathcal{C}\mathcal{O} \\
&= \int_0^\infty e^{At} B C e^{At} \mathrm{d}t \\
&\Rightarrow A W_X + W_X A = -BC
\end{aligned}
$$

- System must be square: $\dim(u) = \dim(y)$
- If system is symmetric: $Ce^A B = (Ce^A B)^T \Rightarrow W_X^2 = W_C W_O$

Direct Truncation

Balance Controllability and Observability:

$$W_X \overset{SVD}{=} UDV^T$$

Partition and Truncate:

$$\sigma_1 > \ldots > \sigma_N$$
$$\rightarrow U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$$

Reduced Order Model:

$$\tilde{A} = U_1^T A U_1$$
$$\tilde{B} = U_1^T B$$
$$\tilde{C} = C U_1$$
$$\tilde{x}_0 = U_1^T x_0$$

Notes on the Cross Gramian

Advantages:

- Compute only 1 instead of 2 gramian matrices
- Significantly less expensive balancing
- Conveys additional information (System Gain, Cauchy Index)

Disadvantages:

- Only sqare systems
- Actually, only symmetric systems
- Without extra effort only one-sided projections

Empirical Gramians [Lall'99]

System Gramians:

$$W_C = \int_0^\infty e^{At} B\, B^T e^{A^T t} \mathrm{d}t$$
$$W_O = \int_0^\infty e^{A^T t} C^T\, C e^{At} \mathrm{d}t$$
$$W_X = \int_0^\infty e^{At} B\, C e^{At} \mathrm{d}t$$

Note:

- Integrands are just Impulse Responses!
- Can be computed empirically.
- Equal to classic gramians for linear systems.
- Extend to nonlinear systems.

Empirical Gramians [Lall'99]

System Gramians:

$$W_C = \int_0^\infty (e^{At}B)\,(e^{At}B)^T \mathrm{d}t$$
$$W_O = \int_0^\infty (e^{A^Tt}C^T)\,(e^{A^Tt}C^T)^T \mathrm{d}t$$
$$W_X = \int_0^\infty (e^{At}B)\,(e^{A^Tt}C^T)^T \mathrm{d}t$$

Note:

- Integrands are just Impulse Responses!
- Can be computed empirically.
- Equal to classic gramians for linear systems.
- Extend to nonlinear systems.

Empirical Gramians [Lall'99]

Empirical Gramians:

$$W_C = \int_0^\infty x(t)x(t)^T \mathrm{d}t$$
$$W_O = \int_0^\infty x^*(t)x^*(t)^T \mathrm{d}t$$
$$W_X = \int_0^\infty x(t)x^*(t)^T \mathrm{d}t$$

Note:

- Integrands are just Impulse Responses!
- Can be computed empirically.
- Equal to classic gramians for linear systems.
- Extend to nonlinear systems.

Averaged Gramians

Empirical Gramians:

$$W_C = \frac{1}{J} \sum_{j}^{J} \int_0^\infty x_{(j)}(t) x_{(j)}(t)^T \mathrm{d}t$$

$$W_O = \frac{1}{J} \sum_{j}^{J} \int_0^\infty x^*_{(j)}(t) x^*_{(j)}(t)^T \mathrm{d}t$$

$$W_X = \frac{1}{J} \sum_{j}^{J} \int_0^\infty x_{(j)}(t) x^*_{(j)}(t)^T \mathrm{d}t$$

Note:

- Controllability: varying impulse inputs
- Observability: varying initial states
- Ensure ROM fits to operating region of system
- Especially for nonlinear systems

Perturbation Sets

Impulse Input Perturbations: $u^{hij}(t) = c_h S_i e_j \delta(t)$

$$E_u = \{e_i \in \mathbb{R}^m; \|e_i\| = 1; e_i e_{j \neq i} = 0; i = 1, \ldots, m\},$$
$$R_u = \{S_i \in \mathbb{R}^{m \times m}; S_i^T S_i = \mathbb{1}; i = 1, \ldots, s\},$$
$$Q_u = \{c_i \in \mathbb{R}; c_i > 0; i = 1, \ldots, q\},$$

