Westfälische Wilhelms-Universität Münster

Applied Mathematics Münster

Combined Reduction for Uncertainty Quantification

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

WWU Münster
Institute for Computational and Applied Mathematics

SIAM Conference on Uncertainty Quantification (#SIAMUQ16)
07.04.2016

Disclaimer

The following presentations contains parts of my PhD thesis:

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

in obscenely abbreviated form!

Outline

- **Motivation:** Dynamic Causal Modelling
- **Means of Execution:** Model Order Reduction
- **Method I:** Gramian-Based Combined Reduction
- **Method II:** Optimization-Based Combined Reduction
- **Showdown:** Numerical Results
- **Post-Mortem**

Practical Motivation

Neuroscientific Application:

- Brain Region Connectivity
- via Functional Neuroimaging Data
  - fMRI & fNIRS
  - EEG & MEG

Numerical Challenge:

- Infer Network Connectivity
- by Inverse Problem

Dynamic Causal Modelling (DCM) [Friston et al.'03]

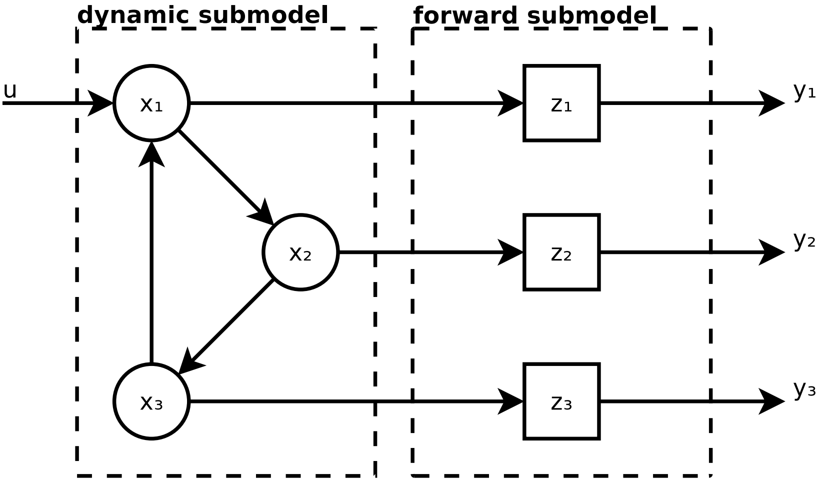
DCM:

- **Connectivity Models**
- Bayesian Inference
  - Hypothesis Testing
  - EM-Algorithm

Model Components:

- Dynamic Submodel
  - Simulate neuronal activity
- Forward Submodel
  - Transform to measurements

DCM Illustration



fMRI & fNIRS DCM [Friston'02, Friston et al.'03, Stephan & Friston'07]

Joint Dynamic and Forward Model for k Regions:

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

$$\begin{pmatrix} \dot{x}(t) \\ \dot{s}(t) \\ \dot{f}(t) \\ \dot{v}(t) \\ \dot{q}(t) \end{pmatrix} = \begin{pmatrix} A(\theta)x(t) \\ x(t) - \kappa s(t) - \gamma(f(t) - \vec{1}_k) \\ s(t) \\ \frac{1}{\tau}(f(t) - v(t)^{\frac{1}{\alpha}}) \\ \frac{1}{\tau}(\frac{1}{\rho} f(t)(\vec{1}_k - ((1-\rho)\,\vec{1}_k)^{\frac{1}{f(t)}}) - v(t)^{\frac{1}{\alpha}-1} \odot q(t)) \end{pmatrix} + \begin{pmatrix} B \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} u(t)$$

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

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

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

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

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

DCM Properties

Common Features:

- SIMO Model (Effective Connectivity)
- Non-Affine Parameter Mapping
- Homogeneous Parameters

fMRI & fNIRS:

- Linear Dynamic Submodel
- Highly Nonlinear Forward Submodel

EEG & MEG:

- Weakly Nonlinear Second-Order Dynamic Submodel
- Linear Forward Submodel

Abstract Motivation

Setting:

- Many-Query Setting (Optimization Problem)
- Nonlinear, Parametric Input-Output System
- "High-Dimensional" State- and **Parameter-Space**

