Westfälische Wilhelms-Universität Münster

Applied Mathematics Münster

Combined State and Parameter Reduction

for Nonlinear Systems with an Application in Neuroscience

Christian Himpe (christian.himpe@uni-muenster.de)

WWU Münster
Institute for Computational and Applied Mathematics

Oberseminar Angewandte Mathematik

May the 4th (be with you) 2016

Disclaimer

1. This presentation is a guided tour through my PhD thesis:

   "Combined State and Parameter Reduction for Nonlinear Systems with an Application in Neuroscience"

2. There will be some system theory.

Introduction

1. Motivation
2. Aim
3. Dual Approach
4. Literature
5. Overview

Motivation I

- Connectivity Inference
- Large-Scale Networks
- From Measurements

Numerical Challenge:

- (Inverse Problem)
- High-Dimensional Parameter-Space
- Many-Query Setting
- High-Dimensional State-Space

Motivation II

Neuroscientific Application:

- Dynamic Causal Modelling (DCM)
- Functional Neuroimaging Data
  (i.e.: fMRI, fNIRS, EEG, MEG)
- Infer Connectivity of Brain Regions
- Controlled Excitation Experiments

Aim

- Combined
  1. State-Space Reduction
  2. Parameter-Space Reduction
- Network Models
- Nonlinear Dynamics

Dual Approach

1. Gramian-Based Combined Reduction
   - System-Theoretic Approach
   - Mini Summary: “Treat parameters as states”
2. Optimization-Based Combined Reduction
   - Greedy Sampling Approach
   - Mini Summary: “Parameter first, state later”

Literature I

Gramian-Based Combined Reduction:

- S. Lall, J.E. Marsden, and S. Glavaski. **Empirical Model Reduction of Controlled Nonlinear Systems**. In Proceedings of the 14th IFAC Congress, volume F, pages 473–478, 1999.
- S. Streif, R. Findeisen, and E. Bullinger. **Relating Cross Gramians and Sensitivity Analysis in Systems Biology**. Theory of Networks and Systems, 10.4:437–442, 2006.
- D. Geffen, R. Findeisen, M. Schliemann, F. Allgöwer, and M. Guay. **Observability Based Parameter Identifiability for Biochemical Reaction Networks**. In Proceedings of the American Control Conference, pages 2130–2135, 2008.

Literature II

Optimization-Based Combined Reduction:

- T. Bui-Thanh, K. Willcox, and O. Ghattas. **Model Reduction for Large-Scale Systems with High-Dimensional Parametric Input Space**. SIAM Journal on Scientific Computing, 30(6):3270–3288, 2008.
- C.E. Lieberman, K. Willcox, and O. Ghattas. **Parameter and State Model Reduction for Large-Scale Statistical Inverse Problems**. SIAM Journal on Scientific Computing , 32(5):2523–2542, 2010.

Overview

Preliminaries

1. Dynamical Systems
2. Control Systems
3. Model Reduction
4. Reduced Order Model Quality
5. Principal Axis Transformation

Dynamical Systems

Initial Value Problem:

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

Steady-State $\bar{x}$:

$$f(t, \bar{x}) = 0$$

Local Linearization:

$$A := \frac{\partial f}{\partial x}(\bar{x}) \Rightarrow \dot{x}(t) \approx Ax(t)$$
$$\Rightarrow x(t) \approx e^{At}$$

Control Systems I

General Time-Invariant Control System:

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

Linear Time-Invariant Control System:

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

Symmetric System:

$$C\,\mathrm{e}^{A}\,B = (C\,\mathrm{e}^{A}\,B)^{\intercal} \Leftrightarrow CA^{-1}B = (CA^{-1}B)^{\intercal}$$

Notation:

- Input Dimension: $M := \dim(u(t))$
- State Dimension: $N := \dim(x(t))$
- Output Dimension: $Q := \dim(y(t))$
- Parameter Dimension: $P := \dim(\theta)$

Model Reduction I

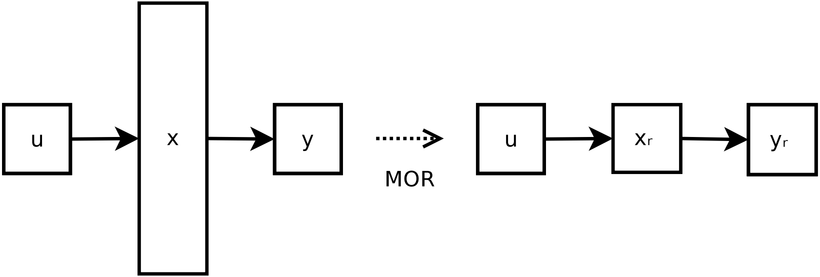


Figure: Full Order Model (FOM) and Reduced Order Model (ROM)

Typical Setting:

- $\dim(x(t)) = N \gg 1$
- $\dim(u(t)) = M \ll N$
- $\dim(y(t)) = Q \ll N$
- $\dim(\theta) = P \gg 1$

Model Reduction II

State-Space ROM:

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

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

Model Reduction III

Parametric State-Space ROM:

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

- $\|y(\theta) - y_r(\theta)\| \ll 1 \quad \forall \theta \in \Theta$

Model Reduction IIII

Parameter-Space ROM:

$$\dot{x}(t) = f_r(x(t), u(t), \theta_r)$$
$$y(t) = g_r(x(t), u(t), \theta_r)$$
$$x(0) = x_0$$

- $\dim(\theta_r) =: p \ll P$
- $\|y(\theta) - y(\theta_r)\| \ll 1 \quad \forall \theta \in \Theta$

Model Reduction V

Combined State and Parameter ROM:

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

- $\|y(\theta) - y_r(\theta_r)\| \ll 1 \quad \forall \theta \in \Theta$

Model Reduction VI

Projection-Based Model Reduction:

- State-space Galerkin Projection $U \in \mathbb{R}^{N \times n}$:
  $x_r = U^\intercal x \rightarrow x \approx U x_r$
- Parameter-Space Galerkin Projection $\Pi \in \mathbb{R}^{P \times p}$:
  $\theta_r = \Pi^\intercal \theta \rightarrow \theta \approx \Pi \theta_r$

Model Reduction VII

Projection-Based ROM:

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

Reduced Order Model Quality [Baur et al.'11]

Time-Domain Only!