Initial State Perturbation: $x_0^{kla} = d_k T_l f_a$

$$E_x = \{f_i \in \mathbb{R}^n; \|f_i\| = 1; f_i f_{j \neq i} = 0; i = 1, \ldots, n\},$$
$$R_x = \{T_i \in \mathbb{R}^{n \times n}; T_i^T T_i = \mathbb{1}; i = 1, \ldots, t\},$$
$$Q_x = \{d_i \in \mathbb{R}; d_i > 0; i = 1, \ldots, r\}.$$

Empirical Controllability Gramian

Analytic Controllability Gramian

$$W_C = \int_0^\infty e^{At} B B^T e^{A^T t} \mathrm{d}t$$

Linear Empirical Controllability Gramian

$$W'_C = \int_0^\infty x(t)\, x^T(t) \mathrm{d}t$$

Empirical Controllability Gramian

$$W''_C = \frac{1}{|Q_u||R_u|} \sum_{h=1}^{|Q_u|} \sum_{i=1}^{|R_u|} \sum_{j=1}^{m} \frac{1}{c_h^2} \int_0^\infty \Psi^{hij}(t) \mathrm{d}t,$$

$$\Psi^{hij}(t) = (x^{hij}(t) - \bar{x})(x^{hij}(t) - \bar{x})^T \in \mathbb{R}^{n \times n}$$

Discrete Empirical Controllability Gramian

$$W'''_C = \frac{1}{|Q_u||R_u|} \sum_{h=1}^{|Q_u|} \sum_{i=1}^{|R_u|} \sum_{j=1}^{m} \frac{\Delta t}{c_h^2} \sum_{t=0}^{\mathfrak{T}} \Psi_t^{hij},$$

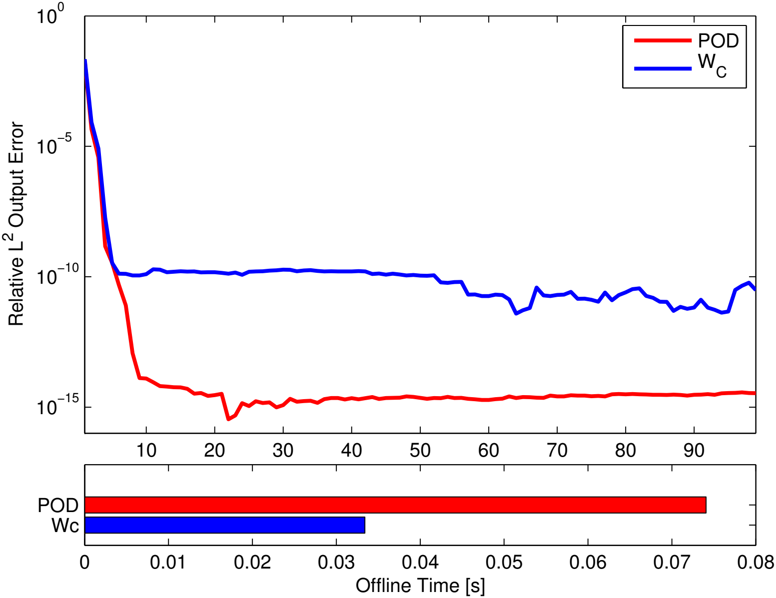
$$\Psi_t^{hij} = (x_t^{hij} - \bar{x})(x_t^{hij} - \bar{x})^T \in \mathbb{R}^{n \times n}$$

Test System

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$
$$x(0) = 0$$
$$u(t) = \delta(t)$$

- A random but stable ($\mathrm{Re}(\lambda_i(A)) < 0, \ \forall i$)
- B, C random

POD vs W_C



Relative L^2 Output Error
10^0
10^{-5}
10^{-10}
10^{-15}
10
20
30
40
50
60
70
80
90
POD
W_C
POD
Wc
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Offline Time [s]

