Westfälische Wilhelms-Universität Münster

Applied Mathematics Münster

The Distributed Empirical Cross Gramian

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

WWU Münster
Institute for Computational and Applied Mathematics

Data-Driven Model Order Reduction and Machine Learning (MORML)
31.03.2016

Outline

Cross Gramian	a tool for model reduction
Empirical	based on experiments, data-driven
Distributed	a parallel approach

Motivation

Application:

- Neurophysiological
- Parameter Inference (ie.: Network Connectivity)
- Functional Neuroimaging Data

Model:

- Nonlinear
- Parametric
- High-Dimensional State- **and** Parameter-Space

Notation (I)

Linear (Time-Invariant) System:

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

- Input: $u(t)$
- State: $x(t)$
- Output: $y(t)$

- $M := \dim(u(t))$
- $N := \dim(x(t))$
- $Q := \dim(y(t))$

- System Matrix: $A \in \mathbb{R}^{N \times N}$
- Input Matrix: $B \in \mathbb{R}^{N \times M}$
- Output Matrix: $C \in \mathbb{R}^{Q \times N}$

Notation (II)

General Input-Output System:

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

- Parameter: $\theta \in \mathbb{R}^P$
- Vector Field: $f : \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^P \to \mathbb{R}^N$
- Output Functional: $g : \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^P \to \mathbb{R}^Q$

Model Order Reduction (MOR)

Relevant Input-Output Mapping:

$$u \mapsto y$$

Actual Input-Output Mapping:

$$u \mapsto x \mapsto y$$

Reduction Rationale:

- $N \gg 1$
- $M \ll N$
- $Q \ll N$

Reduced Order Model (ROM)

Reduced Order System:

$$\dot{x}_r(t) = f_r(x_r(t), u(t), \theta)$$
$$y_r(t) = g_r(x_r(t), u(t), \theta)$$
$$x_r(0) = x_{0,r}$$

- $n := \dim(x_r(t))$
- $n \ll N$
- $\|y(\theta) - y_r(\theta)\| \ll 1$

Projection-Based Model Reduction

Galerkin Projection U:

$$\begin{aligned} \text{rank}(U) &= n \\ UU &= U \\ U^\intercal U &= \mathbb{1} \end{aligned}$$

Projection-Based ROM:

$$\begin{aligned} \dot{x}_r(t) &= U^\intercal f(Ux_r(t), u(t), \theta) \\ y_r(t) &= g(Ux_r(t), u(t), \theta) \\ x_r(0) &= U^\intercal x_0 \end{aligned}$$

Cross Gramian[1] [Fernando & Nicholson'83, Laub et al.'83]

For square systems ($M = Q$):

$$
\begin{aligned}
W_X &:= \mathcal{C} \circ \mathcal{O} \\
&= \int_0^\infty e^{At} BC \, e^{At} \, \mathrm{d}t \\
&\overset{\text{part. Int.}}{\Rightarrow} \; AW_X + W_X A = -BC
\end{aligned}
$$

Special properties:

- state-space symmetric systems
- (orthogonally) symmetric systems

[1] The cross Gramian is not a Gramian matrix!

Direct Truncation

Hankel operator:

$$H := \mathcal{O} \circ \mathcal{C}$$

For symmetric systems:

$$\begin{aligned} H &= H^* \\ \Rightarrow \mathcal{O}\mathcal{C} &= (\mathcal{O}\mathcal{C})^* \\ \Rightarrow \sigma_i(H) &= \sqrt{\lambda_i((\mathcal{O}\mathcal{C})^*\mathcal{O}\mathcal{C})} \\ &= \sqrt{\lambda_i(\mathcal{O}\mathcal{C}\mathcal{O}\mathcal{C})} \\ &= \sqrt{\lambda_i(\mathcal{C}\mathcal{O}\mathcal{C}\mathcal{O})} \\ &= \sqrt{\lambda_i(W_X W_X)} = |\lambda_i(W_X)| \end{aligned}$$

Approximate Hankel Singular Value Decomposition:

$$W_X \overset{SVD}{=} UDV \rightarrow U = \begin{pmatrix} U_1 & U_2 \end{pmatrix}$$

Domains of Application

- Model Reduction [Sorensen & Antoulas'02]
- System Indices:
  - Minimality Test [Fernando & Nicholson'82]
  - System Gain [Gheondea & Ober'99]
  - Singularity Index [Mironovskii & Soloveva'15]
  - Cauchy Index [Fernando & Nicholson'83]
  - Minimum Information Loss Index [Fu et al.'09]
- Sensitivity Analysis [Streif et al.'06, Streif et al.'09]
- Parameter Identification [H. & Ohlberger'14]
- Decentralized Control [Moaveni & Khaki-Sedigh'06]
- and more ...