Objective:

- Combined (State and Parameter) Reduction
- Uncertainty Quantification

General Input-Output System

Full Order Model (FOM):

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

- Input: $u : \mathbb{R} \to \mathbb{R}^M$
- State: $x : \mathbb{R} \to \mathbb{R}^N$
- Output: $y : \mathbb{R} \to \mathbb{R}^Q$
- Parameter: $\theta \in \mathbb{R}^P$

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)

State and Parameter Reduced System:

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

- $x_r : \mathbb{R} \rightarrow \mathbb{R}^{n \ll N}$
- $\theta_r \in \mathbb{R}^{p \ll P}$
- $\|y(\theta) - y_r(\theta_r)\|_? \ll 1$

Projection-Based Combined State and Parameter Reduction

Galerkin Projection V:

$$\mathrm{rank}(V) = n, \quad VV = V, \quad V^\intercal V = \mathbb{1}$$

State-Space Reduction:

$$x_r(t) = U^\intercal x(t) \Rightarrow x(t) \approx U x_r(t)$$

Parameter-Space Reduction:

$$\theta_r = \Pi^\intercal \theta \Rightarrow \theta \approx \Pi \theta_r$$

Combined State and Parameter ROM:

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

Gramian-Based Combined Reduction

- Direct Approach
- Based on Empirical Gramians
- Incorporate Knowledge on Operating Region
- "Separate" State- and Parameter-Space ROMs

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

Linear Time-Invariant System:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t)\\ y(t) &= Cx(t)\end{aligned}$$

For square systems ($M = Q$):

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

Direct Truncation:

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

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

Empirical Linear Cross Gramian (State $x(t)$, Adjoint $z(t)$):

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

Empirical Cross Gramian (State $x(t)$, Output $y(t)$):

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

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

Empirical 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):

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

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

Parameter Projection as Principal Components:

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

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:

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

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

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`

Optimization-Based Combined Reduction

- Iterative Approach
- Based on Greedy Sampling
- Directly Assemble Low-Rank Projections
- Parameter First, State Later

Greedy Sampling [Bui-Thanh et al.'08, Lieberman et al.'10]

Greedy Selection of Next Parameter Base Component:

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

Iterative Parameter Base Assembly:

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

State-Space Reduced Basis Assembly:

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

Regularization

Tikhonov Regularization Operator:

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

Regularized Greedy Selection:

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

Enforced Orthogonality i.e.: [Buhr et al.'14]

Re-Iterated Gram-Schmidt:

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

$b \leftarrow 0$

while $b < \epsilon$

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

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

Algorithm [H. & Ohlberger'15]

Optimization-Based Combined Reduction:

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

$\Pi_0 \leftarrow \theta_0$

$U_0 \leftarrow \mathrm{pod}_1(x(\theta_0))$

for $I = 1 \dots p$

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

  $\Pi_I \leftarrow \mathrm{orth}(\Pi_{I-1}, \theta_I)$

  $U_I \leftarrow \mathrm{orth}(U_{I-1}, \mathrm{pod}_1(x(\theta_I)))$

Greedy-Sampling-Based Combined Reduction

Parameter-space projection:

$$\Pi_1 = \theta_0 \cup \bigcup_{i=1}^{I} \Big(\theta_i \cap \big(\bigcup_{j=0}^{i-1} \theta_j\big)^{\perp}\Big)$$

State-space projection:

$$U_1 = \mathrm{pod}_1(x(\theta_0)) \cup \bigcup_{i=1}^{I} \Big(\mathrm{pod}_1(x(\theta_I)) \cap \big(\bigcup_{j=0}^{i-1} \mathrm{pod}_1(x(\theta_i))\big)^{\perp}\Big)$$

Combined state and parameter ROM:

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

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`

Experimental Setup

MOR Procedure:

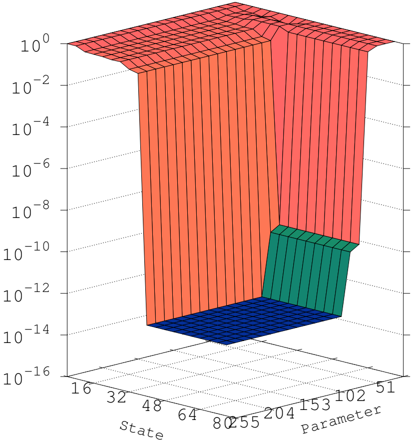
- Combined State and Parameter Reduction
- fMRI/fNIRS and EEG/MEG DCMs for $k = 16$
- Parametrization: $A(\theta) = A_0 + \mathrm{vec}^{-1}(\theta)$
- Linearized and Nonlinear Variants
- $L_2 \otimes L_2$-Norm ($\star$)
- $L_2 \otimes L_\infty$-Norm

System Dimensions:

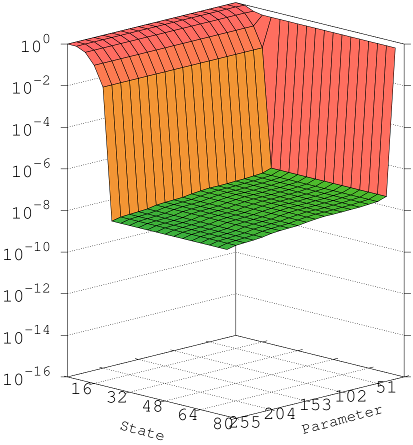
- $\dim(u(t)) = 1$
- $\dim(x_{\mathrm{fMRI}}(t)) = 80$
- $\dim(x_{\mathrm{EEG}}(t)) = 160$
- $\dim(y(t)) = 16$
- $\dim(\theta) = 256$ (!)

Numerical Results fMRI & fNIRS DCM (Linearized)

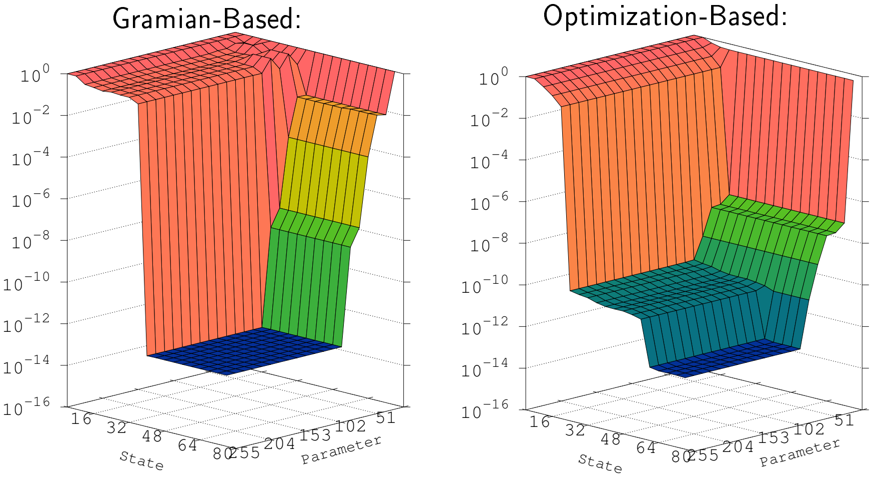
Gramian-Based:



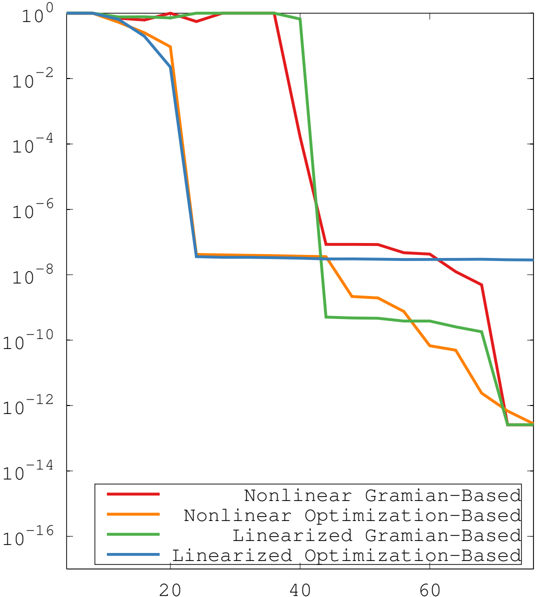
Optimization-Based:



Numerical Results fMRI & fNIRS DCM (Nonlinear)

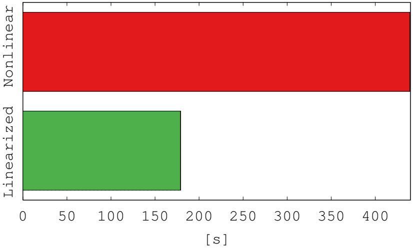


Numerical Results fMRI & fNIRS DCM ($\dim(x_r(t)) = \dim(\theta_r)$)

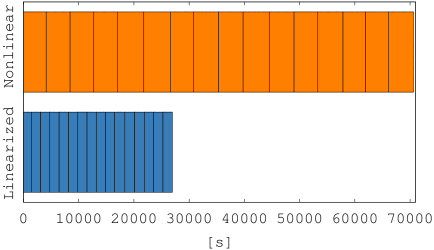


Numerical Results fMRI & fNIRS DCM (Offline Time)

Gramian-Based:

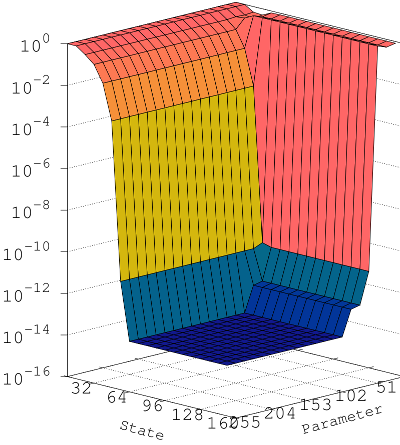


Optimization-Based:

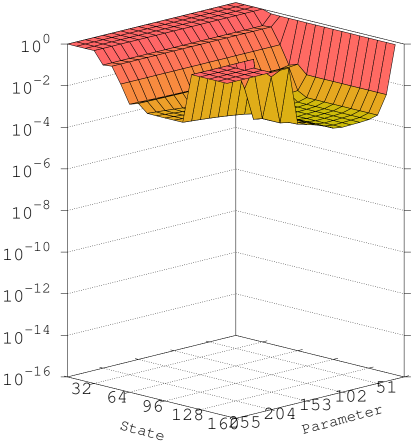


Numerical Results EEG & MEG DCM (Linearized)

Gramian-Based:

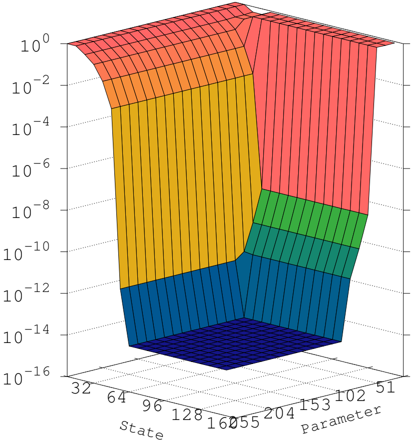


Optimization-Based:

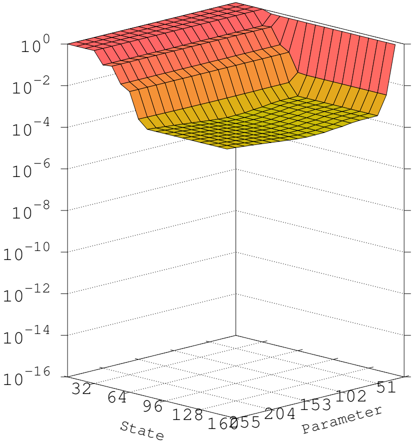


Numerical Results EEG & MEG DCM (Nonlinear)