Empirical Observability Gramian

Analytic Observability Gramian

$$W_O = \int_0^\infty e^{A^T t} C^T C e^{At} \mathrm{d}t$$

Linear Empirical Observability Gramian

$$W_O' = \int_0^\infty x^*(t)\, x^*(t)^T \mathrm{d}t$$

Empirical Observability Gramian

$$W_O'' = \frac{1}{|Q_x||R_x|} \sum_{k=1}^{|Q_x|} \sum_{l=1}^{|R_x|} \frac{1}{d_k^2} \int_0^\infty T_l \Psi^{kl}(t) T_l^T \mathrm{d}t,$$

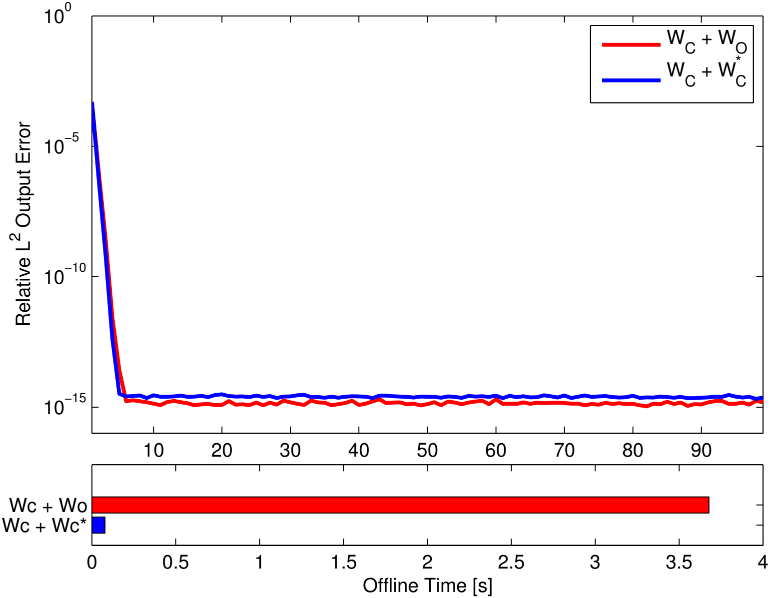
$$\Psi_{ab}^{kl} = (y^{kla}(t) - \bar{y})^T (y^{klb}(t) - \bar{y}) \in \mathbb{R}$$

Discrete Empirical Observability Gramian

$$W_O''' = \frac{1}{|Q_x||R_x|} \sum_{k=1}^{|Q_x|} \sum_{l=1}^{|R_x|} \frac{\Delta t}{d_k^2} \sum_{t=0}^{\mathfrak{T}} T_l \Psi_t kl T_l^T,$$

$$\Psi_{t,ab}^{kl} = (y_t^{kla} - \bar{y})^T (y_t^{klb} - \bar{y}) \in \mathbb{R}$$

BT vs linear BT



Empirical Cross Gramian [Streif'06, H.'14]

Analytic Cross Gramian

$$W_X = \int_0^\infty e^{At} B C e^{At} \mathrm{d}t$$

Linear Empirical Cross Gramian

$$W_X' = \int_0^T x(t) x^*(t)^T \mathrm{d}t$$

Empirical Cross Gramian

$$W_X'' = \frac{1}{|Q_u||R_u|m|Q_x||R_x|} \sum_{h=1}^{|Q_{\boldsymbol{u}}|} \sum_{i=1}^{|R_{\boldsymbol{u}}|} \sum_{j=1}^{m} \sum_{k=1}^{|Q_{\boldsymbol{x}}|} \sum_{l=1}^{|R_{\boldsymbol{x}}|} \frac{1}{c_h d_k} \int_0^\infty T_l \Psi^{hijkl}(t) T_l^T \mathrm{d}t,$$