Empirical Linear Cross Gramian[2] [Fernando & Nicholson'85, Shaker'12]

$$
\begin{aligned}
W_X &= \int_0^\infty e^{At} BC \, e^{At} \, \mathrm{d}t \\
&= \int_0^\infty (e^{At} B)(e^{A^\intercal t} C^\intercal)^\intercal \mathrm{d}t \\
&= \int_0^\infty x(t) z(t)^\intercal \mathrm{d}t
\end{aligned}
$$

- $x(t)$ Impulse Response
- $z(t)$ Adjoint Impulse Response

---

[2] See also: [Moore'81, Lall et al.'99, Lall et al.'02]

Empirical Cross Gramian [Streif et al.'06, H. & Ohlberger'14]

$$\widehat{W}_X = \sum_{m=1}^{M} \int_0^\infty \Psi^m(t)\mathrm{d}t \in \mathbb{R}^{N \times N}$$

$$\Psi_{ij}^m(t) = \langle x_i^m(t), y_m^j(t) \rangle$$

- $x_i^m(t)$ – i-th state component for m-th perturbed input
- $y_m^j(t)$ – m-th output component for j-th perturbed initial state
- For linear systems: $\widehat{W}_X = W_X$
- For parametric systems[3]: $\widehat{W}_X = \sum_{\theta_i \in \Theta_h} \widehat{W}_X(\theta_i)$

---

[3] Based on [Sun & Hahn'06]

Combined State and Parameter Reduction

Reduced Order System:

$$\begin{aligned}\dot{x}_r(t) &= f_r(x_r(t), u(t), \theta_r)\\ y_r(t) &= g_r(x_r(t), u(t), \theta_r)\\ x_r(0) &= x_{0,r}\end{aligned}$$

- $p := \dim(\theta_r)$
- $p \ll P$
- $\|y(\theta) - y_r(\theta_r)\| \ll 1$

Empirical Joint Gramian [Geffen et al.'08, H. & Ohlberger'14]

Augmented System:

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

$$y(t) = g(x(t), u(t), \theta)$$

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

Joint Gramian (Cross Gramian of the Augmented System):

$$W_J = \begin{pmatrix} W_X & W_M \\ W_m & W_\theta \end{pmatrix}$$

Uncontrollable Parameters:

$$W_m = 0$$

$$W_\theta = 0$$

Empirical Cross-Identifiability Gramian [H. & Ohlberger'14]

Schur-Complement of W_θ (Cross-Identifiability Gramian):

$$W_{\ddot{I}} := 0 - \frac{1}{2} W_M^\intercal (W_X + W_X^\intercal)^+ W_M$$

$W_{\ddot{I}}$ encodes the "observability" of parameters.

Parameter Projection as Principal Components:

$$W_{\ddot{I}} \overset{SVD}{=} \Pi \Delta \Lambda \to \Pi = \begin{pmatrix} \Pi_1 & \Pi_2 \end{pmatrix}$$

Cross-Gramian-Based Combined Reduction

State-space projection:

$$W_X \overset{TSVD}{=} U_1 D_1 V_1$$

Parameter-space projection:

$$W_{\ddot{I}} \overset{TSVD}{=} \Pi_1 \Delta_1 \Lambda_1$$

Combined state and parameter ROM:

$$\begin{aligned} \dot{x}_r(t) &= U_1^\intercal f(U_1 x_r(t), u(t), \Pi_1 \theta_r), \\ y_r(t) &= g(U_1 x_r(t), u(t), \Pi_1 \theta_r), \\ x_r(0) &= U_1^\intercal x_0, \\ \theta_r &= \Pi_1^\intercal \theta \end{aligned}$$