State-Space Quality:

$$\|y\|_{L_2} = \sqrt{\int_0^\infty \|y(t)\|_2^2 \mathrm{d}t}$$

Joint State- and Parameter-Space Quality:

$$\|y(\theta)\|_{L_2 \otimes L_2} = \sqrt{\int_\Theta \|y(\theta)\|_{L_2}^2 \mathrm{d}\theta}$$

$$\|y(\theta)\|_{L_2 \otimes L_\infty} = \sup_{\theta \in \Theta} \|y(\theta)\|_{L_2}$$

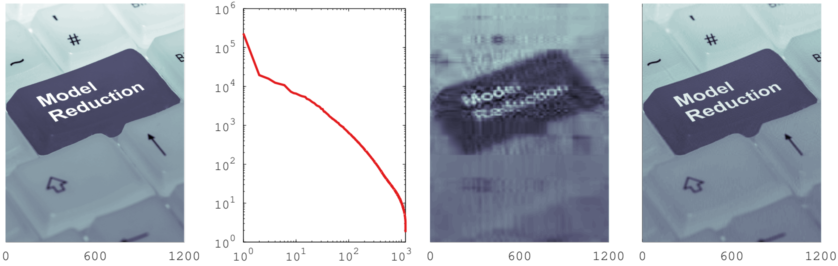
Principal Axis Transformation

Singular Value Decomposition:

$$A = UDV$$

$$UU^{\mathsf{T}} = \mathbb{1},\ V^{\mathsf{T}}V = \mathbb{1},\ UV = \mathbb{1}$$

$$D_{ii} = \sqrt{\lambda_i(AA^{\mathsf{T}})}$$



Gramian-Based Combined Reduction

1. State Reduction
2. System Gramians
3. Empirical Gramians
4. Parameter Reduction
5. Combined Reduction

State Reduction [Francis'87, Antoulas'05]

Convolution Operator S:

$$y(t) = S(u)(t) = \int_0^\infty C\,\mathrm{e}^{A(t-\tau)}\,Bu(\tau)\mathrm{d}\tau$$

Hankel Operator H:

$$\begin{aligned} H(u)(t) = S(u)(-t) &= \int_{-\infty}^0 C\,\mathrm{e}^{A(t-\tau)}\,Bu(\tau)\mathrm{d}\tau \\ &= \mathcal{O} \circ \mathcal{C} \end{aligned}$$

System Gramians I

Gramian Matrix $W(V)$:

$$W_{ij} = \langle V_i, V_j \rangle$$

System Gramians:

- $\mathcal{C}$ontrollability Gramian (Matrix)
- $\mathcal{O}$bservability Gramian (Matrix)
- Cross Gramian (Matrix) $\leftarrow$ Chimera[1]!

[1] Not the ISD, but as versatile [H. & Ohlberger'15].

System Gramians II

Controllability Gramian W_C:

$$
\begin{aligned}
W_C &:= \mathcal{C}\mathcal{C}^* \\
&= \int_0^\infty \mathrm{e}^{At} BB^\mathsf{T} \mathrm{e}^{A^\mathsf{T} t} \,\mathrm{d}t \in \mathbb{R}^{N \times N} \\
&\Rightarrow AW_C + W_C A^\mathsf{T} = -BB^\mathsf{T}
\end{aligned}
$$

System Gramians III

Observability Gramian W_O:

$$\begin{aligned} W_O &:= \mathcal{O}^*\mathcal{O} \\ &= \int_0^\infty \mathrm{e}^{A^\intercal t}\, C^\intercal C\, \mathrm{e}^{At}\, \mathrm{d}t \in \mathbb{R}^{N\times N} \\ &\Rightarrow A^\intercal W_O + W_O A = -C^\intercal C \end{aligned}$$

System Gramians IIII [Fernando & Nicholson'83, Laub et al.'83]

Cross Gramian W_X ($M \stackrel{!}{=} Q$):

$$
\begin{aligned}
W_X &:= \mathcal{C}\mathcal{O} \\
&= \int_0^\infty \mathrm{e}^{At} BC \, \mathrm{e}^{At} \, \mathrm{d}t \in \mathbb{R}^{N \times N} \\
&\Rightarrow AW_X + W_X A = -BC
\end{aligned}
$$

System Gramians V

Why System Gramians?

$$\sigma_i(H) = \sqrt{\lambda_i(W_C W_O)}$$
$$\sigma_i(H) = |\lambda_i(W_X)| \quad \text{(for symmetric systems)}$$

System Gramians VI [Moaveni & Khaki-Sedigh'06, H. & Ohlberger'15]

Non-Symmetric Cross Gramian W_Z:

$$B = \begin{pmatrix} b_1 & \dots & b_M \end{pmatrix},$$

$$C = \begin{pmatrix} c_1 & \dots & c_Q \end{pmatrix}^\intercal,$$

$$W_Z := \int_0^\infty \mathrm{e}^{At} \Big(\sum_{m=1}^{M} b_m\Big)\Big(\sum_{q=1}^{Q} c_q\Big) \mathrm{e}^{At} \, \mathrm{d}t$$

System Gramians VII [Laub et al.'87, Sorensen & Antoulas'02]

Balanced Truncation (Classic Squareroot Method):

$$W_C \overset{\text{Cholesky}}{=} L_C L_C^\mathsf{T}, \quad W_O \overset{\text{Cholesky}}{=} L_O L_O^\mathsf{T}$$

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

$$\to \begin{cases} U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \\ V = \begin{pmatrix} V_1 & V_2 \end{pmatrix}^\mathsf{T} \end{cases}$$

Direct Truncation (Approximate* Balancing):

$$W_X \overset{\text{SVD}}{=} UDV$$

$$\to U = \begin{pmatrix} U_1 & U_2 \end{pmatrix}$$

$$\to V_1 = U_1^\mathsf{T}$$

Empirical Gramians I

Balancing for Nonlinear Systems:

- B. Moore. **Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction**. IEEE Transactions on Automatic Control, 26(1):17–32, 1981.
- U. Pallaske. **Ein Verfahren zur Ordnungsreduktion mathematischer Prozessmodelle**. Chemie Ingenieur Technik, 59(7):604–605, 1987.
- X. Ma and J.A. De Abreu-Garcia. **On the Computation of Reduced Order Models of Nonlinear Systems using Balancing Technique**. In Proceedings of the 27th IEEE Conference on Decision and Control, volume 2, pages 1165–1166, 1988.
- J.M.A. Scherpen. **Balancing for nonlinear systems**. Systems & Control Letters 21(2):143–153, 1993.
- S. Lall, J.E. Marsden, and S. Glavaski. **Empirical Model Reduction of Controlled Nonlinear Systems**. In Proceedings of the 14th IFAC Congress, volume F, pages 473–478, 1999.