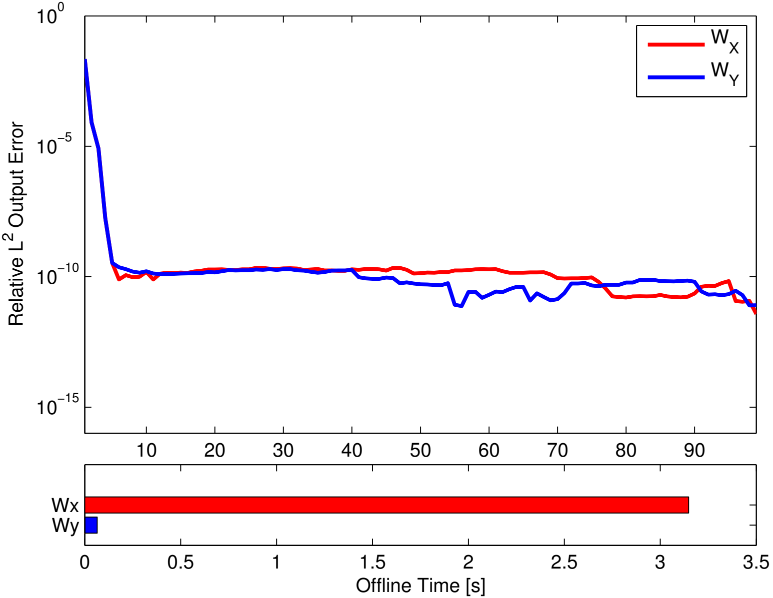
$$\Psi_{ab}^{hijkl}(t) = f_b^T T_k^T (x^{hij}(t) - \bar{x}) e_i^T S_h^T (y^{kla}(t) - \bar{y}) \in \mathbb{R}$$

Discrete Empirical Cross Gramian

$$W_X''' = \frac{1}{|Q_u||R_u|m|Q_x||R_x|} \sum_{h=1}^{|Q_{\boldsymbol{u}}|} \sum_{i=1}^{|R_{\boldsymbol{u}}|} \sum_{j=1}^{m} \sum_{k=1}^{|Q_{\boldsymbol{x}}|} \sum_{l=1}^{|R_{\boldsymbol{x}}|} \frac{\Delta t}{c_h d_k} \sum_{t=0}^{\mathfrak{T}} T_l \Psi_t^{hijkl} T_l^T,$$

$$\Psi_{t,ab}^{hijkl} = f_b^T T_k^T (x_t^{hij} - \bar{x}) e_i^T S_h^T (y_t^{kla} - \bar{y}) \in \mathbb{R}$$

Cross Gramian vs Linear Cross Gramian



General Control Systems

$$\dot{x}(t) = f(x(t), u(t))$$
$$y(t) = g(x(t), u(t))$$
$$x(0) = x_0$$

Input / Control:	$u \in \mathbb{R}^M$
State:	$x \in \mathbb{R}^N$
Output:	$y \in \mathbb{R}^O$
Vector Field:	$f : \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R} \to \mathbb{R}^N$
Output Functional:	$g : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}^O$

Nonlinear Balancing

Balanced Truncation:

$$W_O \stackrel{Cholesky}{=} L_O L_O^T$$

$$W_C \stackrel{Cholesky}{=} L_C L_C^T$$

$$L_C L_O^T \stackrel{SVD}{=} UDV^T$$

$$\sigma_1 > \ldots > \sigma_N$$

$$\rightarrow U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix},\ V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}^T$$

$$\tilde{f} = V_1 f(U_1 x, u, \theta)$$

$$\tilde{g} = g(U_1 x, u, \theta)$$

$$\tilde{x}_0 = V_1 x_0$$

Direct Truncation:

$$W_X \stackrel{SVD}{=} UDV^T$$

$$\sigma_1 > \ldots > \sigma_N$$

$$\rightarrow U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$$

$$\tilde{f} = U_1^T f(U_1 x, u, \theta)$$

$$\tilde{g} = g(U_1 x, u, \theta)$$

$$\tilde{x}_0 = U_1^T x_0$$

Parametrized Control Systems

$$\dot{x}(t) = f(x(t), u(t), \theta)$$
$$y(t) = g(x(t), u(t), \theta)$$
$$x(0) = x_0$$

Input / Control:	$u \in \mathbb{R}^M$
State:	$x \in \mathbb{R}^N$
Output:	$y \in \mathbb{R}^O$
Vector Field:	$f : \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R} \to \mathbb{R}^N$
Output Functional:	$g : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}^O$
Parameter:	$\theta \in \mathbb{R}^P$