Nonlinear Test System[4]

Weakly Nonlinear SISO System:

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

System Dimensions:

- $\dim(u(t)) = 1$
- $\dim(x(t)) = 10^2$
- $\dim(y(t)) = 1$
- $\dim(\theta) = 10^2$

[4] See: Hyperbolic Network Model [Quan et al.'01]

emgr – Empirical Gramian Framework (Version: 3.9, 02/2016)

Empirical Gramians:

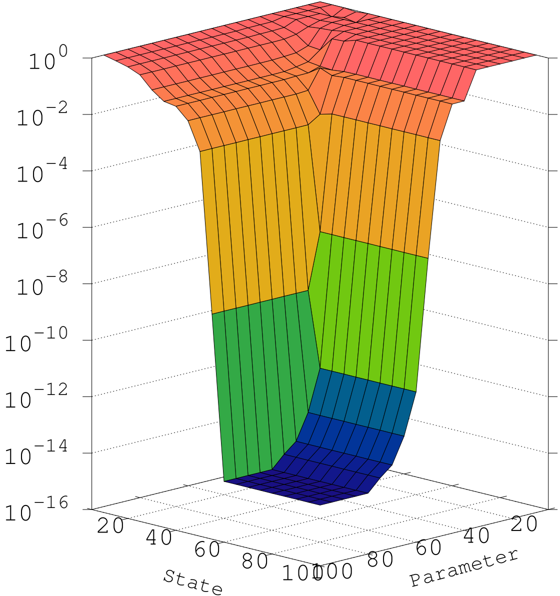
- Empirical Controllability Gramian
- Empirical Observability Gramian
- Empirical Linear Cross Gramian
- Empirical Cross Gramian
- Empirical Sensitivity Gramian
- Empirical Identifiability Gramian
- Empirical Joint Gramians

Features:

- Custom Solver Interface
- Non-Symmetric Cross Gramian
- Compatible with OCTAVE and MATLAB
- Vectorized and Parallelizable
- Open Source License (BSD 2-Clause)

More info at: `gramian.de`

Nonlinear Numerical Results ($L_2 \otimes L_2$ Error)



Distributed Empirical Cross Gramian

Empirical Cross Gramian:

$$\widehat{W}_X = \sum_{m=1}^{M} \int_0^\infty \Psi^m(t)\mathrm{d}t \in \mathbb{R}^{N \times N}$$

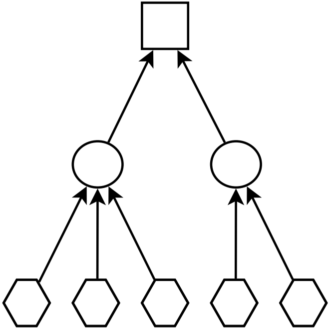
$$\Psi_{ij}^m(t) := \langle x_i^m(t), y_m^j(t) \rangle$$

Empirical Cross Gramian's k-th column:

$$\Rightarrow \widehat{W}_{X,*k} = \sum_{m=1}^{M} \int_0^\infty \psi^{mk}(t)\mathrm{d}t \in \mathbb{R}^{N \times 1}$$

$$\psi_i^{mk}(t) := \langle x_i^m(t), y_m^k(t) \rangle$$

HAPOD – Hierarchical Approximate POD



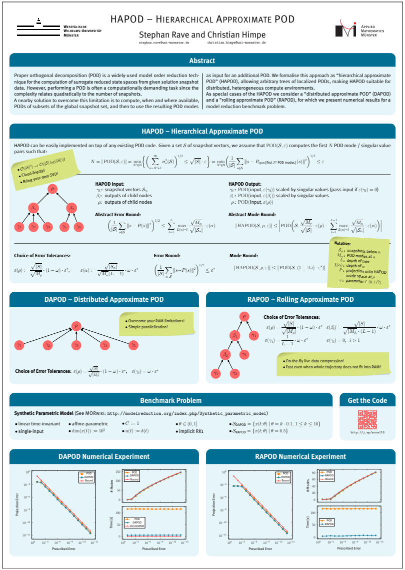
HAPOD:

- Rooted tree with ...
- nodes representing PODs of ...
- child nodes' PODs and ...
- leafs of snapshots.