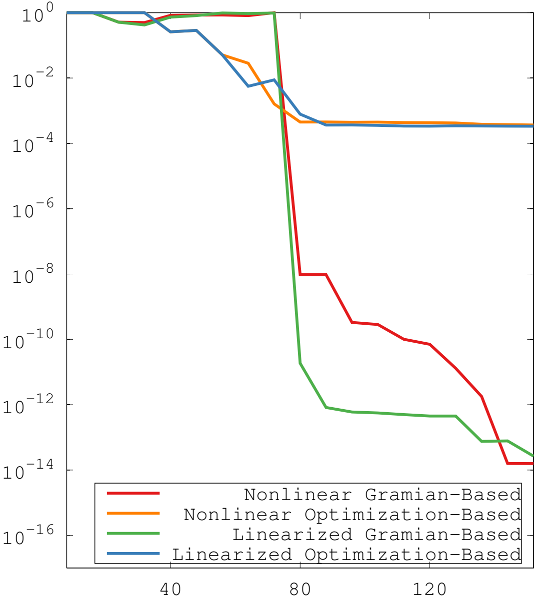
Gramian-Based:



Optimization-Based:

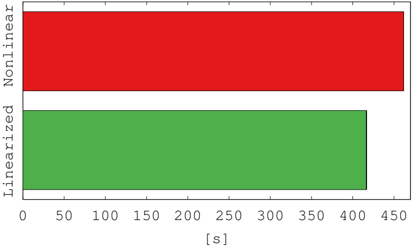


Numerical Results EEG & MEG DCM ($\dim(x_r(t)) = \dim(\theta_r)$)

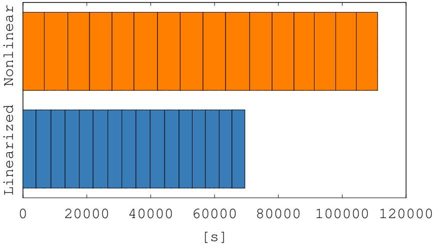


Numerical Results EEG & MEG DCM (Offline Time)

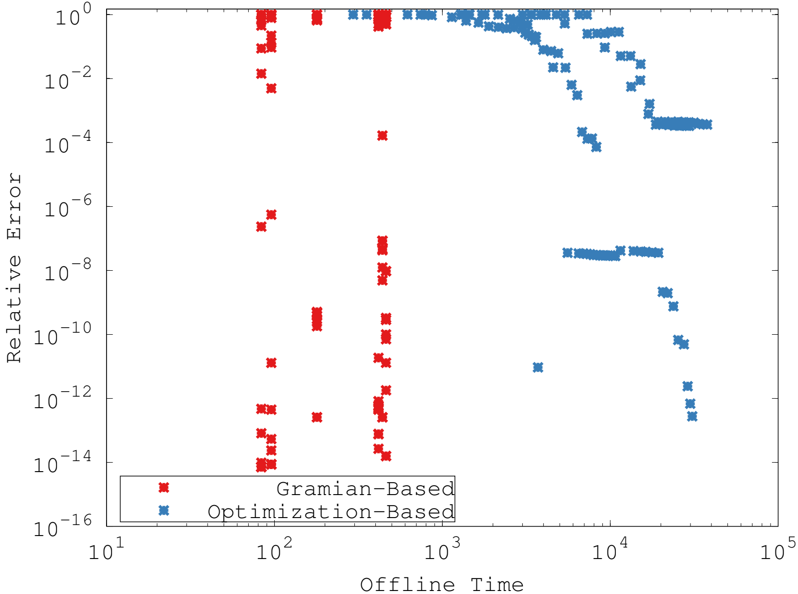
Gramian-Based:



Optimization-Based:



Efficiency



Relative Error
Offline Time
Gramian-Based
Optimization-Based
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}

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	$\| \cdot \|_{L_2 \otimes L_2}$	$\| \cdot \|_{L_2 \otimes L_\infty}$
Nonlinear Systems	✓	✓
Accuracy	☺	☺
Offline Time	☺	☹*

Outlook

- Hyper-Reduction (DMD)
- Gramian-Based:
  - Kernel Methods
  - Distributed Memory Parallelization
- Optimization-Based:
  - Derivative Information (AD | W_I)
  - L_1-Regularization

tl;dl

Summary:

- Combined State and Parameter Reduction
- Gramian-Based Method
- Optimization-Based Method
- Neuroscientific Application

wwwmath.uni-muenster.de/u/himpe

Thanks!