Empirical Gramians II [Fernando & Nicholson'85, Shaker'12, Baur et al.'16]

Empirical Linear Cross Gramian W_x:

$$W_x = \int_0^\infty e^{At} BC\, e^{At}\, \mathrm{d}t$$
$$= \int_0^\infty (e^{At} B)\, (e^{A^\intercal t}\, C^\intercal)^\intercal \mathrm{d}t$$
$$= \int_0^\infty x(t)(z(t))^\intercal \mathrm{d}t$$

With:

- $x(t)$ - state trajectoy
- $z(t)$ - adjoint state trajectory

Empirical Gramians III [Streif et al'06, H. & Ohlberger'14]

Empirical Cross Gramian W_X:

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

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

With:

- $x^m(t)$ - state trajectoy for m-th perturbed input component
- $y^j(t)$ - output trajectory for j-th pertubed initial state component

Parameter Reduction I [Grewal & Glover'76, Singh & Hahn'05, Geffen et al.'08]

Augmented System:

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

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

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

Parameter Reduction II [H. & Ohlberger'14]

(Empirical) Joint Gramian
- Cross Gramian of Augmented System:

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

Uncontrollable Parameters:

$$W_m = 0$$
$$W_\theta = 0$$

Parameter Reduction III [H. & Ohlberger'14]

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

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

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

Parameter Projection as Principal Components:

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

Combined Reduction [H. & Ohlberger'14]

State-space projection:

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

Parameter-space projection:

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

Cross-Gramian-Based Combined Reduction:

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

Optimization-Based Combined Reduction

1. Parameter Reduction
2. State Reduction
3. Combined Reduction
4. Data-Driven Regularization
5. Nonlinear Systems

Parameter Reduction I [Bui-Thanh et al.'08]

Parameter-Space Projection:

$$\Pi = \theta_0 \cup \bigcup_{i=1}^{I} \left(\theta_i \cap \left(\bigcup_{j=0}^{i-1} \theta_j\right)^{\perp}\right)$$

Iterative Parameter Base Assembly:

$$\Pi_{I+1} = \{\Pi_I \cup (\theta_{I+1} \cap \Pi_I^{\perp})\}$$

Parameter Reduction II [Bui-Thanh et al.'08]

Greedy Selection of Next Parameter Base Component:

$$\theta_{I+1} = \mathrm{argmax}_{\theta \in \Theta} \, \|y(\theta) - y_r(\theta_r)\|_{L_2}^2$$

Properties:

- $\theta_{I+1} \neq \theta_{J \leq I}$
- $\|y(\theta) - y(\Pi_{I+1}\Pi_{I+1}^\intercal \theta)\|_{L_2} < \|y(\theta) - y(\Pi_I \Pi_I^\intercal \theta)\|_{L_2}$

Parameter Reduction III [Elden'77]

Tikhonov (L_2) Regularization Operator:

$$\mathcal{R}_{\beta_2} = \beta_2 \|\theta\|_2^2$$

Regularized Greedy Selection:

$$\theta_{I+1} = \mathrm{argmax}_{\theta \in \Theta} \|y(\theta) - y_r(\theta_r)\|_{L_2}^2 - \mathcal{R}_{\beta_2}$$

State Reduction I

State-space projection:

$$U_1 = \Phi_1(x(\theta_0)) \cup \bigcup_{i=1}^{I} \left(\Phi_1(x(\theta_I)) \cap \left(\bigcup_{j=0}^{i-1} \Phi_1(x(\theta_i))\right)^{\perp}\right)$$

State Reduction II [Sirovich'87]

Input-to-State-Based State-Space Reduced Basis Assembly:

$$U_{I+1} = \{U_I \cup (\mathrm{POD}_1(x(\theta_I)) \cap U_I^\perp)\}$$

State Reduction III

State-to-Output-Based State-Space Reduced Basis Assembly:

$$U_{l+1} = \{U_l \cup (\text{POD}_1(z(\theta_l)) \cap U_l^{\perp})\}$$

State Reduction IIII [Willcox et al.'05]

Input-to-Output-Based State-Space Reduced Basis Assembly:

$$U_{I+1} = \{U_I \cup (\mathrm{bPOD}_1(x(\theta_I)) \cap U_I^{\perp})\}$$

Combined Reduction I

Greedy-Sampling-Based Reduction for Inverse Problems:

- T. Bui-Thanh, K. Willcox, and O. Ghattas. **Model Reduction for Large-Scale Systems with High-Dimensional Parametric Input Space**. SIAM Journal on Scientific Computing, 30(6):3270–3288, 2008.
- O. Bashir, K. Willcox, O. Ghattas, B. van Bloemen Waanders, and J. Hill. **Hessian-based model reduction for large-scale systems with initial-condition inputs**. International Journal for Numerical Methods in Engineering, 73(6):844–868, 2008.
- C.E. Lieberman, K. Willcox, and O. Ghattas. **Parameter and State Model Reduction for Large-Scale Statistical Inverse Problems**. SIAM Journal on Scientific Computing , 32(5):2523–2542, 2010.
- C.E. Lieberman, K. Fidkowski, K. Willcox, and B. Van Bloemen Waanders. **Hessian-based model reduction: large-scale inversion and prediction**. International Journal for Numerical Methods in Fluids , 71(2):135–150, 2013.
- D. Galbally, K. Fidkowski, K. Willcox, and O. Ghattas. **Non-linear model reduction for uncertainty quantification in large-scale inverse problems**. International Journal for Numerical Methods in Engineering , 81(12):1581–1608, 2010.

Combined Reduction II [H. & Ohlberger'15]

Optimization-Based Combined Reduction:

$\theta_0 \leftarrow \bar{\theta}$

$\Pi_0 \leftarrow \theta_0$

$U_0 \leftarrow \Phi_1(x(\theta_0))$

for $l = 1 \dots p$

  $\theta_l \leftarrow \operatorname{argmin} \; -\|y(\theta) - y_r(\theta_r)\|_{L_2}^2 + \beta_2 \|\theta\|_2^2$

  $\Pi_l \leftarrow \operatorname{orth}(\Pi_{l-1}, \theta_l)$

  $U_l \leftarrow \operatorname{orth}(U_{l-1}, \Phi_1(x(\theta_l)))$

Combined Reduction III i.e.: [Buhr et al.'14]

Re-Iterated Gram-Schmidt:

$Q^{\intercal} Q = \mathbb{1}$

$b \leftarrow 0$

while $b < \epsilon$

- $v \leftarrow v - Q(Q^{\intercal} v)$
- $b \leftarrow \|v\|_2$
- $v \leftarrow b^{-1} v$

$Q \leftarrow \begin{pmatrix} Q & v \end{pmatrix}$

Data-Driven Regularization [H. & Ohlberger'15]

There is data y_d for an inverse problem. Use it.