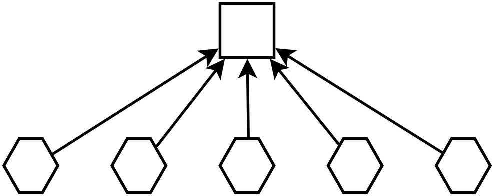
Provides:

- Approximation error bound
- Mode bound

Check Out Stephan's HAPOD Poster



DAPOD – Distributed Approximate POD



DAPOD:

- Flat tree
- Tall and skinny partitioning

Distributed Direct Truncation

Distributed Empirical Cross Gramian:

$$\widetilde{W_X} = \{\widehat{W}_{X,*K_i} | K_i \subset \mathbb{N}, K_i \cap K_j = 0, \bigcup_i K_i = 1 \dots N\}$$

Distributed Approximate POD:

$$\text{DAPOD}(\widetilde{W_X})$$

Linear Model Reduction Benchmark[5]

Affine Parametric SISO System:

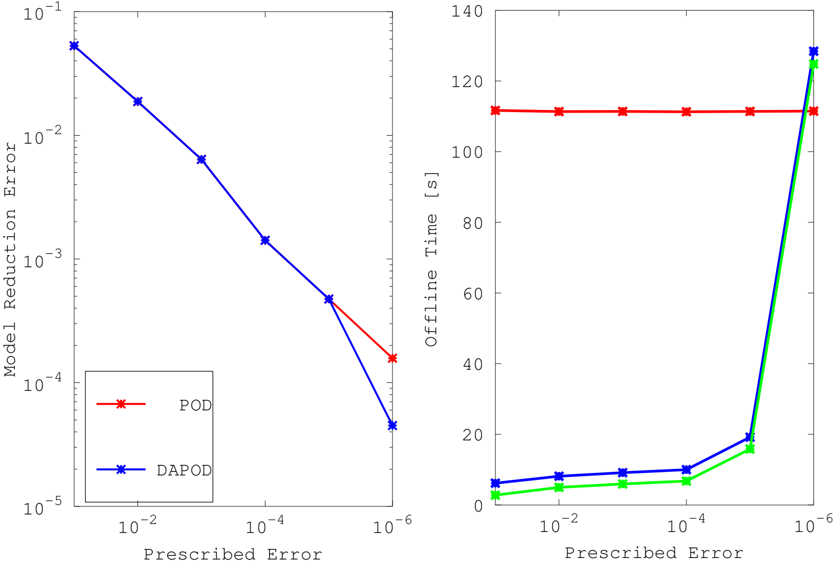
$$\begin{aligned} \dot{x}(t) &= (A_0 + A_\theta \theta)x(t) + Bu(t) \\ y(t) &= Cx(t) \\ x(0) &= 0 \end{aligned}$$

System Dimensions:

- $\dim(u(t)) = 1$
- $\dim(x(t)) = 10^4$
- $\dim(y(t)) = 1$
- $\dim(\theta) = 1$

[5] See: MORwiki modelreduction.org/index.php/Synthetic_parametric_model

Linear Numerical Results (L_2 Error & POD Time)



Parallelization

3 Levels:

- Distributed Memory:

  $\longrightarrow$ Distributed Empirical **Cross** Gramian
- Explicit Shared Memory:

  $\longrightarrow$ "Observability"-Snapshots
- Implicit Shared Memory / SIMD / GPU Offload:

  $\longrightarrow$ "Gramian" Computation

Up Next

- **Non-Symmetric** Cross Gramian[6]
- Distributed Empirical **Joint Gramian**
- **Dynamic Mode Decomposition**[7] Hyperreduction
- **Balanced Gains**[8]

---

6 We have a preprint: [H. & Ohlberger'15]
7 [Proctor et al.'15]
8 [Davidson'86]

tl;dl

Summary:

- Combined State and Parameter Reduction
  for Nonlinear Parametric Systems
- Distributed Empirical Cross Gramian
  utilizing the DAPOD

wwwmath.uni-muenster.de/u/himpe

Thanks!