Parametric Model Order Reduction (pMOR)

- That's the "easy" part!
- Just average the empirical gramians over a discretized parameter space:

$$W_C = \frac{1}{J} \sum_j^J W_C(\theta_j)$$

$$W_O = \frac{1}{J} \sum_j^J W_O(\theta_j)$$

$$W_X = \frac{1}{J} \sum_j^J W_X(\theta_j)$$

Parameter Identification & Parameter Reduction

Parameter Identification:

$$W_\theta \overset{SVD}{=} UDV^T$$

Partition and Truncate:

$$\sigma_1 > \ldots > \sigma_N$$
$$\rightarrow U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$$

Parameter Reduction:

$$\tilde{\theta} = U_1^T \theta$$
$$\theta \approx U_1 \tilde{\theta}$$

Controllability-Based [Sun'06, H.'13]

Parameter Decomposed System:

$$\dot{x} = f(x, u) + \sum_{k=1}^{P} f(x, \theta_k)$$
$$\Rightarrow W_C = W_{C,0} + \sum_{k=1}^{P} W_{C,k}$$

Sensitivity Gramian:

$$W_{S,ii} = \operatorname{tr}(W_{C,i})$$

Observability-Based [Geffen'08]

Parameter Augmented System:

$$\begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} f(x,u,\theta) \\ 0 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix}$$

$$y = \begin{pmatrix} g(x,u,\theta) & 0 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix}$$

$$\begin{pmatrix} x(0) \\ \theta(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ \theta_0 \end{pmatrix}$$

Augmented Observability Gramian:

$$\widetilde{W_O} = \begin{pmatrix} W_O & W_M \\ W_M^T & W_P \end{pmatrix}$$

Identifiability Gramian:
(Schur Complement of Augmented Observability Gramian)

$$W_I := W_P - W_M W_O^{-1} W_M^T \approx W_P$$

Cross-Gramian-Based [H.'14]

Parameter Augmented System:

$$\begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} f(x,u,\theta) \\ 0 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix}$$

$$y = \begin{pmatrix} g(x,u,\theta) & 0 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix}$$

$$\begin{pmatrix} x(0) \\ \theta(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ \theta_0 \end{pmatrix}$$

Joint Gramian (Augmented Cross Gramian):

$$W_J = \begin{pmatrix} W_X & W_M \\ 0 & 0 \end{pmatrix}$$

Cross-Identifiability Gramian:
(Schur Complement of Symmetric Part of Joint Gramian)

$$W_{\tilde{I}} := -\frac{1}{2} W_M (W_X + W_X^T)^{-1} W_M^T$$

Notes on the Joint Gramian

Advantages:

- One gramian to rule them all!
- Can be assembled HPC friendly column-wise
- Computes faster than augmented observability gramian
- Less memory intensive than identifiability gramian
- It is just a cross gramian

Disadvantages:

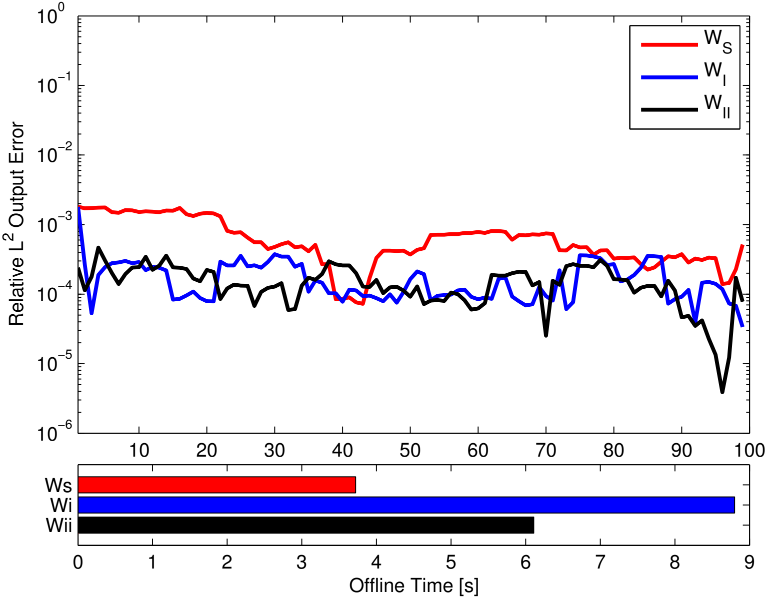
- Only sqare systems
- Actually, only symmetric systems