Data Mismatch as Regularization:

$$\mathcal{R}_d = \beta_d \|y_d - y(\theta)\|_{L_2}^2$$

Extended Cost Functional:

$$J = \|y(\theta) - y_r(\theta_r)\|_{L_2}^2 - \beta_2 \|\theta\|_2^2 - \beta_d \|y(\theta) - y_d\|_{L_2}^2$$

Nonlinear Systems [Kunisch & Volkwein'02]

Prerequisites:

- Nonlinear Optimization
- POD-Based State Reduction

Combined State and Parameter ROM:

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

Software Implementation

1. Design Principles
2. Inverse Lyapunov Procedure
3. emgr - Empirical Gramian Framework
4. optmor - Optimization-Based Model Order Reduction
5. Nonlinear Benchmark

Design Principles

- Programming Language & Compatibility
  - OCTAVE: Open-Source
  - MATLAB: Quasi Standard
- Guided By:
  - Language Best Practices: [Johnson'11]
  - Performance Guidelines: [Altman'15]
- Availability for Replicability
- Configurability for Reproducibility
- Modularity for Reusability

Inverse Lyapunov Procedure [Smith & Fisher'03]

- Generate Random Systems
- Doing Balanced Truncation Backwards
  (Sample W_C and W_O, solve for A)
- Linear Parametrization
- Verification and Validation of Implementations
- Procedural Benchmark in the MORwiki

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 Gramian (+ Empirical Cross-Identifiability Gramian)

Features:

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

More info at: gramian.de

optmor – Optimization-Based Model Order Reduction (Version: 2.5, 02/2016)

Capabilities:

- Iterative Greedy Parameter-Space Sampling
- POD-Based State-Space Projection
- Tikhonov Regularization
- Data-Driven Regularization
- Re-Iterated Orthogonalization

Features:

- Custom Optimizer Interface
- Custom Solver Interface
- Compatible with OCTAVE and MATLAB
- Vectorized and Parallelizable
- Open Source License (BSD 2-Clause)

More info at: github.com/gramian/optmor

Nonlinear Benchmark I [Chen'99, Condon & Ivanov'04]

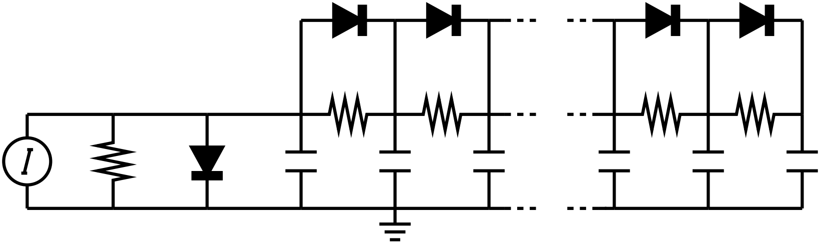
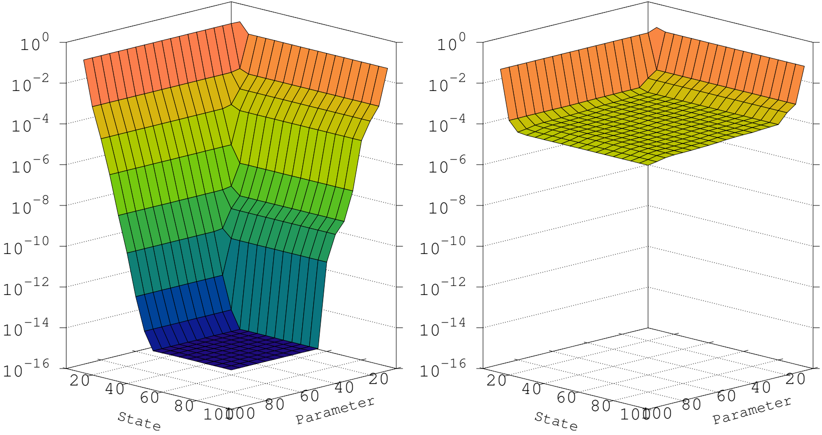


Figure: Nonlinear Resistor-Capacitor Cascade

- SISO System
- Nonlinear Resistors (Diodes)
- Parametrization of Linear Resistors
- Procedural Benchmark in the MORwiki
- Here: $N = P = 100$

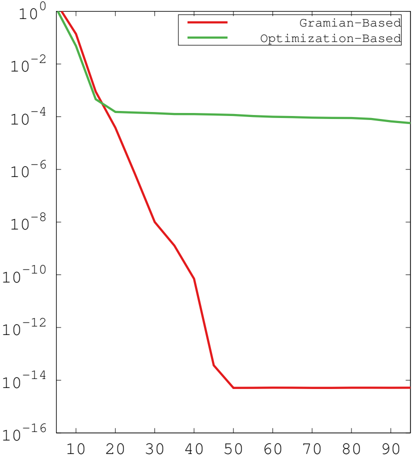
Nonlinear Benchmark II

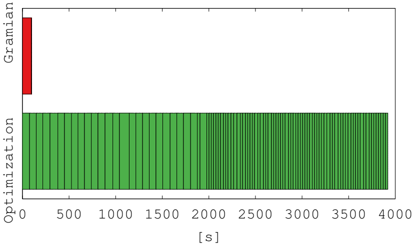
$L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



Nonlinear Benchmark III

$n = p$, Offline Timings, Gramian- vs. Optimization-Based:





An Application in Neuroscience

1. Neuronal Networks
2. Dynamic Causal Modelling
3. fMRI & fNIRS Dynamic Causal Model
4. EEG & MEG Dynamic Causal Model
5. Bayesian Inference

Neuronal Networks [Quan et al.'01]

Adjacency Matrix A:

- Square
- Represents weighted, directed finite graph
- A_{ij} Connection from j-th to i-th region

Hyperbolic Network Model:

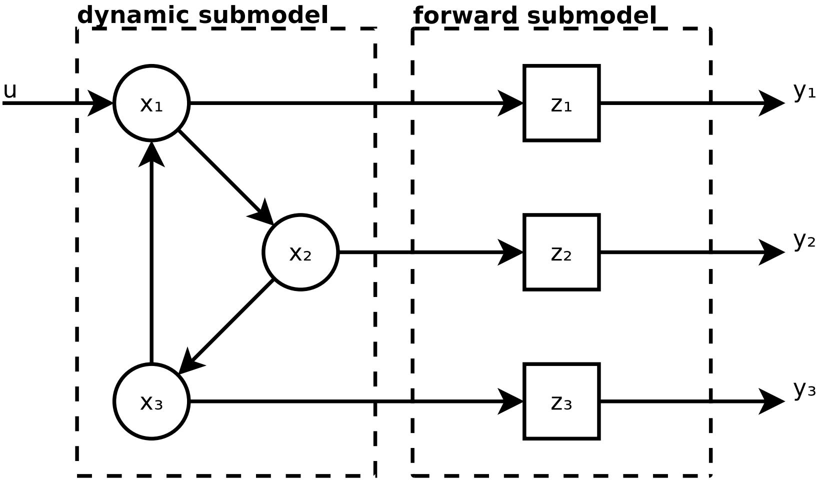
$$\dot{x}(t) = A\tanh(Kx(t)) + Bu(t)$$
$$y(t) = Cx(t)$$