Test System

$$
\begin{aligned}
\dot{x}(t) &= A(\theta)x(t) + Bu(t) \\
y(t) &= Cx(t) \\
x(0) &= 0 \\
u(t) &= \delta(t)
\end{aligned}
$$

- same as before,
- but with parametrized diagonal of A.

WS vs WI vs WJ

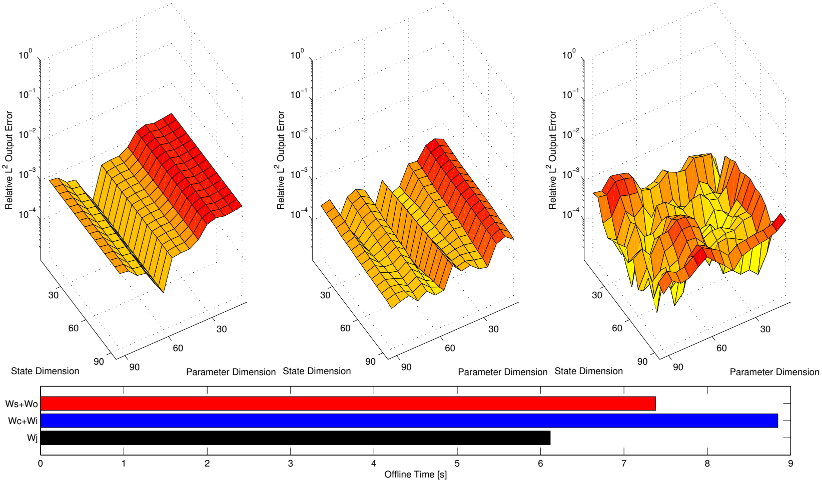


Relative L^2 Output Error
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10 20 30 40 50 60 70 80 90 100
W_S
W_I
W_{II}
Ws
Wi
Wii
0 1 2 3 4 5 6 7 8 9
Offline Time [s]

Combined State and Parameter Reduction

$W_S + W_O$	$W_C + W_I$	W_J
1 Compute W_S	1 Compute W_I	1 Compute $W_{\ddot{I}}$
2 SVD(W_S)	2 SVD(W_I)	2 SVD($W_{\ddot{I}}$)
3 Truncate θ	3 Truncate θ	3 Truncate θ
4 (Extract W_C)	4 (Extract W_O)	4 (Extract W_X)
5 Compute W_O	5 Compute W_C	
6 Balance W_C, W_O	6 Balance W_C, W_O	5 SVD(W_X)
7 Truncate x	7 Truncate x	6 Truncate x

WS+WO vs WC+WI vs WJ



Notes on the SVD

- SVD is expensive!
- Sparse SVD is better than full
- Matlab implementation: eigs$\begin{pmatrix} A & 0 \\ 0 & A^T \end{pmatrix}$ ←Yuck!
- Try Lanczos procedure!
- Make sure to orthogonalize the "singular vectors"
- For example by post-processing with a QR-decomposition

Outlook

- Better $\mathcal{L}^2$ approximation through "Balanced Gains"?
- Nonsymmetric Cross-Gramian
  - The problem reduces to compute a cheap symmetrizer
- Nonsquare Cross-Gramian
  - The problem reduces to computing a cheap adapter matrix
- Extreme-Scale SVD
  - Slicing works best with Cross Gramian
  - Lanczos for slices

tl;dl

- (Combined) State & Parameter Reduction
- for Linear & Nonlinear Systems
- using Empirical Gramians

`http://gramian.de`

Thanks!

Get the Companion Code: `http://j.mp/rbss14`