Dynamic Causal Modelling I [Friston et al.'03]

Concept:

- Two Component Model
  1. Network Submodel (Dynamic Submodel)
  2. Observation Submodel (Forward Submodel)
- Connectivity Parametrization
- SIMO Models (Effective Connectivity)
- Bayesian Inference

Dynamic Causal Modelling II



fMRI & fNIRS DCM I [Friston et al.'03, Stephan & Friston'07]

Network Submodel (Taylor Series):

$$\begin{aligned}\dot{x}(t) &= f(x(t), u(t), \theta) \\ &\approx f(0, 0, \theta) + \frac{\partial f}{\partial x} x(t) + \frac{\partial f}{\partial u} u(t) \\ &= Ax(t) + Bu(t)\end{aligned}$$

Parametrization A:

$$A(\theta) = \mathrm{vec}^{-1}(\theta) \Rightarrow \theta \in \mathbb{R}^{N \times N}$$

fMRI & fNIRS DCM II [Friston'02, Friston et al.'03, Kamrani et al.'12]

Observation Submodel (for the i-th region):

$$\dot{s}_i(t) = x_i(t) - \kappa s_i(t) - \gamma(f_i(t) - 1)$$

$$\dot{f}_i(t) = s_i(t)$$

$$\dot{v}_i(t) = \frac{1}{\tau}(f_i(t) - v_i(t)^{\frac{1}{\alpha}})$$

$$\dot{q}_i(t) = \frac{1}{\tau}(\frac{1}{\rho} f_i(t)(1 - ((1 - \rho))^{\frac{1}{f_i(t)}}) - v_i(t)^{\frac{1}{\alpha}-1} q_i(t))$$

$$y_i(t) = V_0(k_1(1 - q_i(t)) + k_2(1 - v_i(t)))$$

fMRI & fNIRS Dynamic Causal Model III

Joint Dynamic and Forward Model for k Regions:

$(\dim(x(t)) = \dim(s(t)) = \dim(f(t)) = \dim(v(t)) = \dim(q(t)) = \dim(y(t)) = k)$

$$\begin{pmatrix} \dot{x}(t) \\ \dot{s}(t) \\ \dot{f}(t) \\ \dot{v}(t) \\ \dot{q}(t) \end{pmatrix} = \begin{pmatrix} A(\theta)x(t) + Bu(t) \\ x(t) - \kappa s(t) - \gamma(f(t) - \vec{1}_k) \\ s(t) \\ \frac{1}{\tau}(f(t) - v(t)^{\frac{1}{\alpha}}) \\ \frac{1}{\tau}(\frac{1}{\rho}f(t)(\vec{1}_k - ((1-\rho)\,\vec{1}_k)^{\frac{1}{f(t)}}) - v(t)^{\frac{1}{\alpha}-1} \odot q(t)) \end{pmatrix}$$

$$y(t) = V_0(k_1(\vec{1}_k - q(t)) + k_2(\vec{1}_k - v(t)))$$

EEG & MEG DCM [David et al.'05, David et al.'06, Moran et al.'07]

Joint Dynamic and Forward Model for k Regions:
$(\dim(x(t)) = \dim(v(t)) = 5k,\ \dim(y(t)) = k)$

$$\begin{pmatrix} \dot{v}(t) \\ \dot{x}(t) \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{1}_{5k} \\ -T^2 \otimes \mathbb{1}_k & -2T \otimes \mathbb{1}_k \end{pmatrix} \begin{pmatrix} v(t) \\ x(t) \end{pmatrix} + \begin{pmatrix} 0 \\ A_v \end{pmatrix} \bar{\varsigma}_\kappa(A_\varsigma v(t)) + \frac{H_e}{\tau_e}(\delta_{8,1}^{10\times 1} \otimes \vec{1}_k) u(t)$$

$$y(t) = Lx(t)$$

$$A_v = \begin{pmatrix} \frac{H_e}{\tau_e}\gamma_1 \mathbb{1}_k & 0 & 0 & 0 & 0 \\ 0 & \frac{H_i}{\tau_i}\gamma_2 \mathbb{1}_k & 0 & 0 & 0 \\ 0 & 0 & \frac{H_e}{\tau_e}(A_F(\theta) + \gamma_3 \mathbb{1}_k) & 0 & 0 \\ 0 & 0 & 0 & \frac{H_e}{\tau_e}\gamma_4 \mathbb{1}_k & 0 \\ 0 & 0 & 0 & 0 & \frac{H_i}{\tau_i}\gamma_5 \mathbb{1}_k \end{pmatrix}$$

Bayesian Inference

Bayes' Rule:

$$P(\theta|y_d) = \frac{P(y_d|\theta)P(\theta)}{P(y_d)},$$

Data Model Assuming Gaussian Noise:

$$y_d = y(\theta) + \mathrm{N}(0, v)$$

MAP Estimate:

$$P(\theta|y_d) \propto \exp\Big(-\frac{1}{2}\|y(\theta) - y_d\|^2_{\sigma^{-1}_{y|\theta}} - \frac{1}{2}\|\theta - \mu_\theta\|^2_{\sigma^{-1}_\theta}\Big)$$

$$\rightarrow \theta_{\mathrm{MAP}} = \mathrm{argmin}_{\theta\in\mathbb{R}^P} \left(\frac{1}{2}\|f(\theta) - y_d\|^2_{\sigma^{-1}_{y|\theta}} + \frac{1}{2}\|\theta - \mu_\theta\|^2_{\sigma^{-1}_\theta}\right)$$

Numerical Results

1. Experimental Setup
2. Hyperbolic Network Model
3. fMRI & fNIRS Dynamic Causal Model
4. EEG & MEG Dynamic Causal Model
5. Combined Reduction for Inverse Problems

Experimental Setup

Models (Nonlinear + Linearized):

1. Hyperbolic Network Model
2. fMRI & fNIRS Dynamic Causal Model
3. EEG & MEG Dynamic Causal Model

Methods:

1. Gramian-Based Combined Reduction
2. Optimization-Based Combined Reduction

Measures:

- $L_2 \otimes L_2$-Norm
- ($L_2 \otimes L_\infty$-Norm)
- Offline Time

Hyperbolic Network Model I

Nonlinear Model:

$$\dot{x}(t) = A \tanh(K(\theta)x(t)) + Bu(t)$$
$$y(t) = Cx(t)$$

Linearized Model:

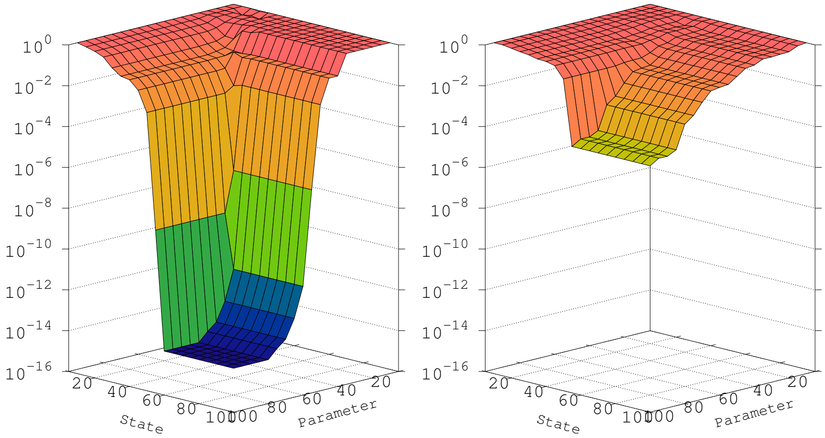
$$\dot{x}(t) = AK(\theta)x(t) + Bu(t)$$
$$y(t) = Cx(t)$$

System Dimensions:

- $M = Q = 1$
- $N = 100$
- $P = 100$

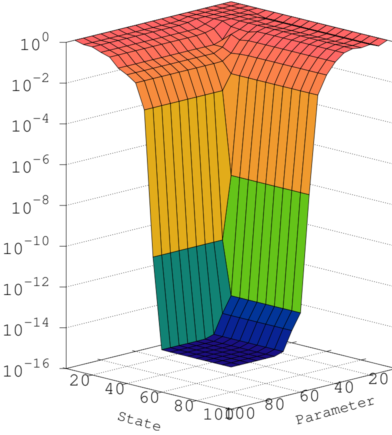
Hyperbolic Network Model II

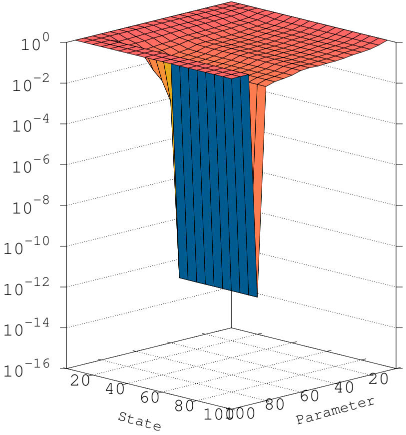
Nonlinear Model, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



Hyperbolic Network Model III

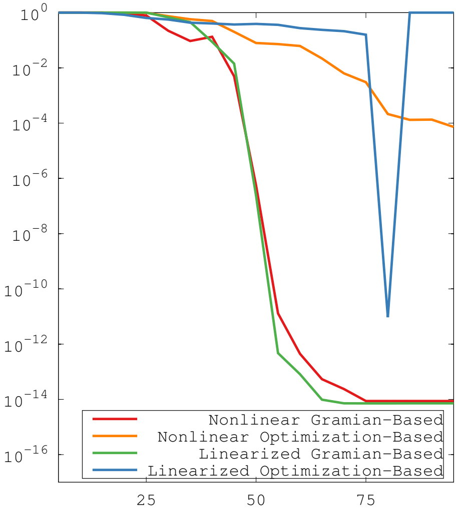
Linearized Model, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:





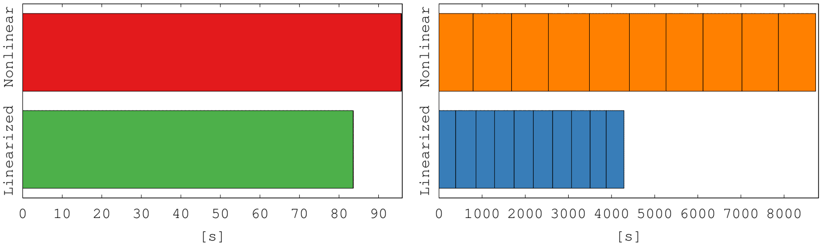
Hyperbolic Network Model IIII

$n = p$, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



Hyperbolic Network Model V

Offline Timings, Gramian- vs. Optimization-Based:



fMRI & fNRIS Dynamic Causal Model I

Properties:

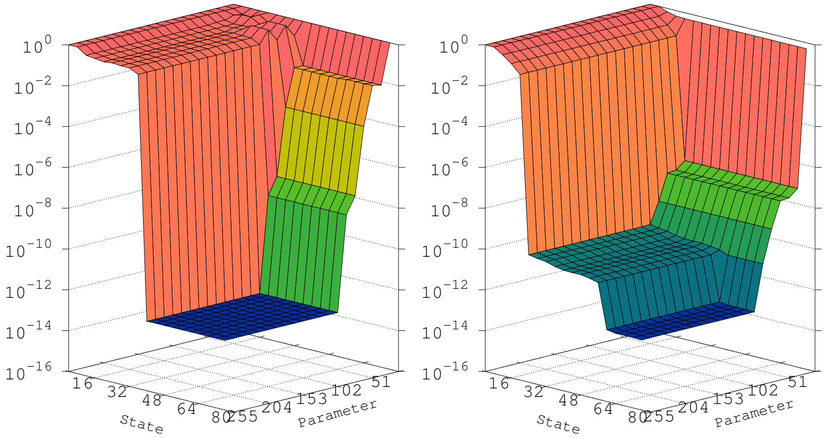
- Linear Network Submodel
- Highly Nonlinear Observation Submodel

System Dimensions ($k = 16$ Regions):

- $M = 1$
- $Q = k$
- $N = 5k = 80$
- $P = k^2 = 256$

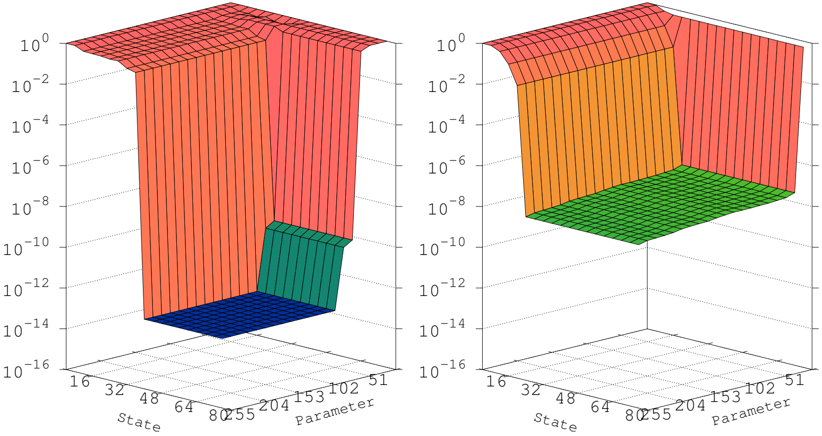
fMRI & fNRIS Dynamic Causal Model II

Nonlinear Model, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



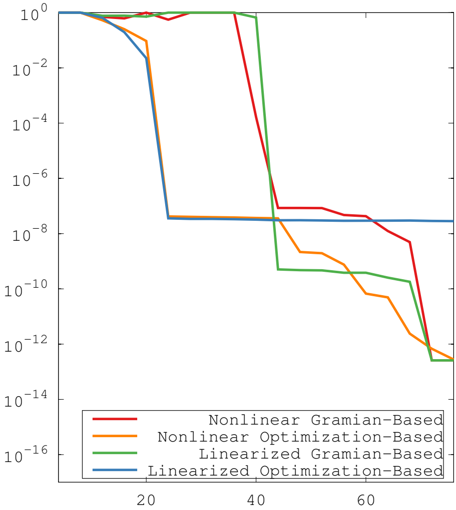
fMRI & fNRIS Dynamic Causal Model III

Linearized Model, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



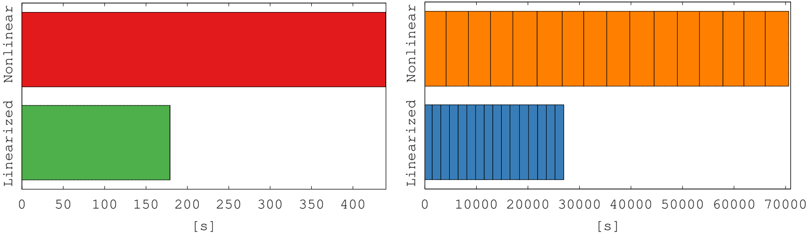
fMRI & fNRIS Dynamic Causal Model IIII

$n = p < 80$, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



fMRI & fNRIS Dynamic Causal Model V

Offline Timings, Gramian- vs. Optimization-Based:



EEG & MEG Dynamic Causal Model I

Properties:

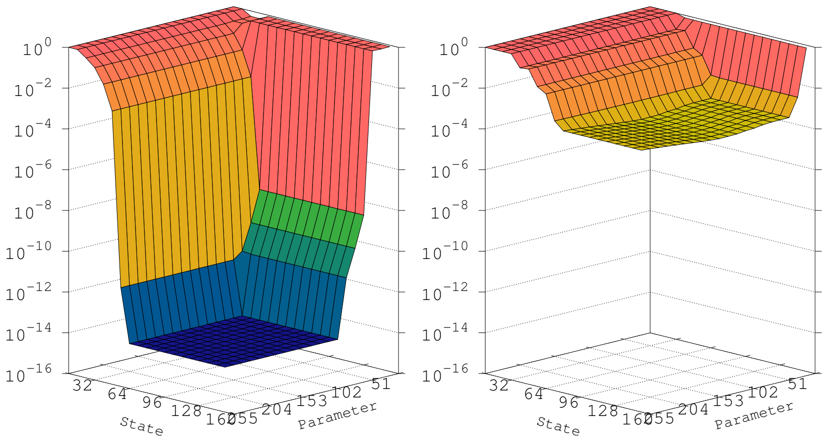
- Nonlinear Second-Order Network Submodel
- Linear Observation Submodel

System Dimensions ($k = 16$ Regions):

- $M = 1$
- $Q = k$
- $N = 10k = 160$
- $P = k^2 = 256$

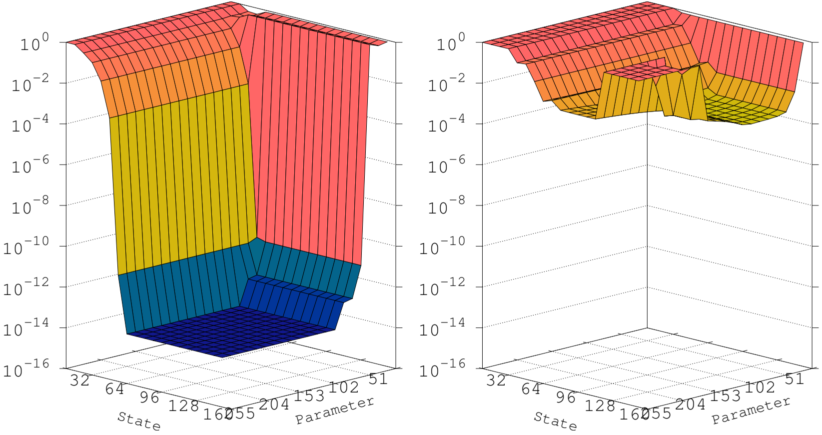
EEG & MEG Dynamic Causal Model II

Nonlinear Model, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



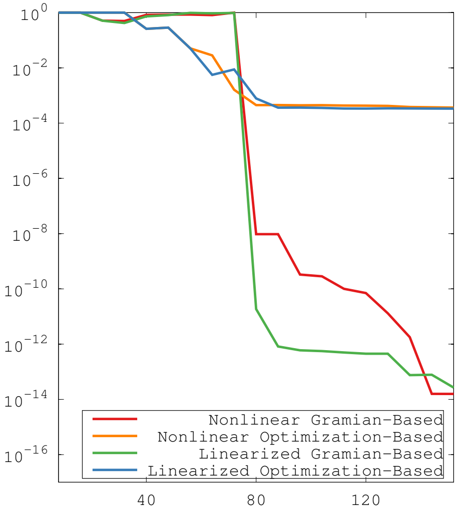
EEG & MEG Dynamic Causal Model III

Linearized Model, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



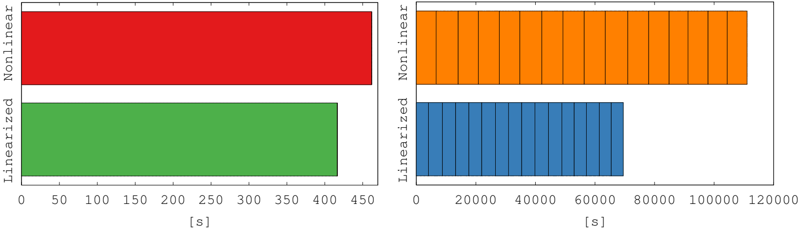
EEG & MEG Dynamic Causal Model IIII

$n = p < 160$, $L_2 \otimes L_2$-Norm, Gramian- vs. Optimization-Based:



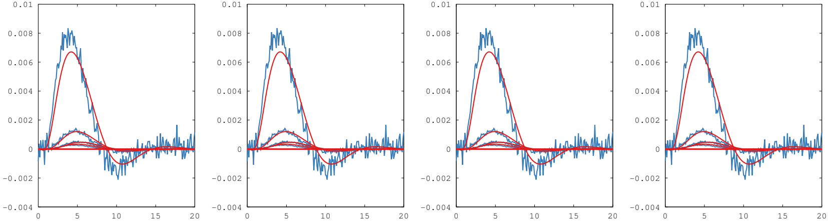
EEG & MEG Dynamic Causal Model V

Offline Timings, Gramian- vs. Optimization-Based:



Combined Reduction for Inverse Problems I

FOM, Gramian-Based ROM, Optimization-Based ROM and Data-Driven Optimization-Based ROM Inversion of fMRI Data:



Combined Reduction for Inverse Problems II

fMRI & fNIRS DCM	FOM	Gramian-Based ROM	Optim.-Based ROM	Optim.-Based ROM (DD)
Parameter Dim.	256	40	24	24
Output Error	0.01	0.01	0.01	0.01
Parameter Error	0.25	0.25	0.25	0.25
Offline Time	-	381s	11957s	9774s
Online Time	1447s	488s	233s	226s
Single Total Time	1447s	869s	12190s	10000s
Multi Breakeven	-	1	10	8

Conclusion

1. Summary
2. Abstract Comparison
3. Numerical Comparison
4. Outlook
5. Concluding Remarks

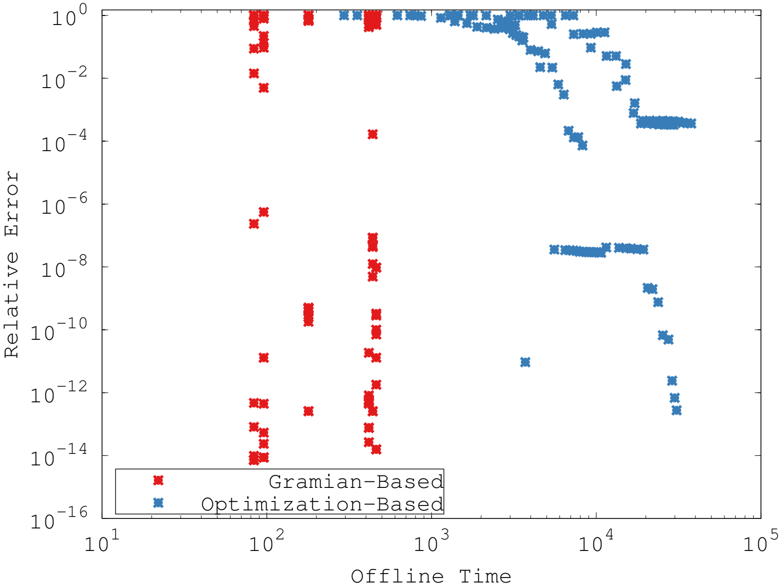
Summary

- Combined State and Parameter Reduction:
  1. Gramian-Based Combined Reduction ✓
  2. Optimization-Based Combined Reduction ✓
- Nonlinear Systems ✓
- Reduced Order Dynamic Causal Models ✓
- Reduced Inverse Problem ✓

Abstract Comparison

	Gramian-Based	**Optimization-Based**
State-Space	Input-to-Output	Input-to-State
Parameter-Space	State-to-Output	Input-to-Output
Sampling Strategy	Sparse	Adaptive
Assembly	Direct	Iterative
Associated Norm	$\| \cdot \|_{L_2 \otimes L_2}$	$\| \cdot \|_{L_2 \otimes L_\infty}$
Complexity	$M + N + P$	$(p-1)(P+1)+1$

Numerical Comparison



Outlook

Gramian-Based:

- Hyper-Reduction (EIM/DEIM, DMD/DMDC/IODMD)
- Kernel Methods (RKHS)
- Parallelization (Distributed Memory, GPU)

Optimization-Based:

- L_1-Regularization (Elastic Net)
- Derivative Information (AD, W_I)

Application-Wise:

- Complex Networks
- Time-Varying (Nonlinear Parametric) Systems

Concluding Remarks

- Both methods work for nonlinear systems
- Empirical gramians are faster and easier
- Reusable software is available

tl;dl

- Empirical-Cross-Gramian-Based
- Greedy-Sampling-Optimization-Based
- Combined State and Parameter Reduction
- Dynamic Causal Models

`wwwmath.uni-muenster.de/u/himpe`

Thanks!

Related Publications

- C. Himpe and M. Ohlberger. **A Unified Software Framework for Empirical Gramians**. Journal of Mathematics, 2013:1–6, 2013.
- C. Himpe and M. Ohlberger. **Model Reduction for Complex Hyperbolic Networks**. In Proceedings of the ECC, pages 2739–2743, 2014.
- C. Himpe and M. Ohlberger. **Cross-Gramian Based Combined State and Parameter Reduction for Large-Scale Control Systems**. Mathematical Problems in Engineering, 2014:1–13, 2014.
- C. Himpe and M. Ohlberger. **Combined State and Parameter Reduction**. In PAMM, volume 14(1), pages 825–826, 2014.
- C. Himpe and M. Ohlberger. **The Empirical Cross Gramian for Parametrized Nonlinear Systems**. In Mathematical Modelling, vol 8, pages 727–728, 2015.
- C. Himpe and M. Ohlberger. **Data-driven combined state and parameter reduction for inverse problems**. Advanced in Computational Mathematics (MoRePaS Special Issue), 41(5):1343–1364, 2015.
- C. Himpe and M. Ohlberger. **A Note on the Non-Symmetric Cross Gramian**. Preprint, math.OC(1501.05519):1–6, 2015.
- U. Baur, P. Benner, B. Haasdonk, C. Himpe, I. Martini, and M. Ohlberger. **Comparison of methods for parametric model order reduction of instationary problems**. In Model Reduction and Approximation: Theory and Algorithms. SIAM, 2016. (To appear)